MEEGID

Volume 8, issue 4,
July 2008
ISSN 1567-1348

Infection, Genetics and Evolution

Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases

doi:10.1016/S1567-1348(08)00099-3

Infection, Genetics and Evolution

Volume 8 (2008)

Special Issue

8th International Meeting on Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases

Royal River Hotel, Bangkok, Thailand

30 November–2 December 2006

Edited by

Michel Tibayrenc

ELSEVIER

Vol. 8, issue 4, July 2008

CONTENTS

Infection, Genetics and Evolution

www.elsevier.com/locate/meegid

Abstracted/indexed in: Biological Abstracts, Chemical Abstracts, BIOBASE, EMBASE/Excerpta Medica, ISI, Photozoological Abstracts. Also covered in the abstract and citation database SCOPUS®. Full text available on ScienceDirect®.

Special Issue: 8th International Meeting on Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases

(Contents Continued on page I)

1567-1348(200807)8:4;1-M

Available online at

ScienceDirect

www.sciencedirect.com

(Contents continued from outside back cover)

Discussions

Abstracts

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 395–396
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Editorial

MEEGID VIII Bangkok, Thailand☆

The 8th session of the international congress, Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases, was held at the Royal River Hotel, Bangkok, Thailand, 30th November–2nd December 2006.

The MEEGID congresses were born in 1996 at the Centers for Disease Control (CDC) in Atlanta. My CDC counterpart, Altaf Lal (chief, malaria section), and I felt that molecular epidemiology was fast gaining in topicality at that time and fully deserved to be supported by a specific congress. Initially, it only covered pathogens, already original in its scope in that all pathogens were considered at MEEGID I, whether they be parasites, bacteria or viruses, of medical, veterinary or agronomical relevance. Soon came the vision that pathogen specialists needed to cooperate with entomologists and host geneticists to bring out a global picture of the entire infectious diseases transmission chain. This concept proved to be so successful in the MEEGIDs that followed that it became the backbone of the newborn journal *Infection, Genetics and Evolution* started in 2001. Another key concept of IGE and the MEEGID is the close interaction between basic science and biomedical research. This original formula has since then inspired several new journals.

Unfortunately, IGE was included in the Thompson Scientific/ISI database only in 2007, so that impact factor will be available only in 2009. However, an unofficial IF was computed by Elsevier experts using the same software as the ISI official IF: 3.554, which is excellent for such a young journal (Straub and Tibayrenc, 2007). Thanks to Elsevier's very active and professional support, our journal is on the way to fulfilling the role that I assigned to it from the very beginning: becoming the major tribune of all infectious disease specialists using evolutionary concepts and of all geneticists/evolutionists/bioinformaticians interested in the fascinating models offered by transmissible diseases.

MEEGID VIII proved again how successful the formula was. For three days, microbiologists, virologists, parasitologists, entomologists, evolutionists, geneticists, and bioinformaticians interacted in the dreamy and very professional environment of the Royal River Hotel, enjoying the legendary and smiling Thai hospitality.

According to its classical formula, MEEGID VIII specifically focused on those endemies that are the most relevant to the host country and its neighbors: malaria, dengue, avian flu, and cattle pathogens, among others.

No fewer than 26 different nationalities were present at the congress. A strong (37%) and enthusiastic minority of Thai scientists animated the sessions and illustrated the dynamism of Thai science in this topical research domain. Following its tradition, MEEGID VIII has catered particularly students, who were exempted from registration fees and were granted a specific symposium, chaired by students and with only student speakers. No less than 28% of MEEGID VIII participants were students.

A last tradition was fulfilled with the attribution of the MEEGID medals. The medal for the best oral communication was granted to Bruce Wilcox (University of Hawaii). Niyaz Ahmed (University of Hyderabad, India) was awarded the prize for the best oral communication by a scientist from a southern country. The prize for the best communication by a student was granted to Thierry Lefèvre (IRD, Montpellier, France). Last, the prize for the best poster communication was given to Panatda Saenkham (Chulabhorn Research Institute, Luksi, Bangkok).

It is my pleasure to open this special issue of *Infection, Genetics and Evolution*, which features the abstracts of all the papers delivered at the congress, together with 15 selected papers featuring a balanced sample of MEEGID and IGE topics (host genetics, microbiology and entomology).

The location of 2008's MEEGID IX will be Nairobi, Kenya. We will enjoy the hospitality and scientific expertise of our colleagues from ICIPE (African Insect Science for Food and Health; http://www.icipe.org/), the host institute. Apart from IRD (http://www.ird.fr), other organizers will include CNRS (http://www.cnrs.fr), KEMRI (http://www.kemri.org/), ILRI (http://www.ilri.org/), CDC Kenya (http://www.cdc.gov/malaria/cdcactivities/kenya.htm), the University of Nairobi (http://www.uonbi.ac.ke/), CIRAD (http://www.cirad.fr/fr/) and the French Embassy in Kenya (http://www.ambafrance-ke.org). According to tradition, MEEGID IX will emphasize problems that are more specifically relevant to Africa and the

☆ See information on the MEEGID IX congress at http://www.th.ird.fr//site_meegid/menu.htm.

1567-1348/$ – see front matter
doi:10.1016/j.meegid.2008.01.009

Indian Ocean: AIDS, malaria, tuberculosis, ebola, HIV, human African trypanosomiasis (sleeping sickness), chikungunya, as well as cattle and crop pathologies.

Welcome to Nairobi for another successful MEEGID!

References

Straub, B., Tibayrenc, M., 2007. Infection, genetics and evolution: a journal with a high impact but no impact factor (as yet). Infect. Genet. Evol. 7, 145–146.

Editor-in-Chief
Michel Tibayrenc*
IRD representative in Thailand,
IRD Representative Office, French Embassy,
29, Thanon Sathorn Tai, Bangkok 10120, Thailand

*Tel.: +66 2 627 2190; fax: +66 2 627 2194
E-mail address: Michel.Tibayrenc@ird.fr
URL:http://www.th.ird.fr

Available online 20 January 2008

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 397–405
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Review

The SARS-CoV nucleocapsid protein: A protein with multifarious activities

Milan Surjit, Sunil K. Lal *

Virology Group, International Centre for Genetic Engineering & Biotechnology, Aruna Asaf Ali Road, New Delhi 110067, India

Received 15 April 2007; received in revised form 10 July 2007; accepted 11 July 2007
Available online 20 July 2007

Abstract

Ever since the discovery of SARS-CoV in the year 2003, numerous researchers around the world have been working relentlessly to understand the biology of this virus. As in other coronaviruses, nucleocapsid (N) is one of the most crucial structural components of the SARS-CoV. Hence major attention has been focused on characterization of this protein. Independent studies conducted by several laboratories have elucidated significant insight into the primary function of this protein, which is to encapsidate the viral genome. In addition, many reports also suggest that this protein interferes with different cellular pathways, thus implying it to be a key regulatory component of the virus too. In the first part of this review, we will discuss these different properties of the N-protein in a consolidated manner. Further, this protein has also been proposed to be an efficient diagnostic tool and a candidate vaccine against the SARS-CoV. Hence, towards the end of this article, we will discuss some recent progress regarding the possible clinically relevant use of the N-protein.

Keywords: SARS; Nucleocapsid protein; Coronavirus; SARS diagnosis; SARS-CoV assembly; RNA virus

Contents

* Corresponding author at: Virology Group, ICGEB, P.O. Box 10504, Aruna Asaf Ali Road, New Delhi 110067, India. Tel.: +91 9818522900.
E-mail address: sunillal@icgeb.res.in (S.K. Lal).

1567-1348/$ – see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.meegid.2007.07.004

1. N-protein: structure and composition

The nucleocapsid (N) protein is encoded by the 9th ORF of SARS-CoV. The same ORF also codes for another unique accessory protein called ORF9b, though in a different reading frame, whose function is yet to be defined. The N-protein is a 46 kDa protein composed of 422 amino acids (Rota et al., 2003). Its N-terminal region consists mostly of positively charged amino acids, which is responsible for RNA binding. A lysine rich region is present between 373 and 390 amino acids at the C-terminus, which is predicted to be the nuclear localization signal. Besides that, a SR-rich motif is present in the middle region encompassing 177–207 amino acids. Biophysical studies done by Chang et al. (2006) have suggested that this protein is composed of two independent structural domains and a linker region. The first domain is present at the N-terminus, inside the putative RNA binding domain. The second domain consists of the C-terminal region that is capable of self-association. Between these two structural domains, there lies a highly disordered region, which serves as a linker. This region has been reported to interact with the membrane (M) protein and human cellular hnRNPA1 protein (Fang et al., 2006; Luo et al., 2005a,b). Besides, this region is also predicted to be a hot spot for phosphorylation. Hence, in summary, the N-protein can be classified into three distinct regions (Fig. 1), which may serve completely different functions during different stages of the viral life cycle. A similar mode of organization has been reported for other coronavirus nucleocapsid proteins.

2. Stability of the N-protein

In vitro thermodynamic studies done by C. Luo et al. (2004) and H. Luo et al. (2004) using purified recombinant N-protein have shown it to be stable between pH 7 and 10, with maximum conformational stability near pH 9. Further, it was observed to undergo irreversible thermal-induced denaturation. It starts to unfold at 35 °C and is completely denatured at 55 °C (Wang et al., 2004). However, chemicals such as urea or guanidium chloride-induced denaturation of the N-protein is a reversible process.

3. Post-translational modification

As in other coronavirus N-proteins, SARS-CoV N-protein has been predicted and later experimentally proven to undergo various post-translational modifications such as acetylation, phosphorylation, and sumoylation.

Acetylation is the first modification of the N-protein to be experimentally proven. By mass spectrometric analysis of convalescent sera from several SARS patients, it has been shown that the N-terminal methionine of N is removed as well as all other methionines are oxidised and the resulting N-terminal serine is acetylated. However, the functional relevance of this modification, if any, remains to be elucidated (Krokhin et al., 2003).

Another unique modification of the N-protein is its ability to get sumoylated. Studies done by Y.H. Li et al. (2005) and F.Q. Li et al. (2005) have clearly established that heterologously expressed N in mammalian cells is sumoylated. Using a site-directed mutagenesis approach, the sumoylation motif has been mapped to the lysine residue 62, which is present in a putative sumo-modification domain ($GK^{62}EE$). Their data further suggests that sumoylation may play a key role in modulating homo-oligomerization, nucleolar translocation and cell-cycle deregulatory property of the N-protein. Further experimental support regarding sumoylation of N-protein came from another independent study carried out by Fan et al. (2006) wherein they have demonstrated an association between the N-protein and Hubc9, which is a ubiquitin conjugating enzyme of the sumoylation system. They have also mapped the interaction domain to the SR-rich motif, which is in agreement with the earlier report. However, they failed to detect the involvement of the GKEE motif in mediating this interaction (Fan et al., 2006).

Initially, the SARS-CoV N-protein was predicted to be heavily phosphorylated. Later on, from results obtained in our laboratory as well as by other researchers, it is now clear that the N-protein is a substrate of multiple cellular kinases. First experimental evidence for the phosphorylation status of the N-protein came from the study done by Zakhartchouk et al. (2005) wherein by ^{32}P orthophosphate labelling, they were able to observe phosphorylation of adenovirus vector expressed N-protein in 293T cells. Further studies done in our laboratory clearly confirmed this observation. Majority of the N-protein was found to be phosphorylated at its serine residues (although the involvement of threonine and tyrosine residues could not be detected, they may be occurring in vivo). In addition, using a variety of biochemical assays, it was proved that, at least in vitro, the N-protein could get phosphorylated by mitogen activated protein kinase (MAP kinase), cyclin-dependent kinase (CDK), glycogen synthase kinase 3 (GSK3), and casein kinase 2 (CK2). Also, this data provides preliminary indication regarding phosphorylation-dependent nucleo-cytoplasmic shuttling of the N-protein (Surjit et al., 2005). Whether these events actually bear any functional significance in vivo, remains yet to be established.

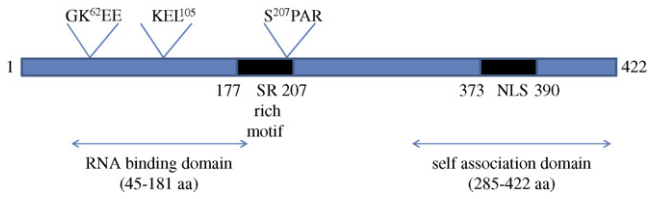


Fig. 1. Structure of the SARS-CoV nucleocapsid protein. A schematic diagram showing various different domains identified to-date. The numbers 1–422 correspond to the length in amino acids of the N gene. GKEE represents the sumoylation motif (lysine residue). KEL is the RXL motif, responsible for binding with cyclin D and SPAR is the motif that gets phosphorylated by cyclin–CDK complex (serine residue).

4. Localization of the N-protein

In contrast to the N-protein of many other coronaviruses, the SARS-CoV N-protein is predominantly distributed in the cytoplasm, when expressed heterologously or in infected cells (Surjit et al., 2005; You et al., 2005; Rowland et al., 2005). In

infected cells, a few cells exhibited nucleolar localization (You et al., 2005). As reported by You et al. (2005) the N-protein contains pat4, pat7, and bipartite-type nuclear localization signals. It has also been predicted to possess a potential CRM-1 dependent nuclear export signal. However, no clear experimental evidence could be obtained regarding the involvement of these signature sequences in regulating the localization of the N-protein. Interestingly, studies done in our laboratory revealed that the majority of N-protein localized to the nucleus in serum starved cells. This phenomenon could be reproducibly observed both in biochemical fractionation as well as immunofluorescence studies. In addition, treatment of cells with specific inhibitors of different cellular kinases such as CK2 inhibitor and CDK inhibitor, resulted in retention of a fraction of the N-protein in the nucleus, whereas GSK3 and MAPK inhibitor had very little effect. Further, N was found to be efficiently phosphorylated by the cyclin–CDK complex, which is known to be active only in the nucleus. The N-protein was also found to associate with 14-3-3 protein in a phospho-specific manner and inhibition of the 14-3-3θ protein level by siRNA resulted in nuclear accumulation of the N-protein. Although these experiments are too preliminary to conclusively provide any answer regarding the intra-cellular localization of N-protein, nevertheless they do provide substantial clues regarding the physical presence of the N-protein in the nucleus, under certain circumstances, which may be a very dynamic phenomenon. Another study done by Timani et al. (2005) using different deletion mutants of the N-protein fused to EGFP, showed that the N-terminal of N, which contains the NLS 1 (aa 38–44) localizes to the nucleus, whereas the C-terminal region containing both NLS 2 (aa 257–265) and NLS 3 (aa 369–390) localizes to the cytoplasm and nucleolus. Using a combination of different deletion mutants, they concluded that the N-protein may act as a shuttle protein between cytoplasm–nucleus and nucleolus. Taken together, all these results further suggest that the N-protein *per se* has the physical ability to localize to the nucleus. Whether this localization is regulated through phosphorylation mediated activation of a potential NLS or piggy-backing by association with another cellular nuclear protein or through any other mechanism remains to be established.

5. Genome encapsidation: primary function of a viral capsid protein

Being the capsid protein, the primary function of the N-protein is to package the genomic RNA in a protective covering. In order to achieve this structure, the N-protein must be equipped with two different characteristic properties such as: (i) being able to recognize the genomic RNA and associate with it; (ii) self-associate into an oligomer to form the capsid. The N-protein of SARS-CoV has been experimentally proven to possess these properties in vitro, as discussed below.

5.1. Recognition and binding with the genomic RNA

First experimental evidence regarding the RNA binding property of the N-protein came from the work of Huang et al. (2004), wherein by NMR studies, they proved the ability of N-terminal domain to associate with several viral 3′ untranslated RNA sequence. Additionally, a recent report by Chen et al. (2007) proved the presence of another RNA binding domain at the C-terminal region (residues 248–365) of the N-protein, which is proposed to be a stronger interaction than that at the N-terminus. Based on the structural analysis of the RNA–protein interaction, they have further suggested that the genomic RNA is packaged in a helical manner by the N-protein. The RNA binding motif of N-protein has been mapped to 363–382 amino acid residues by Luo et al. (2006).

Perhaps, the most convincing proof till date regarding the ability of the N-protein to package the genomic RNA came from the work of Hsieh et al. (2005). They have established a system to produce SARS-CoV VLPs by cotransfection of spike, membrane, envelope, and nucleocapsid cDNAs into Vero E6 cells. While testing the packaging of an RNA bearing GFP fused to SARS-CoV packaging signal into this particle, they observed that the presence of the N-protein is an absolute requirement. However, the N-protein was not essential for the assembly of the empty particle *per se*. Further, by performing a filter binding assay using recombinant N-protein, they were able to identify two independent RNA binding domains in the N-protein; one at the N-terminus (1–235 aa) and the other at the C-terminus (236–384 aa). These results are in agreement with previous findings and further suggest that these two regions may be functional in vivo. Future experiments using a model infection system will confirm these observations.

5.2. Formation of the capsid

One of the most crucial properties required by the N-protein for genome encapsidation is its ability to self-associate. Therefore, many laboratories have focused on characterizing this phenomenon, with an eye to develop possible interference strategies that may help in limiting virus propagation.

Initial studies done in our laboratory using a yeast two-hybrid assay revealed that N-protein is able to self-associate through its C-terminal 209 amino acid residues (Surjit et al., 2004a,b,c). A parallel study done by He et al. (2004) using the mammalian two-hybrid system and sucrose gradient fractionation also proved the ability of the N-protein to self-associate to form an oligomer. They further mapped the interaction region to 184–196 amino acid residues, encompassing the SR-rich motif. However, there were some discrepancies regarding the interaction domain mapped in these two studies. Later on, extensive biophysical and biochemical analysis done by Chen lab (Yu et al., 2005, 2006) and Jiang lab (Luo et al., 2005a,b, 2006) have enriched our understanding of the oligomerisation process of the N-protein. In summary, the SR-rich motif does possess binding affinity, but this is specific for the central region (211–290 aa) of another molecule of N-protein, instead of the SR-rich motif itself. The C-terminal region (283–422 aa) possess binding affinity for itself and to associate into a dimer, trimer, tetramer or hexamer, in a concentration-dependent manner. The essential sequence for oligomerisation of the N-protein was identified to be residues 343–402. Interestingly, this

region also encompasses the RNA binding motif of the N-protein, which prompts us to speculate that there might be mutual interplay between RNA binding and oligomerisation activities of the N-protein. Further, the oligomerisation was observed to be independent of electrostatic interactions and addition of single strand DNA to the reaction mixture containing tetramers of the N-protein promoted oligomerisation. Thus, it has been proposed that once the tetramer is formed by protein–protein interaction between nucleocapsid molecules, binding with genomic RNA prompts further assembly of the complete nucleocapsid structure.

6. Perturbation of host cellular process by the N-protein

Besides being the capsid protein of the virus, the N-protein of many coronaviruses is known to double up as a regulatory protein. The N-protein of the SARS-CoV too has been shown to modulate the host cellular machinery in vitro, thereby indicating its possible regulatory role during its viral life cycle. Some of the major cellular processes perturbed by heterologous expression of the N-protein are discussed below.

6.1. Deregulation of host cell cycle

Two groups have reported the ability of the N-protein to interfere with the host cell cycle in vitro. Work done by Y.H. Li et al. (2005) and F.Q. Li et al. (2005) proved that mutation of the sumoylation motif in the N-protein leads to cell cycle arrest. Work done in our laboratory has shown the inhibition of S phase progress in cells expressing the N-protein (Surjit et al., 2006). Further, we have observed down-regulation of S phase specific gene products like cyclin E and CDK2 in SARS-CoV infected cell lysate, which suggest that the observed phenomenon may be relevant in vivo. In an attempt to further characterize the mechanism of cell cycle blockage induced by the N-protein, several biochemical and mutational analysis were carried out. The results thus obtained demonstrated that the N-protein directly inhibits the activity of the cyclin–CDK complex, resulting in hypophosphorylation of retinoblastoma protein with a concomitant down-regulation of E2F1-mediated transactivation. Analysis of RXL and CDK phosphorylation mutant N-protein identified the mechanism of inhibition of CDK4 and CDK2 activity to be different. Whereas the N-protein could directly bind to cyclin D and inhibit the activity of CDK4–cyclin D complex; inhibition of CDK2 activity appeared to be achieved in two different ways: indirectly by down-regulation of protein levels of CDK2, cyclin E, and cyclin A, and by direct binding of N-protein to CDK2–cyclin complex. Nevertheless, the mechanism of cell cycle deregulation in vivo, if any, remains to be understood.

6.2. Inhibition of interferon production

Production of interferon is one of the primary host defense mechanism. However, SARS-CoV infection does not result in IFN production. Nevertheless, pretreatment of cells with IFN blocks SARS-CoV infection (Spiegel et al., 2005; Zheng et al., 2004). Based on this observation, Palese lab has studied the IFN inhibitory property of different SARS-CoV proteins, which revealed that ORF3b, ORF6 as well as the N-protein have the ability to independently inhibit IFN production through different mechanisms. The N-protein was found to inhibit the activity of IRF3 and NFkB in host cells, resulting in inhibition of IFN synthesis. IRF3 activity was also blocked by 3b and ORF6 proteins, but inhibition of NFkB activity was a property unique to the N-protein. In addition, 3b and ORF6 proteins were able to block STAT1 activity through different mechanisms (Kopecky-Bromberg et al., 2007). All these data suggest that SARS-CoV may employ multiple factors to check the activity of host immune system and N-protein may be one of the major partners in this process. It may be possible that these different factors act independently during different stages of viral life cycle. In that case, regulatory activity of the N-protein will be as indispensible as its structural activity.

6.3. Up-regulation of COX2 production

Another major proinflammatory factor-induced during viral infection is the cyclo oxygenase-2 protein. Using 293T cells expressing the N-protein, Yan et al. (2006) have shown that expression of the N-protein leads to upregulation of COX2 protein production in a transcriptional manner. They have further demonstrated that the N-protein directly binds to the NFkB response element present in the COX2 promoter through a 68 aa residue binding domain (136–204 aa) and activates its transcription.

Although the N-protein is known to associate with stretches of nucleic acids, till date there is no other documentation or prediction of its sequence specific DNA binding activity (as a transcription factor). In such a scenario, the above observation, if reproducible in vivo, may really be a unique property of the N-protein and may further add to the established regulatory functions of the N-protein.

6.4. Up-regulation of AP1 activity

Exogenously expressed N-protein has been reported to enhance the DNA binding activity of c-fos, ATF-2, CREB-1, and fos B in an ELISA-based assay, thus suggesting an increase in AP1 activity in these cells (He et al., 2003). Mechanistic details and functional significance of this phenomenon remains to be elucidated.

6.5. Induction of apoptosis in serum starved monkey kidney cells

Earlier work done in our laboratory has shown that N-protein, when expressed in COS-1 monkey kidney cells, induces apoptosis in the absence of growth factors. Attempts to understand the mechanism of programmed cell death revealed that the N-protein down modulated the activity of pro-survival factors such as extracellular regulated kinase, Akt, and bcl 2, and upregulated the activity of pro-apoptotic factors like

caspase-3 and caspase-7. This phenomenon was also associated with reorganization of the actin cytoskeleton (Surjit et al., 2004a,b,c). However, this phenomenon was not observed in another cell line of epithelial lineage (huh7). Recently, Zhang et al. (2007) have reported that serum starvation apoptosis of N expressing COS-1 cells involves activation of mitochondrial pathway. It remains to be studied whether this phenomenon is actually recapitulated in vivo.

6.6. Association with host cell proteins

C. Luo et al. (2004) and H. Luo et al. (2004) have reported the interaction between hnRNPA1 and N-protein by using a variety of biochemical and genetic assays. The interaction was found to be mediated through the middle region (161–210 aa) of N-protein. If relevant in vivo, this interaction may play a significant role in regulation of the viral RNA synthesis.

Another interesting study done by C. Luo et al. (2004) and H. Luo et al. (2004) have reported association between the N-protein and human cyclophylin A. By SPR analysis they have shown it to be a high affinity interaction. Although the significance of this interaction is not known in vivo, they have proposed that this interaction might be crucial for viral infection. Notable is the fact that HIV-1 gag also binds with human cyclophylin A and this interaction is crucial for HIV infection (Gamble et al., 1996).

In summary, although several regulatory roles have been proposed for the SARS-CoV N-protein using a variety of in vitro experimental systems, no clear evidence exists for their occurrence in vivo. In the absence of a suitable in vivo experimental system, all these functions remain speculative.

7. N-protein: an efficient diagnostic tool

One of the most essential steps to limit the outbreak of any infectious disease is the ability to diagnose the causative agent, at the earliest possible time, which can be achieved by detecting some of the markers that are specifically expressed by the pathogen or by identifying some of the host factors that are specifically produced during infection. N-protein, being one of the predominantly expressed proteins at the early stage of SARS-CoV infection, against which a strong antibody response is initiated by the host; has been proposed to be an attractive diagnostic tool.

In serum of SARS-CoV patients, the N-protein has been detected as early as day 1 of infection by ELISA using monoclonal antibodies against it (Che et al., 2004). Further, a comparative study to detect SARS-CoV specific IgG, SARS-CoV RNA, and the N-protein during early stages of infection has demonstrated that detection efficiency of the N-protein is significantly higher than the other two markers (Y.H. Li et al., 2005; F.Q. Li et al., 2005).

Researchers have been mainly focussing on two different strategies by which nucleocapsid can be used as a diagnostic tool (i) development of efficient monoclonal antibodies against the N-protein, (ii) production of recombinantly expressed, highly purified N-protein for detection of N-specific antibody in the host.

Using a phage display approach, Flego et al. (2005) have identified human antibody fragments that recognise distinct epitopes of the N-protein. These may help to develop efficient reagents to detect N-protein in the infected host. Further, several laboratories have been trying to develop efficient monoclonal antibodies against the major immunodominant epitopes of the N-protein, that can be used in ELISA to detect SARS-CoV at an early stage of infection (Shang et al., 2005; Liu et al., 2003; He et al., 2005; Woo et al., 2005). In another interesting study, Liu et al. (2005) have developed an immunofluorescence assay using antirabbit N-antibody that can specifically detect N-protein from throat wash sample of SARS-CoV patients at day 2 of illness.

Several other workers have focused on economical production of highly purified recombinant N-protein using a variety of heterologous expression systems that can be used in ELISA to detect N-specific antibody in the patient sample. N-protein has been produced in abundant quantity using codon optimised gene in *Escherichia coli* (Das and Suresh, 2006). Saijo et al. (2005) have successfully expressed recombinant N-protein using a baculovirus expression system, which was found to be 92% efficient in neutralizing antibody assay. In another study, Liu et al. (2004), have expressed full length N-protein using yeast expression system. However diagnostic use of recombinant N-protein has been a problematic issue because of several reasons as discussed below.

Bacterially expressed N-protein has been reported to produce false sero-positivity owing to interference of bacterially derived antigens (Leung et al., 2006; Yip et al., 2007). In addition, several studies have shown cross-reactivity between full-length N-protein of SARS and polyclonal antisera of group 1 animal coronaviruses, which may lead to faulty detection (Sun and Meng, 2004). Another study done by Woo et al., have also reported crossreactivity of full-length recombinant N-protein with antisera of H CoV-OC43 and H CoV-229E infected patients, thus giving false positive results. They were able to minimise this false positivity by further verifying the ELISA results with Western blot assay using recombinant N and spike protein of SARS-CoV (Woo et al., 2004).

Later on, studies done by Qiu et al. (2005) and Bussmann et al. (2006), showed that recombinantly expressed C-terminal of the N-protein acts more specifically in detecting SARS-CoV specific antisera in comparision to full-length N-protein. It is noteworthy that this region is predicted to encompass major antigenic sites of the N-protein.

Also, several reports have been published dealing with the detection of N specific IgM by ELISA or indirect immunofluorescent assay (Chang et al., 2004; Hsueh et al., 2004; Woo et al., 2004). However, in these studies, IgM antibodies became detectable later than IgG antibodies, which is in contrast to the phenomena observed in most other pathogens.

A recent report published by Yu et al. (2007), have attempted to solve this problem by using a truncated N-protein (122–422 aa) as an antigen in IgM ELISA. They found the IgM response to appear 3 days before detection of IgG response, which is in

agreement with the results obtained from other known pathogens. Further, their results showed 100% specificity and sensitivity of the truncated protein in detecting N-specific IgM from patients with laboratory confirmed SARS cases in comparison to healthy volunteers. The authors have suggested that the IgM capture ELISA using this truncated N-protein may be more effective in serodiagnosis of SARS-CoV at an earlier time.

In another interesting report, Woo et al. (2005) have carried out comparative studies to evaluate the relative diagnostic efficacy of recombinantly expressed N- and S-proteins. They observed sensitivity of recombinant N-IgG ELISA to be significantly higher than that of recombinant S-IgG ELISA. The reverse was true in case of IgM ELISA using recombinant N- and S-proteins. Based on this data, they have suggested the practice of an ELISA for detection of IgM against both S- and N-protein instead of N alone (Woo et al., 2005).

Taken together, all these data does support the notion that the N-protein may be used as an efficient diagnostic tool for detection of SARS-CoV infection. Nevertheless, production scale-up and further validation of specificity using patient samples will determine the possible clinical use of these reagents.

8. N-protein: a suitable vaccine candidate

One of the most clinically relevant uses of the N-protein can be its use as a protective vaccine against SARS-CoV infection. N-protein is one of the major antigens of the SARS-CoV. Also, N-protein analysed from different patient samples shows least variation in the gene sequence (Tong et al., 2004), therefore indicating it to be a stable protein, which is a primary requirement for an efficient vaccine candidate.

Earlier studies carried in Collins lab, Rao lab, and Li lab have clearly shown that anti-serum to the N-protein does not contain neutralizing antibodies against SARS-CoV (Buchholz et al., 2004; Pang et al., 2004, and Liang et al., 2005). This may be attributed to the localization of N-protein inside the viral envelope, which will not be accessible to the antibody during infection. It is noteworthy that the most effective SARS-CoV structural protein that can induce neutralizing antibody production is the S-protein (Buchholz et al., 2004). The S antibody could block viral infection with 100% efficiency. On the other hand, although unable to induce humoral immunity, expression of N-protein-induced significant cytotoxic T lymphocyte (CTL) response (Buchholz et al., 2004; Gao et al., 2003; Zhu et al., 2004). Induction of N specific CTLs will help limit the infection by lysing virus infected cells. This will also limit the spread of virus. Thus, N-based vaccines may further augment the protection efficiency when co-administered with S-based vaccine. Several laboratories have been exploring various strategies to evaluate the potential of N-protein as a vaccine candidate.

In an elegant work done by Kim et al. (2004), calreticulin fused N-protein expressing vaccinia virus has been shown to generate potent N-specific humoral and T-cell immune responses in mice. As reported by the authors, fusion with calreticulin specifically enhanced the efficiency and significantly reduced the titre of challenging vector (vaccinia virus). The authors have proposed that N-protein may be the logical choice as a target antigen in the event of S antibody dependent enhancement (ADE) of infection. However, ADE phenomenon has not been observed during spike-mediated vaccination (Buchholz et al., 2004). Another study done by Wang et al. (2005) has attempted to use plasmid DNA expressing S-, M-, and N-proteins as an efficient vaccine candidate. Although they report the production of some B-cell and T-cell responses against N-protein, however stringent immune response was obtained for the S- and M-proteins, thus scaling down the choice of N-protein as a suitable candidate vaccine (Wang et al., 2005). A similar plasmid mediated vaccination approach has also been reported by Zhao et al. (2004) wherein they have immunised mice with the DNA construct (pCI vector) expressing the N-protein. They too have reported the generation of a robust B-cell and T-cell immune response in animals. Another group of workers have also reported successful use of the N-protein as a DNA vaccine. They immunised mice by intra-mucosal injection of the N-protein expressing plasmid vector and were able to obtain specific humoral and T-cell responses (Zhu et al., 2004).

The N-protein has also been reported to be of potential interest as a peptide-based vaccine. A systematic study done by Liu et al. (2006) has revealed the immunodominant epitopes of the N-protein which could efficiently stimulate immune response. They have also deduced some conserved immunodominant epitopes in mouse, monkey, and humans, which may help in design of the vaccine.

A recent report published by Gao's laboratory provides further evidence regarding the efficiency of an N-based vaccine (Zhao et al., 2007). By using overlapping synthetic peptides spanning the N-protein, they have identified dominant helper T-cell epitopes in the nucleocapsid protein of SARS-CoV. Immunization of mice with peptides emcompassing these dominant TH-cell epitopes resulted in strong cellular immunity in vivo. Priming with the helper peptides significantly accelerated the immune response induced by the N-protein. Further, by fusing with a conserved neutralizing epitope from the spike protein of SARS-CoV, two of the TH-cell epitope bearing peptides assisted in the production of higher titre neutralizing antibodies in vivo, in comparision to spike epitope alone or its mixture with TH epitope of N. Thus, it is practically possible to generate a better immune response by using a fusion of N- and S-protein. However, the TH epitopes identified in their report is specific to mouse. Therefore, TH epitopes identified in that study will not be useful for human. Nevertheless, their data provides useful information for the design of peptide-based anti-SARS-CoV vaccines.

Another interesting study conducted by Pei et al. (2005) reports the possible use of the N-protein as a mucosal vaccine candidate. They expressed the N-protein in Lactobacillus lactis, which is a food grade bacteria, and challenged the mice either orally or intra-mucosally. As preliminary evidence, they were able to observe significant N-specific IgG in the sera of orally challenged animals.

9. Future perspective

It is a significant achievement for the research community that within a short span of time, we have been able to obtain more or less a clear understanding regarding the structural and functional properties of the N-protein. However, it is a fact worth mentioning that all the studies done here were performed with in vitro experiments, using recombinantly expressed N-protein, in isolation. So at present, all we can conclude is that, the N-protein *per se* has the physical ability to perform the above described functions, in other words N-protein does bear the necessary signature sequence or motifs or conformation to perform these functions under suitable circumstances. Whether a similar event is recapitulated in vivo during viral infection, will be dependent on several criteria: (i) net effect of other viral factors on the activity of N-protein, (ii) net translation and turn over rate of N-protein, (iii) a conducive intra-cellular milieu, and (iv) net modulation of an already skewed cellular pathway by other viral factors. Hence, it will be interesting to re-evaluate the properties of N-protein in a SARS-CoV infection model. However, owing to the limited user friendliness and accessibility of an infection system, probably we still have to resort to in vitro systems for further analysis of the characteristics of N-protein. One of the better experimental system has already been established by the Chang lab (Hsieh et al., 2005), wherein all the structural proteins were co-expressed to form VLP in 293T cells. If this system can be further improved to optimise the rate of synthesis of these different proteins to a near in vivo level, it will at least enable us to study the net effect of the N-protein with respect to other viral proteins. Further establishment of a replicon system may also be helpful. In addition, some of the interesting preliminary observations reported by several laboratories need to be elaborately analysed. To begin with, the reported interaction of the N-protein with genomic RNA packaging signal needs to be further characterised and mapped. Since the oligomerisation domain and the RNA binding regions of the N-protein overlap with each other, the suggested possibility of regulated genome incorporation and capsid assembly should be further characterised with the aid of a replicon system or a particle assembly system. In addition, the reported ability of the N-protein to modulate different cellular pathways should be further characterised in the particle assembly system or at least in the presence of other viral accessory proteins.

The most unique and significant property of the N-protein revealed by preliminary studies is its ability to act as sequence specific DNA binding factor. It has been shown to bind NFkB response element of COX2 promoter and enhance COX2 gene expression. This activity may be further empowering the N-protein to manipulate the entire gene expression programme of the infected cell. Therefore, studies should be initiated to elaborately analyse this phenomenon. It seems to deserve so much attention because another study done by Palese lab has proved the ability of the N-protein to inhibit NFkB activity, which results in inhibition of IFN synthesis. Further, Liao et al. (2005) have reported the activation of NFkB by N-protein in Vero E6 cells and He et al. (2005) failed to detect any change in NFkB activity in the same cells. Therefore it needs to be clarified whether N enhances NFkB activity and if yes; whether upregulation of COX2 transcription by direct DNA binding is a property specific to that promoter or it is a global phenomenon. In such a scenario, there may be complicated crosstalk between the ability of N-protein to deregulate the expression of COX2 and IFN in infected cells.

Lastly, the N-protein is known to be the most abundantly expressed protein of the SARS-CoV. Therefore, any information generated from the analysis of this protein, whether in vivo or ex vivo, will definitely help to increase our understanding of the biology of SARS-CoV and may someday help design better protective tools against it.

Acknowledgements

The authors wish to thank Ms. Alisha Lal for helping out in typing and formatting this review. We apologize to all those colleagues whose work we might have missed to cite in this article.

References

Buchholz, U.J., Bukreyev, A., Yang, L., Lamirande, E.W., Murphy, B.R., Subbarao, K., Collins, P.L., 2004. Contributions of the structural proteins of severe acute respiratory syndrome coronavirus to protective immunity. Proc. Natl. Acad. Sci. U.S.A. 101, 9804–9809.

Bussmann, B.M., Reiche, S., Jacob, L.H., Braun, J.M., Jassoy, C., 2006. Antigenic and cellular localisation analysis of the severe acute respiratory syndrome coronavirus nucleocapsid protein using monoclonal antibodies. Virus Res. 122, 119–126.

Chang, W.T., Kao, C.L., Chung, M.Y., Chen, S.C., et al., 2004. SARS exposure and emergency department workers. Emerg. Infect. Dis. 10, 1117–1119.

Chang, C.K., Sue, S.C., Yu, T.H., et al., 2006. Modular organization of SARS coronavirus nucleocapsid protein. J. Biomed. Sci. 13, 59–72.

Che, X.Y., Hao, W., Wang, Y., et al., 2004. Nucleocapsid protein as early diagnostic marker for SARS. Emerg. Infect. Dis. 10, 1947–1949.

Chen, C.Y., Chang, C.K., Chang, Y.W., et al., 2007. Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. J. Mol. Biol. 368, 1075–1086.

Das, D., Suresh, M.R., 2006. Copious production of SARS-CoV nucleocapsid protein employing codon optimized synthetic gene. J. Virol. Methods 137, 343–346.

Fan, Z., Zhuo, Y., Tan, X., et al., 2006. SARS-CoV nucleocapsid protein binds to hUbc9, a ubiquitin conjugating enzyme of the sumoylation system. J. Med. Virol. 78, 1365–1373.

Fang, X., Ye, L.B., Zhang, Y., et al., 2006. Nucleocapsid amino acids 211 to 254, in particular, tetrad glutamines, are essential for the interaction between the nucleocapsid and membrane proteins of SARS-associated coronavirus. J. Microbiol. 44, 577–580.

Flego, M., Di Bonito, P., Ascione, et al., 2005. Generation of human antibody fragments recognizing distinct epitopes of the nucleocapsid (N) SARS-CoV protein using a phage display approach. BMC Infect. Dis. 5, 73.

Gamble, T.R., Vajdos, F.F., Yoo, S., et al., 1996. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell 87, 1285–1294.

Gao, W., Tamin, A., Soloff, A., D'Aiuto, L., Nwanegbo, E., Robbins, P.D., Bellini, W.J., Barratt-Boyes, S., Gambotto, A., 2003. Effects of a SARS-associated coronavirus vaccine in monkeys. Lancet 362, 1895–1896.

He, R., Leeson, A., Andonov, A., et al., 2003. Activation of AP-1 signal transduction pathway by SARS coronavirus nucleocapsid protein. Biochem. Biophys. Res. Commun. 311, 870–876.

He, R., Dobie, F., Ballantine, M., et al., 2004. Analysis of multimerization of the SARS coronavirus nucleocapsid protein. Biochem. Biophys. Res. Commun. 316, 476–483.

He, Q., Du, Q., Lau, S., et al., 2005. Characterization of monoclonal antibody against SARS coronavirus nucleocapsid antigen and development of an antigen capture ELISA. J. Virol. Methods 127, 46–53.

Hsieh, P.K., Chang, S.C., Huang, C.C., et al., 2005. Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. J. Virol. 79, 13848–13855.

Hsueh, P.R., Huang, L.M., Chen, P.J., et al., 2004. Chronological evolution of IgM, IgA, IgG and neutralisation antibodies after infection with SARS-associated coronavirus. Clin. Microbiol. Infect. 10, 1062–1066.

Huang, Q., Yu, L., Petros, A.M., et al., 2004. Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. Biochemistry 43, 6059–6063.

Kim, T.W., Lee, J.H., Hung, C.F., et al., 2004. Generation and characterization of DNA vaccines targeting the nucleocapsid protein of severe acute respiratory syndrome coronavirus. J. Virol. 78, 4638–4645.

Kopecky-Bromberg, S.A., Martinez-Sobrido, L., Frieman, M., et al., 2007. Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. J. Virol. 81, 548–557.

Krokhin, O., Li, Y., Andonov, A., et al., 2003. Mass spectrometric characterization of proteins from the sars virus: a preliminary report. Mol. Cell Proteomics 2, 346–356.

Leung, D.T., van Maren, W.W., Chan, F.K., Chan, W.S., Lo, A.W., Ma, C.H., Tam, F.C., To, K.F., Chan, P.K., Sung, J.J., Lim, P.L., 2006. Extremely low exposure of a community to severe acute respiratory syndrome coronavirus: false seropositivity due to use of bacterially derived antigens. J. Virol. 80, 8920–8928.

Li, F.Q., Xiao, H., Tam, J.P., et al., 2005. Sumoylation of the nucleocapsid protein of severe acute respiratory syndrome coronavirus. FEBS Lett. 579, 2387–2396.

Li, Y.H., Li, J., Liu, X.E., et al., 2005. Detection of the nucleocapsid protein of severe acute respiratory syndrome coronavirus in serum: comparison with results of other viral markers. J. Virol. Methods 130, 45–50.

Liang, M.F., Du, R.L., Liu, J.Z., et al., 2005. SARS patients-derived human recombinant antibodies to S and M proteins efficiently neutralize SARS-coronavirus infectivity. Biomed. Environ. Sci. 18, 363–374.

Liao, Q.J., Ye, L.B., Timani, K.A., et al., 2005. Activation of NF-kappaB by the full-length nucleocapsid protein of the SARS coronavirus. Acta Biochim. Biophys. Sin. (Shanghai) 37, 607–612.

Liu, G., Hu, S., Hu, Y., et al., 2003. The C-terminal portion of the nucleocapsid protein demonstrates SARS-CoV antigenicity. Genom. Proteom. Bioinform. 1, 193–197.

Liu, R.S., Yang, K.Y., Lin, J., et al., 2004. High-yield expression of recombinant SARS coronavirus nucleocapsid protein in methylotrophic yeast Pichia pastoris. World J. Gastroenterol. 10, 3602–3607.

Liu, I.J., Chen, P.J., Yeh, S.H., et al., 2005. Immunofluorescence assay for detection of the nucleocapsid antigen of the severe acute respiratory syndrome (SARS)-associated coronavirus in cells derived from throat wash samples of patients with SARS. J. Clin. Microbiol. 43, 2444–2448.

Liu, S.J., Leng, C.H., Lien, S.P., et al., 2006. Immunological characterizations of the nucleocapsid protein based SARS vaccine candidates. Vaccine 24, 3100–3108.

Luo, C., Luo, H., Zheng, S., et al., 2004. Nucleocapsid protein of SARS coronavirus tightly binds to human cyclophilin A. Biochem. Biophys. Res. Commun. 321, 557–565.

Luo, H., Ye, F., Sun, T.Y., et al., 2004. In vitro biochemical and thermodynamic characterization of nucleocapsid protein of SARS. Biophys. Chem. 112, 15–25.

Luo, H., Ye, F., Chen, K., et al., 2005a. SR-rich motif plays a pivotal role in recombinant SARS coronavirus nucleocapsid protein multimerization. Biochemistry 44, 15351–15358.

Luo, H., Chen, Q., Chen, J., et al., 2005b. The nucleocapsid protein of SARS coronavirus has a high binding affinity to the human cellular heterogeneous nuclear ribonucleoprotein A1. FEBS Lett. 579, 2623–2628.

Luo, H., Chen, J., Chen, K., et al., 2006. Carboxyl terminus of severe acute respiratory syndrome coronavirus nucleocapsid protein: self-association analysis and nucleic acid binding characterization. Biochemistry 45, 11827–11835.

Pang, H., Liu, Y., Han, X., Xu, Y., Jiang, F., Wu, D., Kong, X., Bartlam, M., Rao, Z., 2004. Protective humoral responses to severe acute respiratory syndrome-associated coronavirus: implications for the design of an effective protein-based vaccine. J. Gen. Virol. 85, 3109–3113.

Pei, H., Liu, J., Cheng, Y., et al., 2005. Expression of SARS-coronavirus nucleocapsid protein in *Escherichia coli* and *Lactococcus lactis* for serodiagnosis and mucosal vaccination. Appl. Microbiol. Biotechnol. 68, 220–227.

Qiu, M., Wang, J., Wang, H., Chen, Z., Dai, E., Guo, Z., Wang, X., Pang, X., Fan, B., Wen, J., Wang, J., Yang, R., 2005. Use of the COOH portion of the nucleocapsid protein in an antigen-capturing enzyme-linked immunosorbent assay for specific and sensitive detection of severe acute respiratory syndrome coronavirus. Clin. Diagn. Lab. Immunol. 12, 474–476.

Rota, P.A., Oberste, M.S., Monroe, S.S., et al., 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300, 1394–1399.

Rowland, R.R., Chauhan, V., Fang, Y., et al., 2005. Intracellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein: absence of nucleolar accumulation during infection and after expression as a recombinant protein in vero cells. J. Virol. 79, 11507–11512.

Saijo, M., Ogino, T., Taguchi, F., et al., 2005. Recombinant nucleocapsid protein-based IgG enzyme-linked immunosorbent assay for the serological diagnosis of SARS. J. Virol. Methods 125, 181–186.

Shang, B., Wang, X.Y., Yuan, J.W., et al., 2005. Characterization and application of monoclonal antibodies against N protein of SARS-coronavirus. Biochem. Biophys. Res. Commun. 336, 110–117.

Spiegel, M., Pichlmair, A., Martinez-Sobrido, L., et al., 2005. Inhibition of beta interferon induction by severe acute respiratory syndrome coronavirus suggests a two-step model for activation of interferon regulatory factor 3. J. Virol. 79, 2079–2086.

Sun, Z.F., Meng, X.J., 2004. Antigenic cross-reactivity between the nucleocapsid protein of severe acute respiratory syndrome (SARS) coronavirus and polyclonal antisera of antigenic group I animal coronaviruses: implication for SARS diagnosis. J. Clin. Microbiol. 42 (5), 2351–2352.

Surjit, M., Liu, B., Chow, V.T., et al., 2004a. The nucleocapsid protein of the SARS coronavirus is capable of self-association through a C-terminal 209 amino acid interaction domain. Biochem. Biophys. Res. Commun. 317, 1030–1036.

Surjit, M., Liu, B., Jameel, S., et al., 2004b. The SARS coronavirus nucleocapsid protein induces actin reorganization and apoptosis in COS-1 cells in the absence of growth factors. Biochem. J. 383, 13–18.

Surjit, M., Liu, B., Kumar, P., et al., 2004c. The nucleocapsid protein of the SARS coronavirus is capable of self-association through a C-terminal 209 amino acid interaction domain. Biochem. Biophys. Res. Commun. 317, 1030–1036.

Surjit, M., Kumar, R., Mishra, R.N., et al., 2005. The severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated and localizes in the cytoplasm by 14-3-3-mediated translocation. J. Virol. 79, 11476–11486.

Surjit, M., Liu, B., Chow, V.T., et al., 2006. The nucleocapsid protein of severe acute respiratory syndrome-coronavirus inhibits the activity of cyclin-cyclin-dependent kinase complex and blocks S phase progression in mammalian cells. J. Biol. Chem. 281, 10669–10681.

Timani, K.A., Liao, Q., Ye, L., et al., 2005. Nuclear/nucleolar localization properties of C-terminal nucleocapsid protein of SARS coronavirus. Virus Res. 114, 23–34.

Tong, S., Lingappa, J.R., Chen, Q., et al., 2004. Direct sequencing of SARS-coronavirus S and N genes from clinical specimens shows limited variation. J. Infect. Dis. 190, 1127–1131.

Wang, Y., Wu, X., Wang, Y., et al., 2004. Low stability of nucleocapsid protein in SARS virus. Biochemistry 43, 11103–11108.

Wang, Z., Yuan, Z., Matsumoto, M., et al., 2005. Immune responses with DNA vaccines encoded different gene fragments of severe acute respiratory

syndrome coronavirus in BALB/c mice. Biochem. Biophys. Res. Commun. 327, 130–135.

Woo, P.C.Y., Lau, S.K.P., Wong, B.H.L., Chan, K.H., Chu, C.M., Tsoi, H.W., Huang, Y., Peiris, J.S.M., Yuen, K.Y., 2004. Longitudinal profile of immunoglobulin G (IgG), IgM, and IgA antibodies against the severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein in patients with pneumonia due to the SARS coronavirus. Clin. Diagn. Lab. Immunol. 11, 665–668.

Woo, P.C., Lau, S.K., Wong, B.H., et al., 2005. Differential sensitivities of severe acute respiratory syndrome (SARS) coronavirus spike polypeptide enzyme-linked immunosorbent assay (ELISA) and SARS coronavirus nucleocapsid protein ELISA for serodiagnosis of SARS coronavirus pneumonia. J. Clin. Microbiol. 43, 3054–3058.

Yan, X., Hao, Q., Mu, Y., et al., 2006. Nucleocapsid protein of SARS-CoV activates the expression of cyclooxygenase-2 by binding directly to regulatory elements for nuclear factor-kappa B and CCAAT/enhancer binding protein. Int. J. Biochem. Cell Biol. 38, 1417–1428.

Yip, C.W., Hon, C.C., Zeng, F., Chow, K.Y., Chan, K.H., Peiris, J.S., Leung, F.C., 2007. Naturally occurring anti-*Escherichia coli* protein antibodies in the sera of healthy humans cause analytical interference in a recombinant nucleocapsid protein-based enzyme-linked immunosorbent assay for serodiagnosis of severe acute respiratory syndrome. Clin. Vaccine Immunol. 14, 99–101.

You, J., Dove, B.K., Enjuanes, L., et al., 2005. Subcellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein. J. Gen. Virol. 86, 3303–3310.

Yu, I.M., Gustafson, C.L., Diao, J., et al., 2005. Recombinant severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein forms a dimer through its C-terminal domain. J. Biol. Chem. 280, 23280–23286.

Yu, I.M., Oldham, M.L., Zhang, J., et al., 2006. Crystal structure of the severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein dimerization domain reveals evolutionary linkage between corona- and arteriviridae. J. Biol. Chem. 281, 17134–17139.

Yu, F., Le, M.Q., Inoue, S., et al., 2007. Recombinant truncated nucleocapsid protein as antigen in a novel immunoglobulin M capture enzyme-linked immunosorbent assay for diagnosis of severe acute respiratory syndrome coronavirus infection. Clin. Vaccine Immunol. 14, 146–149.

Zakhartchouk, A.N., Viswanathan, S., Mahony, J.B., et al., 2005. Severe acute respiratory syndrome coronavirus nucleocapsid protein expressed by an adenovirus vector is phosphorylated and immunogenic in mice. J. Gen. Virol. 86, 211–215.

Zhang, L., Wei, L., Jiang, D., Wang, J., Cong, X., Fei, R., 2007. SARS-CoV nucleocapsid protein induced apoptosis of COS-1 mediated by the mitochondrial pathway. Artif. Cells Blood Substit. Immobil. Biotechnol. 35, 237–253.

Zhao, P., Cao, J., Zhao, L.J., et al., 2004. Immune responses against SARS-coronavirus nucleocapsid protein induced by DNA vaccine. Virology 331, 128–135.

Zhao, J., Huang, Q., Wang, W., Zhang, Y., Lv, P., Gao, X.-M., 2007. Identification and characterization of dominant helper T-cell epitopes in the nucleocapsid protein of severe acute respiratory syndrome coronavirus. J. Virol. 81 (11), 6079–6088.

Zheng, B., He, M.L., Wong, K.L., et al., 2004. Potent inhibition of SARS-associated coronavirus (SCOV) infection and replication by type I interferons (IFNalpha/beta) but not by type II interferon (IFN-gamma). J. Interferon Cytokine Res. 24, 388–390.

Zhu, M.S., Pan, Y., Chen, H.Q., et al., 2004. Induction of SARS-nucleoprotein-specific immune response by use of DNA vaccine. Immunol. Lett. 92, 237–243.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 406–413
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Phylogenetic analysis of the promoter region of the *CD40L* gene in primates and other mammals☆

Michael E. Steiper [a,b,*], Sonia J. Parikh [a], Julia M. Zichello [a]

[a] *Department of Anthropology, Hunter College of the City University of New York, 695 Park Avenue, New York, NY 10021, United States*
[b] *Departments of Anthropology and Biology, The Graduate Center of the City University of New York, 365 Fifth Avenue, New York, NY 10016, United States*

Received 18 October 2006; received in revised form 12 December 2006; accepted 14 December 2006
Available online 28 December 2006

Abstract

CD40L is a type II membrane protein comprised of 261 amino acids. CD40L plays a crucial role in the immune system where it is primarily expressed on activated T cells and triggers immunoglobulin class switching. The genetic disease X-linked hypergammaglobulinemia (HIGM1, XHIGM or XHIM) is caused by mutations in the *CD40L* gene. Individuals with HIGM1 are susceptible to recurrent infections to pathogens and a relationship has been shown to exist with malaria [Sabeti, P., Usen, S., Farhadian, S., Jallow, M., Doherty, T., Newport, M., Pinder, M., Ward, R., Kwiatkowski, D., 2002a. CD40L association with protection from severe malaria. Genes Immun. 3, 286–291]. In this paper, we phylogenetically examine the promoter region of *CD40L* in primates and other mammals via phylogenetic shadowing. This analysis revealed several regions of the *CD40L* promoter that were highly constrained and thereby inferred to be functional. These constrained regions confirmed many known regulatory sites. In addition, a novel, highly constrained upstream region was also identified which had an NF-AT recognition motif. These analyses also showed that the different mammal groups do not share an exactly similar set of promoter binding sites and taxon-specific promoters have evolved.

Keywords: Evolution; XHIM; Hyper IgM; NF-AT; XHIGM; Phylogenetic shadowing; Immune system; Gene regulation; CD154

1. Introduction

Differences in gene regulation have been hypothesized to be crucial to the evolutionary process (Carroll, 2005; King and Wilson, 1975). While direct functional studies on the evolution of gene regulation are becoming increasingly common, bioinformatical studies of regulatory elements have proven useful both for confirming known regulatory elements and generating target regions for study. Generally, these bioinformatical methods, known as phylogenetic footprinting (Gumucio et al., 1992, 1993; Tagle et al., 1988), examine sequence alignments of two or more species and derive regions that have been conserved over phylogenetic distance. Genetic regions that remain conserved, despite mutational pressure over evolutionary time, are forwarded as candidate functional elements. Recently, a related method, phylogenetic shadowing (Boffelli et al., 2003), has been introduced. The shadowing method examines alignments of closely related taxa and has proven useful in confirming sequence conservation among known sequence elements, such as exons (Boffelli et al., 2003; Ovcharenko et al., 2004). Furthermore, this method also has identified conserved upstream genetic regions which were subsequently determined to be important in gene regulation *in vitro*, in the case of the *APO*(*A*) gene (Boffelli et al., 2003).

In this paper, we utilize phylogenetic shadowing to examine the promoter region of *CD40L* in 18 primates and 4 other mammals. Also known as CD154, gp39, TNFSF5 or TRAP, CD40L is an X-linked, type II membrane protein, comprised of 261 amino acids and encoded by 5 exons (Allen et al., 1993; Gauchat et al., 1993; Hollenbaugh et al., 1992; Padayachee et al., 1992; Villa et al., 1994). CD40L plays a crucial role in the immune system, where it is primarily expressed on activated T cells and triggers immunoglobulin class switching upon

☆ *Note*: Nucleotide sequence data reported in this paper are available in the GenBank under the accession numbers: EF055456–EF055471.

* Corresponding author at: Department of Anthropology, Hunter College of the City University of New York, 695 Park Avenue, New York, NY 10021, United States. Fax: +1 212 772 5423.
E-mail address: msteiper@hunter.cuny.edu (M.E. Steiper).

1567-1348/$ – see front matter
doi:10.1016/j.meegid.2006.12.004

ligation with CD40. Many mutations in the *CD40L* gene result in X-linked hypergammaglobulinemia (HIGM1, XHIGM or XHIM) (Allen et al., 1993; Aruffo et al., 1993; DiSanto et al., 1993; Fuleihan et al., 1993; Korthauer et al., 1993), a disease characterized by elevated IgM and absence of IgG, IgA and IgE. Individuals with XHIGM are immune deficient and are susceptible to recurrent infections from pathogens such as *Pneumocystis carinii* and *Cryptosporidium* (Levy et al., 1997).

The promoter region of the *CD40L* gene is an important target for phylogenetic shadowing analysis for a number of reasons. First, *CD40L* is under strong, recent natural selection in humans (Sabeti et al., 2002b). This study will detect whether this gene is evolving under constraints across humans and other primates and mammals. Second, *CD40L* interacts with malaria. An allele of *CD40L*, which has a mutation in the promoter region, has been identified as ameliorating malaria in Gambian males (Sabeti et al., 2002a). In mice, *CD40L* is implicated in progression to severe malaria (Piguet et al., 2001). Third, there are many known regulatory elements in the *CD40L* promoter that have been examined in both human and mouse cells (Fig. 1). This study will detect whether these known functional regions have been conserved across different species and whether the *CD40L* promoter has species-specific or taxon-specific (e.g. primate-specific) regulatory elements. Fourth, the phylogenetic shadowing method can potentially provide insight into novel conserved sequence elements that can serve as candidates for further functional analyses.

There is a considerable amount of experimental evidence relating to the regulation of *CD40L*, which is expressed transiently and under tight control at various stages, including mRNA transcription (reviewed in Howard and Miller, 2004). Regulation of *CD40L* mRNA transcription is inferred to occur via many elements, including NF-AT (Lobo et al., 2000; Schubert et al., 1995; Tsytsykova et al., 1996), EGR (Cron et al., 2006; Lindgren et al., 2001), AP-1 (Tsytsykova et al., 1996), NF-κB (Srahna et al., 2001), AKNA (Siddiqa et al., 2001), a CD28 response element (Parra et al., 2001) and TFE3/TFEB (Huan et al., 2006). These proteins regulate *CD40L* expression through recognition of known sequence motifs in the promoter region of *CD40L* (summarized in Fig. 1). Previous studies have utilized direct functional approaches in order to implicate these mechanisms in both mouse and human cells. The phylogenetic shadowing method will test whether these regions are phylogenetically conserved or taxon-specific. Furthermore, the method will also find regions that are conserved, but have not yet been examined experimentally for regulatory function.

2. Materials and methods

2.1. Samples

The primate samples outlined in Table 1A were obtained from the Coriell Institute and were targeted for PCR and sequencing of the promoter and first exon of the *CD40L* gene as described in the following sections. These sequences were added to existing DNA sequences from two other primates and four outgroups from GenBank (Table 1B).

2.2. PCR and sequencing

The region targeted for sequencing in these primates was the same region studied by Sabeti et al. (2002a). We followed the established PCR protocols outlined previously for humans (Sabeti et al., 2002a). However, as sequences from species more distantly related to humans were studied, the PCR protocols were necessarily altered. The alterations included usage of different primer combinations to generate amplicons, refinement of PCR parameters, molecular cloning of amplicons for sequencing using plasmid-based primers and generation of new PCR and sequencing primers. The primer sets used in each species, as well as the novel primers, are shown in Supplementary Table 1. PCRs were done using the Eppendorf Triplemaster kit. Where indicated, amplicons were cloned via the Invitrogen TA cloning kit. Plasmid DNA from multiple colonies was miniprepped using the Eppendorf FastPlasmid kit and screened for the insert via clone-test PCR and/or *Eco*RI restriction digestion. These plasmids were sequenced with the M13R and T7 sequencing primer, as well as PCR primers. DNA sequencing was performed by the Dana-Farber/Harvard Cancer Center (DF/HCC) High-Throughput DNA Sequencing Facility. Regions were sequenced in both directions for the entire

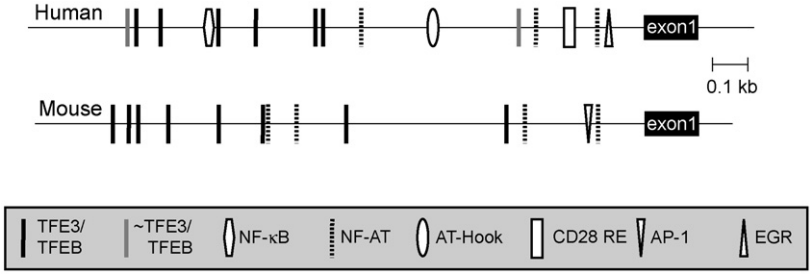


Fig. 1. Schematic representation of the *CD40L* promoters of human and mouse, based on studies from the literature that are described in Section 1. Elements examined in mice are represented only in the mouse and those examined in humans are represented only in the human. However, elements may actually be present in both. The TFE3/TFEB sites examined functionally by Huan et al. (2006) in mice are also represented in humans, following their supplementary figure. The black bars in mouse indicate functionally verified sites, while the gray bars denote 'non-conserved' 'non-optimal' motifs. These same two colors were used for humans, although these sites were not experimentally verified (see Huan et al., 2006, Supplementary Fig. 3). The EGR and AP-1 sites were located adjacent to the NF-AT site, but their precise locations and sequence motifs were not given in the original publications.

Table 1
Samples

Latin name	Common name	Coriell ID	Accession number
(A) Sequenced			
Pan paniscus	Bonobo	NG05253	EF055457
Gorilla gorilla	Gorilla	NG05251	EF055456
Pongo pygmaeus	Sumatran orangutan	NA04272	EF055458
Hylobates gabriellae	Buff-cheeked gibbon	PR00381	EF055459
Colobus guereza	Colobus monkey	PR00980	EF055467
Papio anubis	Olive baboon	PR00036	EF055461
Erythrocebus patas	Patas monkey	NG06116	EF055466
Macaca nemestrina	Pigtailed macaque	NG08452	EF055460
M. mulatta	Rhesus macaque	NG07109	EF055464
M. nigra	Celebes crested macaque	NG07101	EF055462
M. fascicularis	Crab-eating macaque	GM03446	EF055463
M. arctoides	Stumptailed macaque	GM03443	EF055465
Lagothrix lagotricha	Wooly monkey	NG05356	EF055469
Callicebus moloch	Dusky titi monkey	NG06115	EF055470
Saguinus labiatus	Red-chested mustached tamarin	NG05308	EF055468
Lemur catta	Ringtailed lemur	NG07099	EF055471
(B) Obtained from GenBank			
Homo sapiens	Human	–	6983480
Pan troglodytes	Chimpanzee	–	CT005235.2
Mus musculus	Mouse	–	NT_039706
Rattus norvegicus	Rat	–	NW_048052.2
Bos taurus	Cow	–	NW_001030216.1
Canis familiaris	Dog	–	NW_879563.1

sequence, except for the region surrounding a poly-A repeat region between sequence primers −211 and −88. This region was difficult to sequence in both directions 'through' the poly-A repeats. After removal of the poly-A region from the analyses, approximately 110 bp remained in this region between −211 and −88. Because sequences internal to the known regions were targeted in the creation of novel primers, a slightly shorter alignment resulted than in Sabeti et al. (2002a). These sequences are available in GenBank (accession nos. #EF055456–EF055471).

2.3. Sequence alignment and annotation

These 22 sequences were aligned using the program ClustalX (Thompson et al., 1997) using a profile alignment, and subsequently modified by eye. A profile alignment was used because the cow had a large insert. Initially, 21 sequences were aligned without cow, and subsequently the cow sequence was aligned to this alignment via the 'profile alignment' option. The sequence alignment was annotated based on known and putative sequence elements in this region of the CD40L gene reviewed in Fig. 1.

2.4. Phylogenetic analyses and shadowing

A phylogenetic tree was generated from this alignment by maximum parsimony and maximum likelihood using PAUP* (Swofford, 2001). The most parsimonious (MP) trees for the alignment were found using the branch and bound searching technique. Gaps were not scored as informative. Bootstrap values were calculated using the same parameters, with 1000 resampling replicates of 2 heuristic search additions each. The most likely (ML) tree reconstruction was estimated from 10 heuristic search replicates. Initially, MODELTEST (Posada and Crandall, 1998) was used to find the most likely model of molecular evolution, which was used for the ML search. The model was TVM with a γ shape parameter α value of 1.14. In all analyses, a monophyletic dog and cow clade was chosen as the outgroup, based on our current understanding of mammalian phylogenetics (Murphy et al., 2004). Also, the poly-A region and the basepairs corresponding to the large insert within the cow sequence were removed.

The alignment was examined by the 'phylogenetic shadowing' method (Boffelli et al., 2003) using the eShadow program (Loots and Ovcharenko, 2005; Ovcharenko et al., 2004). Phylogenetic shadowing compares closely related sequences and parses sequence alignments into those regions that are conserved and putatively functional and those that are 'neutral' or putatively non-functional. Two statistical methods are implemented in eShadow, a hidden Markov model island approach (HMMI) and a divergence threshold (DT) approach. The DT approach simply determines a conserved element on the basis of the number of matches in an alignment region. A threshold percentage of 20% variation was chosen, and all regions with sequence conservation in excess of this threshold were determined to be conserved for a given 'window' length. Window length was 80 bp. The program also uses a two state HMMI, which predicts regions that are conserved, based on three parameters: the probability of being a slowly mutating site (eS), a fast mutating site (eF) or the probability that a site changes between fast and slow (T). Here, for the HMMI analyses these parameters were optimized by a known

functional region of the *CD40L* alignment (exon 1) using the eShadow maximum likelihood implementation. Four groups were examined: (i) all species; (ii) primates and rodents; (iii) primates, cow and dog; (iv) primates. The HMMI parameters were (i) eS = 0.7244, eF = 0.4242 and $T = 0.0009$; (ii) eS = 0.7281, eF = 0.5234 and $T = 0.0009$; (iii) eS = 0.7949, eF = 0.5678 and $T = 0.0009$; (iv) eS = 0.8910, eF = 0.7413 and $T = 0.0009$.

3. Results

3.1. Alignment

In total, 22 sequences were aligned. The sequences formed a 2438 basepair alignment of the *CD40L* promoter region and first exon. After alignment, the cow (*Bos*) sequence was determined to have an insert of 227 basepairs, which was removed from subsequent analyses. Also, a variable poly-A region of 43 basepairs was removed from the alignment. This 2166 basepair alignment was used for further analyses.

3.2. Phylogenetic analysis

Phylogenetic analyses of the alignment resulted in trees that were consistent with our current understanding of mammalian (Murphy et al., 2004) and primate (Goodman et al., 1998) phylogenetic relationships. The MP analysis recovered one MP tree of 1544 steps and a consistency index of 0.82 (Fig. 2A). An ML analysis recovered one tree identical to the MP consensus tree (Fig. 2B; −ln likelihood = 10195.544). One region within the tree bears on phylogenetic questions not firmly resolved within primates. Within the platyrrhines, the three species examined are representative of each of the three main families of platyrrhines. The phylogeny recovered here weakly supports a closer relationship between Atelidae (*Lagothrix*) and Cebidae (*Saguinus*). The relationships among these two groups, and the remaining platyrrhine group (Pitheciidae), have proven difficult to resolve (Opazo et al., 2006).

Table 2
Conserved regions

Species groups	HMM islands[a]	DT regions[b]
All species	332–463 (132 bps, 526.936) 1332–1912 (581 bps, 5091.558)	1427–1568 (142 bps) 1538–1638 (101 bps) 1562–1740 (179 bps) 1811–1890 (80 bps)
Primates and rodents	332–464 (133 bps, 518.714) 1327–1912 (586 bps, 4750.271)	1377–1473 (97 bps) 1413–1745 (333 bps) 1670–1766 (97 bps) 1802–1915 (114 bps)
Primates, dog and cow	1278–1917 (640 bps, 4478.436)	1412–1571 (160 bps) 1499–1584 (86 bps) 1517–1597 (81 bps) 1533–1767 (235 bps) 1691–1771 (81 bps) 1694–1774 (81 bps) 1747–1830 (84 bps) 1776–1875 (100 bps) 1801–1919 (119 bps)
Primates	189–458 (270 bps, 816.137) 1278–1913 (636 bps, 4103.606)	1–43 (43 bps) 150–529 (380 bps) 466–546 (81 bps) 470–550 (81 bps) 867–948 (82 bps) 1041–1126 (86 bps) 1240–1960 (721 bps)

[a] Conserved regions based on the HMMI method. Regions are followed by their length and the score.
[b] Conserved regions based on the DT method. Regions are followed by their length.

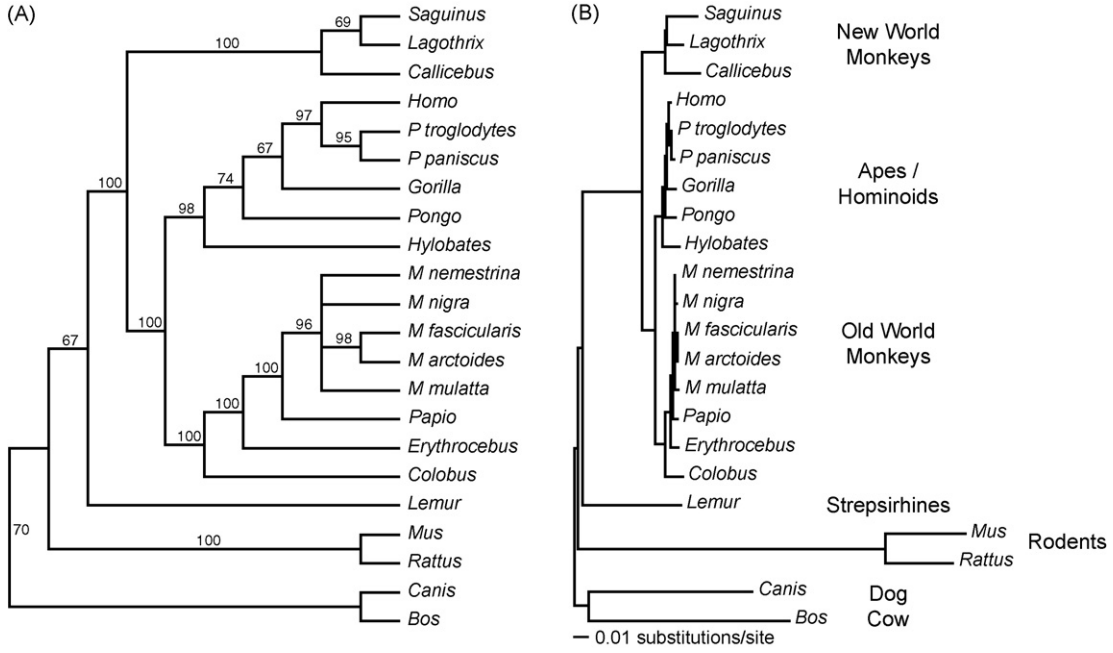


Fig. 2. (A) Maximum parsimony tree. Bootstrap values appear below branches. (B) Maximum likelihood tree.

3.3. Phylogenetic shadowing

Phylogenetic shadowing was used to determine regions of the alignments that were conserved across phylogenetic distance. Four groups were examined: (i) all species; (ii) primates and rodents; (iii) primates, cow and dog; (iv) primates. In all alignments, the exon and its immediate upstream region were strongly conserved, based on both methods for assessing sequence change, DT and HMMI (Table 2 and Fig. 3). This region was conserved in all species groups analyzed and averaged 611 basepairs in length, based on the HMMI method. Of this 611 bp, 155 bp were from exon 1, meaning that only 25% of this conserved region was exonic. This conserved region included known or suspected regulatory sites (two NF-AT sites, an EGR site, an AP-1 site, a CD28 response element and a TFE3/TFEB site).

Further upstream, in three of the four species groups, a second conserved region was recovered. This region was largest in the primate group (270 bps based on HHMI and 400 bps based on the DT method). The DT method encompassed the NF-κB site, but the HMMI method did not. This region did encompass some of the recently described TFE3/TFEB control elements.

Because of the high level of conservation of this upstream region, especially in primates, and its relative lack of regulatory elements, this alignment region was examined for regulatory sites already known to have importance in the regulation of *CD40L*. In this region, an additional NF-AT transcription site was found (Figs. 3 and 4). The site was completely conserved in all primates, as well as dog and cow, but was part of a larger deleted region in rodents (Fig. 4). This site will be referred to as the 'novel NF-AT site'.

Between the upstream and downstream conserved regions, few regions of sequence conservation were found, despite the regulatory elements in this region. Only in primates were two conserved regions identified using the DT method. The regulatory sites in this region include the 'distal' NF-AT site, determined to be most important among the three examined human NF-AT sites (Lobo et al., 2000). The boundary of the DT conserved region 867–948, however, only covered part of the 6 bp NF-AT site. A TFE3/TFEB site was in the conserved region. The conserved region from 1041 to 1126 approached

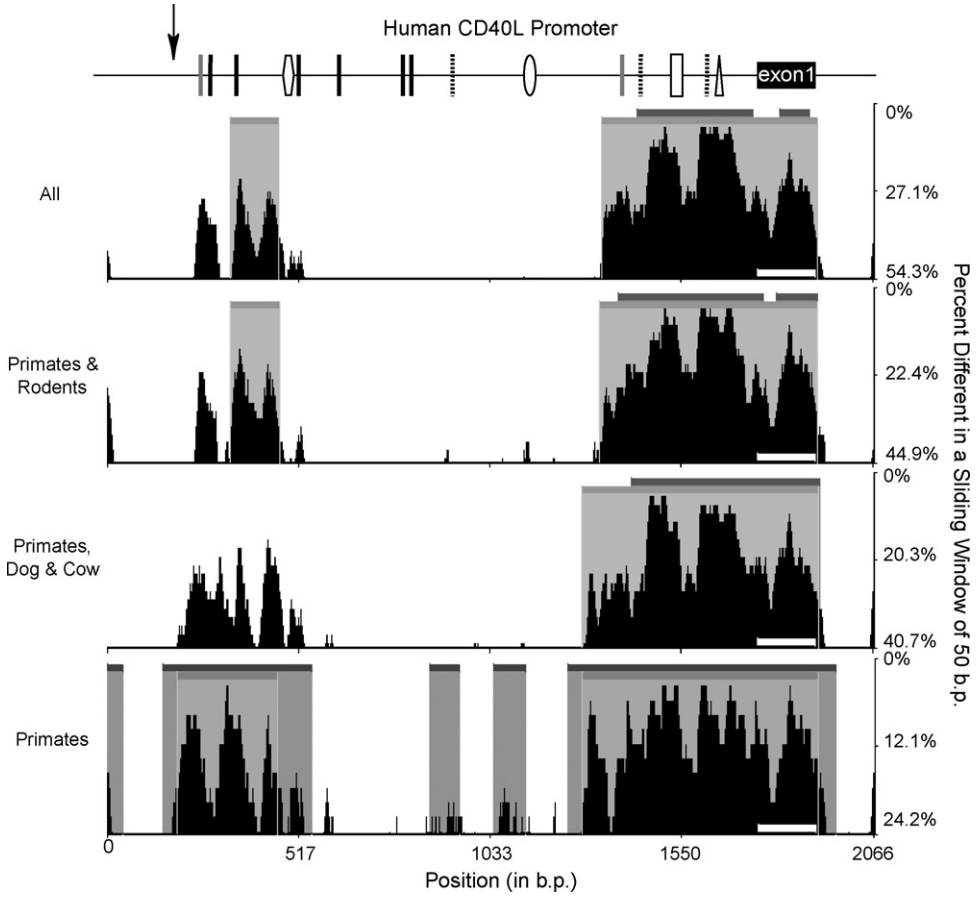


Fig. 3. Phylogenetic shadowing results for all four species groups. At top of figure is the human promoter region as a reference, drawn to be on the same scale as the results of the four shadowing diagrams. The symbols on the human promoter follow Fig. 1. On the *X*-axis is the position along the human sequence (total length excluding gaps was 2066 bp). On the *Y*-axis is the percentage difference based on sliding windows 50 bp in length (which is different from the window size used by the DT method). The level of difference appears in black. Regions identified as conserved by the DT method are in dark gray. Regions identified as conserved by the HMMI method are in light gray. The exon is included as a white bar. These images were generated by eShadow in color and modified to grayscale. Arrow denotes placement of the novel NF-AT described in text and Fig. 4.

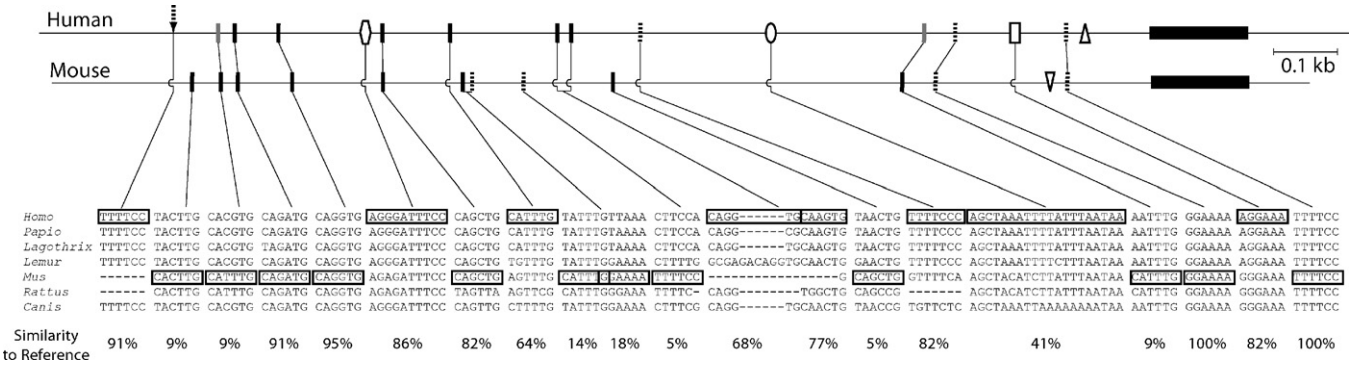


Fig. 4. Alignments of the regulatory regions. All regulatory regions from human and mouse are shown. Homologies between the regulatory motifs are indicated by horizontal lines connecting the elements. For each regulatory site a sampling of species is shown as an alignment. Gaps are indicated with a '-'. The putative novel NF-AT site is indicated by an arrow. "Similarity to reference" is the percentage of species (based on the entire 22 species alignment) exactly matching the species sequence motif indicated by the box. The box was placed with the species where the direct functional studies were done, or where the sequence motif was described.

the beginning of the AKNA site, but did not encompass it. None of these regions were conserved in the other species groups. The polymorphism linked to resistance to cerebral malaria in males (Sabeti et al., 2002a) was located in this region (position 976) but was not determined to be in a conserved region.

3.4. Phylogenetic specificity

Only the two NF-AT sites closest to exon 1 were completely conserved among all 22 species (Fig. 4). These sites were located in a region of high sequence conservation. In this region, the CD28 response element motif was found in 18 of the 22 species studied (82%), all of them primates. This was the same pattern as found in the 'distal' NF-AT site. Phylogenetically, these two sites were 'primate specific'. Other NF-AT sites were found in only one species (5%), while others were found in only a minority of species (18%).

In the upstream conserved region, many of the TFE3/TFEB sites were highly conserved among taxa (e.g. 95, 91 and 82% of species). Both the novel NF-AT site and the NF-κB site were conserved in almost all primates, and dog, and cow, but were not recovered in mouse or rat.

4. Discussion

Phylogenetic shadowing is a useful technique that can assist in understanding gene regulation *in silico*. At the *CD40L* gene, many functional studies have examined regulatory regions in both human and mice cells. Many of these functional regions are conserved, particularly those in the vicinity immediately upstream of the first exon. This region includes TFE3/TFEB, EGR, AP-1, NF-AT and a CD28 response element. Two of the NF-AT sites are found in all taxa studied, possibly representing a core set of *CD40L* regulatory elements that is consistent among all mammals and required for the expression of the *CD40L* gene.

Upstream from this 'core', sequence conservation lessens, despite the presence of known regulatory sites. This region includes elements such as AKNA, which was not in a statistically conserved region. Another noteworthy site was the 'distal' NF-AT site that was determined to have the strongest known effect on *CD40L* regulation of the three examined human NF-AT sites (Lobo et al., 2000). Although this 'distal' NF-AT site was not completely within a conserved region, the specific 6 bp motif was conserved in all primates and the dog and cow. This example may be instructive as to the limits of the resolution of the phylogenetic shadowing method.

Further upstream, a smaller conserved region was found. Only in the primates did the region of sequence conservation include the NF-κB regulatory site. Other TFE3/TFEB regulatory sites located in this upstream conserved region were described only recently (Huan et al., 2006). The bioinformatical identification of this region as conserved dovetails with these functional results. Interestingly, this conserved region contains a novel NF-AT site. NF-AT sites are strongly implicated in the regulation of *CD40L* by a number of studies (Lobo et al., 2000; Schubert et al., 1995; Tsytsykova et al., 1996) and are highly conserved elsewhere in the *CD40L* promoter. This novel NF-AT site is conserved in all primates, dog and cow, while in rodents there is a deletion of this region. This is a similar phylogenetic pattern to the 'distal' NF-AT site described above. A greater apparent level of conservation between primates and dog and cow (rather than between primates and rodents) is surprising because rodents and primates are closer phylogenetic relatives. Parsimony suggests that these three NF-AT sites were lost in rodents. This phylogenetic pattern has also been found in the regulatory sites of other genes (Perry et al., 2006), and may perhaps hint at a more general phenomenon affecting the rodent lineage.

This evidence strongly suggests that the upstream conserved region is worthy of additional study to understand its role in the regulation of *CD40L* expression in non-rodent mammals. This region may include three different regulatory elements, including NF-AT. This level of sequence conservation, on par with the region closer to the transcription start site, suggests that this region plays a role in the expression of *CD40L*.

Similarly, this study has shown that some regulatory sites are found across all mammals, while other regulatory sites are specific to particular orders of mammals. For example, the novel NF-AT site and the NF-κB site are parsimoniously

reconstructed as having been lost in rodents. Other elements, such as the 'distal' NF-AT site and the CD28 response element, are primate specific. Other sites are completely conserved across all examined orders. This highlights the importance of comparative bioinformatics in understanding potential for translating the results of studies in one organism (e.g. mice) and others (e.g. humans). Furthermore, this reveals that certain elements may be critical to *CD40L* regulation in all mammals, while other elements may be taxon specific. These specificities may have evolved by natural selection in response to unique sets of selective pressures facing different orders of mammals, such as pathogens.

This study also has relevance to human health, as mutations in the *CD40L* gene are implicated in XHIGM. *CD40L* mutations which lead to XHIGM listed in the *CD40L* mutation database and elsewhere point towards mutations within the coding region of the gene, e.g. frameshift, nonsense, missense and splice site mutations (Etzioni and Ochs, 2004; Vihinen, 2006). These different mutations can present with different phenotypes (Lopez-Granados et al., 2003). In 15% of HIGM patients, however, a genetic etiology is unknown (Etzioni and Ochs, 2004). Based on the strong conservation in the promoter of the gene, one hypothesis is that some of these patients have mutations within the promoter of *CD40L* that lead to an HIGM phenotype. Additional sequencing and analysis of these unclassified cases would enable a test of this hypothesis. The extreme conservation of many of these promoter regions and their role in gene function suggests that changes in these sequences could disrupt *CD40L* expression.

Acknowledgements

The authors thank Weigang Qiu and Derrick Brazill for comments on this manuscript. This research was supported by Grant NIGMS S06-060654. The infrastructure of the Anthropological Genetics Lab at Hunter College was supported by Grant Number RR03037 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NCRR or NIH.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at: doi:10.1016/j.meegid.2006.12.004.

References

Allen, R.C., Armitage, R.J., Conley, M.E., Rosenblatt, H., Jenkins, N.A., Copeland, N.G., Bedell, M.A., Edelhoff, S., Disteche, C.M., Simoneaux, D.K., et al., 1993. CD40 ligand gene defects responsible for X-linked hyper-IgM syndrome. Science 259, 990–993.

Aruffo, A., Farrington, M., Hollenbaugh, D., Li, X., Milatovich, A., Nonoyama, S., Bajorath, J., Grosmaire, L.S., Stenkamp, R., Neubauer, M., et al., 1993. The CD40 ligand, gp39, is defective in activated T cells from patients with X-linked hyper-IgM syndrome. Cell 72, 291–300.

Boffelli, D., McAuliffe, J., Ovcharenko, D., Lewis, K.D., Ovcharenko, I., Pachter, L., Rubin, E.M., 2003. Phylogenetic shadowing of primate sequences to find functional regions of the human genome. Science 299, 1391–1394.

Carroll, S.B., 2005. Evolution at two levels: on genes and form. PLoS Biol. 3, e245.

Cron, R.Q., Bandyopadhyay, R., Genin, A., Brunner, M., Kersh, G.J., Yin, J., Finkel, T.H., Crow, M.K., 2006. Early growth response-1 is required for CD154 transcription. J. Immunol. 176, 811–818.

DiSanto, J.P., Bonnefoy, J.Y., Gauchat, J.F., Fischer, A., de Saint Basile, G., 1993. CD40 ligand mutations in X-linked immunodeficiency with hyper-IgM. Nature 361, 541–543.

Etzioni, A., Ochs, H.D., 2004. The hyper IgM syndrome—an evolving story. Pediatr. Res. 56, 519–525.

Fuleihan, R., Ramesh, N., Loh, R., Jabara, H., Rosen, R.S., Chatila, T., Fu, S.M., Stamenkovic, I., Geha, R.S., 1993. Defective expression of the CD40 ligand in X chromosome-linked immunoglobulin deficiency with normal or elevated IgM. Proc. Natl. Acad. Sci. U.S.A. 90, 2170–2173.

Gauchat, J.F., Aubry, J.P., Mazzei, G., Life, P., Jomotte, T., Elson, G., Bonnefoy, J.Y., 1993. Human CD40-ligand: molecular cloning, cellular distribution and regulation of expression by factors controlling IgE production. FEBS Lett. 315, 259–266.

Goodman, M., Porter, C.A., Czelusniak, J., Page, S.L., Schneider, H., Shoshani, J., Gunnell, G., Groves, C.P., 1998. Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence. Mol. Phylogenet. Evol. 9, 585–598.

Gumucio, D.L., Heilstedt-Williamson, H., Gray, T.A., Tarle, S.A., Shelton, D.A., Tagle, D.A., Slightom, J.L., Goodman, M., Collins, F.S., 1992. Phylogenetic footprinting reveals a nuclear protein which binds to silencer sequences in the human gamma and epsilon globin genes. Mol. Cell. Biol. 12, 4919–4929.

Gumucio, D.L., Shelton, D.A., Bailey, W.J., Slightom, J.L., Goodman, M., 1993. Phylogenetic footprinting reveals unexpected complexity in trans factor binding upstream from the epsilon-globin gene. Proc. Natl. Acad. Sci. U.S.A. 90, 6018–6022.

Hollenbaugh, D., Grosmaire, L.S., Kullas, C.D., Chalupny, N.J., Braesch-Andersen, S., Noelle, R.J., Stamenkovic, I., Ledbetter, J.A., Aruffo, A., 1992. The human T cell antigen gp39, a member of the TNF gene family, is a ligand for the CD40 receptor: expression of a soluble form of gp39 with B cell co-stimulatory activity. EMBO J. 11, 4313–4321.

Howard, L.M., Miller, S.D., 2004. Immunotherapy targeting the CD40/CD154 costimulatory pathway for treatment of autoimmune disease. Autoimmunity 37, 411–418.

Huan, C., Kelly, M.L., Steele, R., Shapira, I., Gottesman, S.R., Roman, C.A., 2006. Transcription factors TFE3 and TFEB are critical for CD40 ligand expression and thymus-dependent humoral immunity. Nat. Immunol. 7, 1082–1091.

King, M.C., Wilson, A.C., 1975. Evolution at two levels in humans and chimpanzees. Science 188, 107–116.

Korthauer, U., Graf, D., Mages, H.W., Briere, F., Padayachee, M., Malcolm, S., Ugazio, A.G., Notarangelo, L.D., Levinsky, R.J., Kroczek, R.A., 1993. Defective expression of T-cell CD40 ligand causes X-linked immunodeficiency with hyper-IgM. Nature 361, 539–541.

Levy, J., Espanol-Boren, T., Thomas, C., Fischer, A., Tovo, P., Bordigoni, P., Resnick, I., Fasth, A., Baer, M., Gomez, L., Sanders, E.A., Tabone, M.D., Plantaz, D., Etzioni, A., Monafo, V., Abinun, M., Hammarstrom, L., Abrahamsen, T., Jones, A., Finn, A., Klemola, T., DeVries, E., Sanal, O., Peitsch, M.C., Notarangelo, L.D., 1997. Clinical spectrum of X-linked hyper-IgM syndrome. J. Pediatr. 131, 47–54.

Lindgren, H., Axcrona, K., Leanderson, T., 2001. Regulation of transcriptional activity of the murine CD40 ligand promoter in response to signals through TCR and the costimulatory molecules CD28 and CD2. J. Immunol. 166, 4578–4585.

Lobo, F.M., Xu, S., Lee, C., Fuleihan, R.L., 2000. Transcriptional activity of the distal CD40 ligand promoter. Biochem. Biophys. Res. Commun. 279, 245–250.

Loots, G.G., Ovcharenko, I., 2005. Dcode.org anthology of comparative genomic tools. Nucleic Acids Res. 33, W56–W64.

Lopez-Granados, E., Cambronero, R., Ferreira, A., Fontan, G., Garcia-Rodriguez, M.C., 2003. Three novel mutations reflect the variety of defects causing phenotypically diverse X-linked hyper-IgM syndrome. Clin. Exp. Immunol. 133, 123–131.

Murphy, W.J., Pevzner, P.A., O'Brien, S.J., 2004. Mammalian phylogenomics comes of age. Trends Genet. 20, 631–639.

Opazo, J.C., Wildman, D.E., Prychitko, T., Johnson, R.M., Goodman, M., 2006. Phylogenetic relationships and divergence times among New World monkeys (Platyrrhini primates). Mol. Phylogenet. Evol. 40, 274–280.

Ovcharenko, I., Boffelli, D., Loots, G.G., 2004. eShadow: a tool for comparing closely related sequences. Genome Res. 14, 1191–1198.

Padayachee, M., Feighery, C., Finn, A., McKeown, C., Levinsky, R.J., Kinnon, C., Malcolm, S., 1992. Mapping of the X-linked form of hyper-IgM syndrome (HIGM1) to Xq26 by close linkage to HPRT. Genomics 14, 551–553.

Parra, E., Mustelin, T., Dohlsten, M., Mercola, D., 2001. Identification of a CD28 response element in the CD40 ligand promoter. J. Immunol. 166, 2437–2443.

Perry, G.H., Verrelli, B.C., Stone, A.C., 2006. Molecular evolution of the primate developmental genes MSX1 and PAX9. Mol. Biol. Evol. 23, 644–654.

Piguet, P.F., Kan, C.D., Vesin, C., Rochat, A., Donati, Y., Barazzone, C., 2001. Role of CD40-CD40L in mouse severe malaria. Am. J. Pathol. 159, 733–742.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Sabeti, P., Usen, S., Farhadian, S., Jallow, M., Doherty, T., Newport, M., Pinder, M., Ward, R., Kwiatkowski, D., 2002a. CD40L association with protection from severe malaria. Genes Immun. 3, 286–291.

Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J., Ackerman, H.C., Campbell, S.J., Altshuler, D., Cooper, R., Kwiatkowski, D., Ward, R., Lander, E.S., 2002b. Detecting recent positive selection in the human genome from haplotype structure. Nature 419, 832–837.

Schubert, L.A., King, G., Cron, R.Q., Lewis, D.B., Aruffo, A., Hollenbaugh, D., 1995. The human gp39 promoter. Two distinct nuclear factors of activated T cell protein-binding elements contribute independently to transcriptional activation. J. Biol. Chem. 270, 29624–29627.

Siddiqa, A., Sims-Mourtada, J.C., Guzman-Rojas, L., Rangel, R., Guret, C., Madrid-Marina, V., Sun, Y., Martinez-Valdez, H., 2001. Regulation of CD40 and CD40 ligand by the AT-hook transcription factor AKNA. Nature 410, 383–387.

Srahna, M., Remacle, J.E., Annamalai, K., Pype, S., Huylebroeck, D., Boogaerts, M.A., Vandenberghe, P., 2001. NF-kappaB is involved in the regulation of CD154 (CD40 ligand) expression in primary human T cells. Clin. Exp. Immunol. 125, 229–236.

Swofford, D.L., 2001. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4. Sinauer Associates, Sunderland, MA.

Tagle, D.A., Koop, B.F., Goodman, M., Slightom, J.L., Hess, D.L., Jones, R.T., 1988. Embryonic epsilon and gamma globin genes of a prosimian primate (*Galago crassicaudatus*). Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints. J. Mol. Biol. 203, 439–455.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Tsytsykova, A.V., Tsitsikov, E.N., Geha, R.S., 1996. The CD40L promoter contains nuclear factor of activated T cells-binding motifs which require AP-1 binding for activation of transcription. J. Biol. Chem. 271, 3763–3770.

Vihinen, M., 2006. CD40Lbase v.2.04.

Villa, A., Notarangelo, L.D., Di Santo, J.P., Macchi, P.P., Strina, D., Frattini, A., Lucchini, F., Patrosso, C.M., Giliani, S., Mantuano, E., et al., 1994. Organization of the human CD40L gene: implications for molecular defects in X chromosome-linked hyper-IgM syndrome and prenatal diagnosis. Proc. Natl. Acad. Sci. U.S.A. 91, 2110–2114.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 414–421
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

The geometry of the wing of *Aedes* (*Stegomyia*) *aegypti* in isofemale lines through successive generations

Nuananong Jirakanjanakit [a], Somjai Leemingsawat [b], Jean Pierre Dujardin [c,*]

[a] *Center for Vaccine Development, Institute of Science and Technology for Research and Development, Mahidol University, Salaya, Puttamonthon, Nakhonpathom, Thailand*
[b] *Department of Medical Entomology, Faculty of Tropical Medicine, Mahidol University, Rajvithee, Bangkok, Thailand*
[c] *UMR IRD-CNRS 2724, Avenue Agropolis, 911, BP 64501, 34394 Montpellier, France*

Received 1 February 2007; received in revised form 3 May 2007; accepted 4 May 2007
Available online 22 May 2007

Abstract

Under a common laboratory environment, three isofemale lines of *Aedes (Stegomyia) aegypti* were used to score metric properties (size and shape) of the wings during 10 generations. Since the number of generations was much higher than the number of founders in each line, genetic drift was expected to occur.

Size tended to slightly increase with time, but its variation among the successive generations did not show any detectable information specific to each line. Shape could discriminate among lines in females. Males of lines A and B were not discriminated before generation 8, after which they became completely separated.

For each line at each generation, the variance of size and the metric disparity index (as an estimate of shape variance) were higher in females than in males. From one generation to another, the within line shape variance decreased in both sexes, while shape similarity progressively and consistently decreased between males of lines A and B. At each generation, in both sexes, shape variance of the pooled lines was higher than that of each line separately, a pattern not observed for size variance.

In conclusion, shape, as a metric property, was closer to a genetic character than size: (i) it showed modifications compatible with the hypothesis of genetic drift and (ii) its variation was related to the complexity of the sample.

Keywords: *Aedes aegypti*; Isofemale lines; Geometric morphometrics; Metric disparity

1. Introduction

Aedes (*Stegomyia*) *aegypti* is the main vector of arthropod-borne viruses in the tropical and subtropical regions. It originated from Africa and expanded around the tropical world with a pantropical distribution in 1930 (Rodhain, 1996). Known as the yellow fever mosquito in Africa and South America, it is the principal vector of dengue worldwide (Christophers, 1960). To fight this latter disease, vector control and entomological surveillance remain the most important issues.

Many laboratory lines were reared to perform tests on sensitivity to control measures (Brogdon et al., 1999; Ganesh et al., 2002), on vectorial capacity (Mourya et al., 2001; Bennett et al., 2005), or to improve our knowledge about the biology or the taxonomy of the mosquito vector (Cook et al., 2005). Isofemale lines were used to assess the genetic basis for dengue susceptibility among females (Miller and Mitchell, 1991; Tardieux et al., 1991). In the present study, laboratory lines were used to explore the morphometric variability of the wings *Ae. aegypti*.

Wings are bi-dimensional rigid structures well suited for morphometric work. For *Ae. aegypti*, they have been used to demonstrate the power of geometric morphometrics in distinguishing among laboratory lines from different geographic origins in Thailand. Although reared in the same insectarium, laboratory lines did not show any evidence of morphological convergence, since on the basis of shape only, 97% of the wings were assigned to their correct origin (Jirakanjanakit and Dujardin, 2005). However, it was not clear which factor could cause the discriminating power of the wing,

* Corresponding author.
E-mail address: fnjpd@diamond.mahidol.ac.th (J.P. Dujardin).

1567-1348/$ – see front matter
doi:10.1016/j.meegid.2007.05.004

either the different geographic origins or the complete isolation among lines (genetic drift, inbreeding), or both.

The present study applied the same technique to compare three laboratory lines which had high probability to be adapted to the same geographic area and to be affected by genetic drift. The first condition, the same environmental adaptation, was approached by collecting the founders from the same breeding site, at the same time and the same stage of development. The second condition, a likely genetic drift, was obtained by examining progenies of a single parental pair during 10 generations.

Among such isofemale lines, the possible differences in wing shape would be attributed to initial genetic differences among parents, to isolation effects after a few generations (genetic drift) as well as to possible micro-environmental effects. If maintained throughout various generations, these shape differences would then be attributed more to genetic, heritable effects than to micro-environmental ones. In that case, we would be allowed to give this technique a certain potential as a genetic marker for studying natural population.

2. Materials and methods

2.1. Sample collection and isofemale line establishment

Three isofemale lines A, B and C of *Ae. aegypti* were established from the larvae collected from the same source in Chanthaburi province, located in the East of Thailand. The emerging inseminated female mosquitoes were allowed to oviposit in 20 ml cup half filled with water and individually put inside a larger, 1.2 l cup covered with nylon mesh. The eggs were allowed to hatch in clean water. Larvae were distributed to six to eight trays and fed daily on dog food (Alpo low fat, to avoid buildup of oil film on water surface) until they became pupae. All the pupae were then transferred to 30 cm × 30 cm cages and the emerging adults were provided with 10% sugar solution. Blood from Swiss mice was also given twice a week for colonization. The three lines were raised in the insectarium for 10 generations at 25–28 °C and 50–60% RH. The development time, as well as the sex of the first and last emerging adults, were scored during 10 generations.

2.2. Data processing and analyses

2.2.1. The geometry of the wing

A total of 1272 male wings and 1270 female wings have been digitized. In each sex, from 80 to 90 wings were dissected at generations F2, F4, F6, F8 and F10, mounted using Euparal medium and stored in slide boxes at room temperature. Digital pictures of mosquito wings were taken under the phase contrast microscope with 4× objective lens. For each wing, 20 landmarks were collected (Fig. 1). Raw coordinates of the 20 landmarks were submitted to Generalized Procrustes and thin-plate spline analyses to generate ''partial warp'' scores (Rohlf, 1993). Affine variation (the ''uniform component''), which was computed separately (Rohlf, 1990), was added to the partial warp scores to constitute the final set of shape variables, i.e. variables describing the displacement of each landmark relative to a consensus wing.

2.2.2. Centroid size

The global size of the wing was described by the ''centroid size'' (Bookstein, 1991). The variance of size was compared between sexes in each line and generation, and among lines, using non-parametric techniques (permutation tests, 1000 runs).

2.2.3. Metric disparity

Because we were interested more in examining the range of phenotypic variation than in comparing average phenotypes, we used the metric disparity index (MD). This index provides a single value measuring the variety of forms in a sample; it is based on the Euclidian distances of each form with the consensus one. For a given sample of individuals, MD may be estimated by the trace of the variance–covariance matrix (the sum of its diagonal elements) of the partial warp scores (including the uniform component) computed using the sample mean as the consensus (Zelditch et al., 2004). Measuring the variety of forms within a sample, the MD value is zero if all wings in the sample are identical. Unless too few individuals are scored (we scored more than 40 in each group in each sex), the MD score is not correlated to the number of individuals (Zelditch et al., 2004).

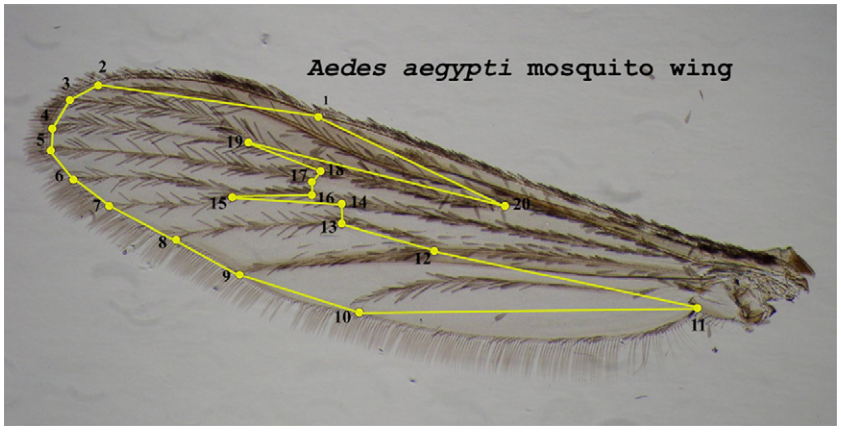


Fig. 1. Selected landmarks on *Aedes aegypti* mosquito wing.

Within each generation, the values of MD for the sum of individuals over the three lines were computed (Table 2, ABC rows). Within each line, as well as for the total of them, the individuals of generations F2, F4, F6, F8 and F10 were pooled into one group, and their metric disparities were computed (Table 2, rightmost column). All these analyses were performed separately on each sex. Statistical comparisons made use of bootstrap methods as described by Zelditch et al. (2004).

2.2.4. Allometry

The residual allometry within lines was estimated by multivariate regression of PW on size, separately in each sex, and statistical significance estimated by 1000 runs permutation tests (Good, 2000). The contribution of size to the discrimination among lines was examined by the linear regression of the two canonical factors derived from shape variation, against size variation. To explore a possible allometric effect in MD comparisons, a regression of MD against size and variance of size was performed within each sex, as well as mixing both sexes.

2.2.5. Euclidian distances, classification trees and reclassification method

The average shape of each generation for each line was computed from the partial warps, and the Euclidian distances between these mean shapes (see Fig. 4) allowed the construction of an UPGMA tree (Figs. 2 and 3).

The individual shapes (i.e. sets of partial warps) were classified according to their Mahalanobis distances to the mean shape of each isofemale line (Table 3).

2.3. Software

Landmarks were scored using tpsDig (Rohlf, 2003). We used our own modules (under General Public License) to

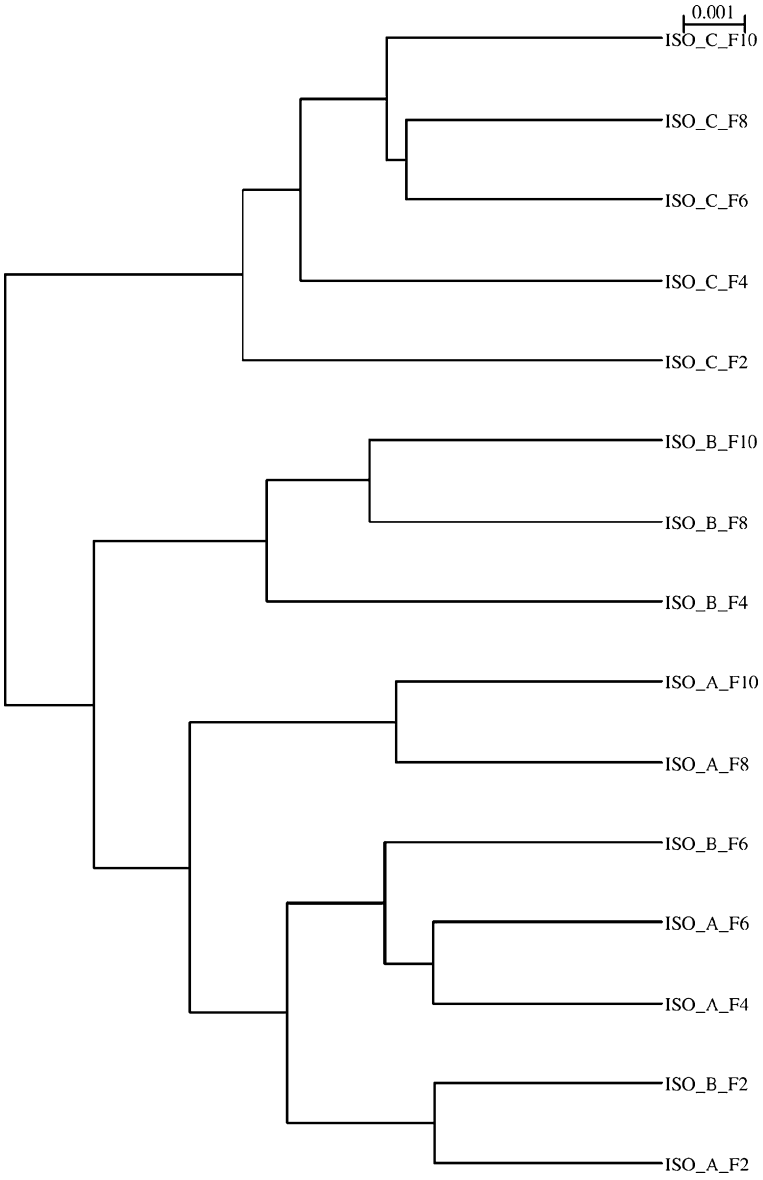


Fig. 2. UPGMA tree of male *Ae. aegypti* at different generations of isofemale lines A, B and C based on Euclidian distances between mean shapes.

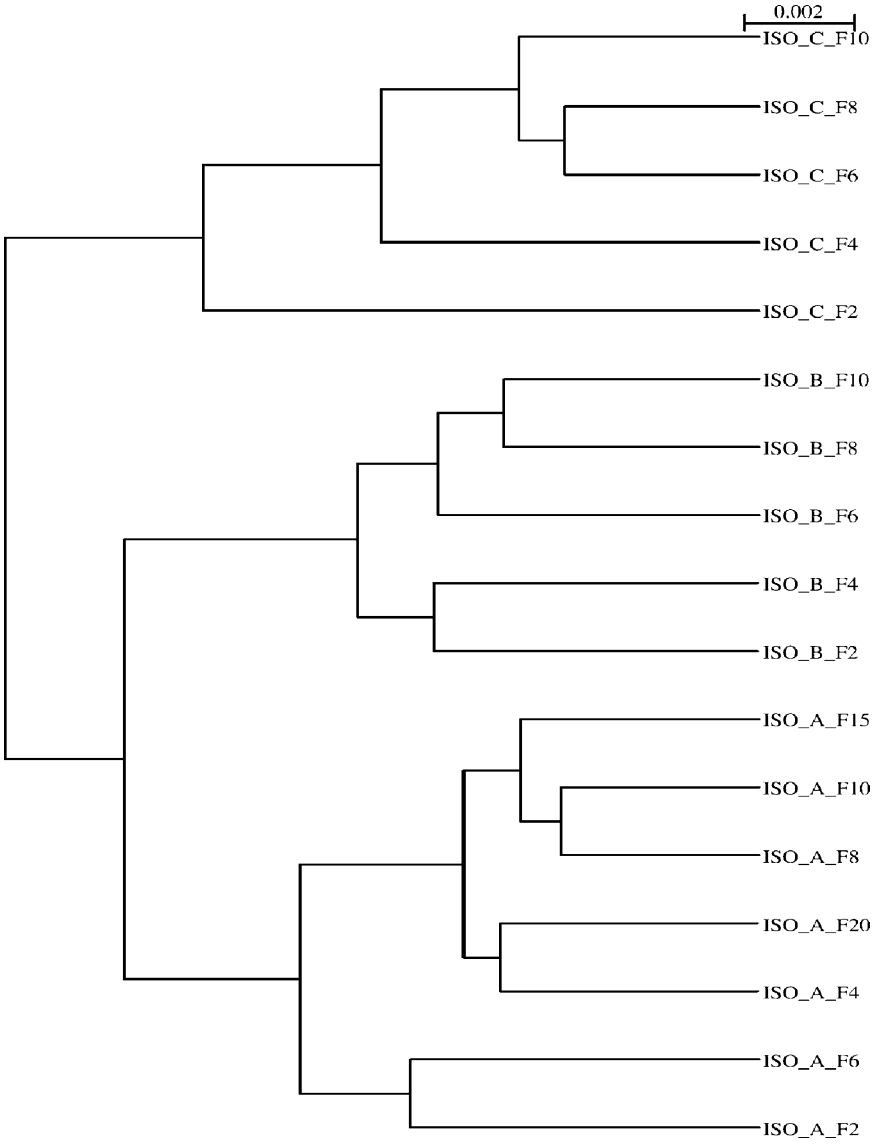


Fig. 3. UPGMA tree of female *Ae. aegypti* at different generations of isofemale lines A, B and C based on Euclidian distances between mean shapes.

perform morphometric analyses: MOG to generate partial warps and centroid sizes, COV to compute metric disparity, Euclidian distance and perform non-parametric comparisons, PAD to compute Mahalanobis distance and classification table. These modules are available at http://www.mpl.ird.fr/morphometrics.

The UPGMA tree was constructed using the PHYLIP package, neighbor module (by J. Felsenstein, http://evolution.genetics.washington.edu/phylip.html).

3. Results

3.1. Mosquitoes

Isofemale lines founded by the three parental pairs of Chanthaburi (Thailand) have been successfully propagated for 10 generations. Life history traits were comparable across lines and generations. It took 8–10 days from hatching to adult eclosion under laboratory condition. The first emerging adult mosquito was always male, but the last was either male or female. The number of larvae was frequently close to 200 individuals by tray and similar among lines.

3.2. Metric properties

3.2.1. Centroid size

On average, females were always significantly larger than males (Table 1). Variance of centroid size (CS) had larger values in females (with one exception, line B, generation 10, Table 1), although not always significantly larger (12/15). Size slightly increased with time in both sexes. Only in 3 comparisons out of 10, the variance of CS over lines was higher than the maximum observed for a single line (males, F4 and F6; females, F2; see Table 1). Over generations, although the number of individuals was increased approximately five times, the variances were lower than the maximum value observed for a single generation (with one exception, females of line C, Table 1).

Table 1
Mean and variance of centroid sizes in each sex at each generation, over the three lines at each generation (row ABC) and over generations for each lines (column F2 to F10)

	F2	F4	F6	F8	F10	F2toF10
Males						
A						
Mean	1011	1029	1012	1069	1030	1031
S.D.	39.4	30.0	34.9	57.7*	36.3	45.6*
n	82	88	88	84	90	432
B						
Mean	1037	1057	1054	1076	1076	1061
S.D.	34.3	26.0*	33.3*	24.7*	36.8*	34.2*
n	72	84	86	86	84	412
C						
Mean	1042	1050	1037	1071	1067	1053
S.D.	29.8*	30.4	32.8*	25.5*	24.8*	31.5*
n	90	80	88	86	84	428
ABC						
Mean	1030	1046	1034	1058	1072	1048
S.D.	37.1	31.1	37.8	43.2	33.2	39.8
n	244	252	262	256	258	1272
Females						
A						
Mean	1368	1373	1355	1426	1344	1375
S.D.	57.3	65.2	35.7*	83.3*	49.3*	68.9*
n	84	80	86	80	90	420
B						
Mean	1447	1434	1410	1449	1419	1432
S.D.	36.0*	51.5*	62.4	54.6*	34.3*	50.9*
n	82	86	84	84	88	424
C						
Mean	1385	1368	1348	1422	1412	1388
S.D.	46.0*	40.1*	49.6*	39.2*	45.0	51.7*
n	82	88	82	86	88	426
ABC						
Mean	1340	1392	1372	1393	1433	1318
S.D.	58.0	60.8	57.3	69.6	44.8	62.2
n	248	254	252	250	266	1270

For each sex, at each scored generation (columns F2, F4, F6, F8 and F10), the values of mean centroid size and its variance are shown within each line (rows A, B and C) and on the pooled lines (row ABC). For each line, it is also shown for the pooled generations (column F2 to F10). F2, F4, F6, F8 and F10 are generations 2, 4, 6, 8 and 10 of lines A, B and C. F2 to F10 is the pooled generations, i.e. the total of individuals from each scored generation for one line, or for the pooled lines. S.D., standard deviation; *n*, number of specimens. *$P < 0.05$ after Bonferroni correction for the comparison of the variance A, B or C with the ABC value of the pooled lines variance.

3.2.2. Metric disparity

Brothers and sisters did not show the same variety of forms: in most of the comparisons throughout generations and lines (13/15, except line A-F6 and line B-F2), the MD values were significantly higher in females. Another observation was visible across time counted in generations: the decrease of MD. In males of the three lines, there was an obvious decrease of MD scores with time. In females it was observed in lines A and B, while the variation in line C did not show visible trend. Over lines, in each sex, at each generation, the MD score produced higher values than the maximum observed for a single line (Table 2, last row), with one exception (males F2, Table 2). Over generations, although the number of individuals was increased five times, the MD scores were lower than the maximum value observed for a single generation (Table 2, last column).

3.2.3. Allometry

Within each line, in each sex, there was a significant correlation of shape with size (detailed results not shown). The contribution of allometric effects to discrimination among lines based on shape was 3% in females and 1% in males for the first canonical factor while it was 3% and 4% for the second one.

3.2.4. Euclidian distances, classification trees and reclassification tests

At generation 2, the closest lines were A and B, the most distant ones were B and C (Fig. 4). The distances among lines did not remain unchanged after 10 generations, especially in males. The males of lines A and B, the closest lines at generation 2, strongly diverged with time (Fig. 4). The males of lines B and C slightly converged with time (Fig. 4).

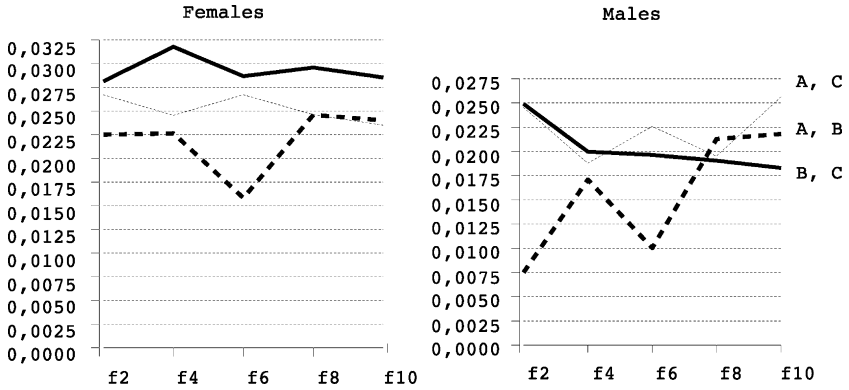


Fig. 4. Euclidian pair-wise distances among lines at each generation. Left graph: females, right graph: males; dotted lines represent distances between A and C; dashed lines represent distances between A and B; plain lines represent distances between B and C.

When using these distances to generate a classification tree, every generation of females was clustered in agreement with the corresponding founder lines. In males, the three lines were consistently separated only after eight generations. Before that, generations F2 and F6 of line B were ''wrongly'' grouped with line A (Fig. 2).

As a result, shape variables classified the wings of the total sample in a more satisfying way for females than for males. Nevertheless, mean classification scores based on Mahalanobis distances were high in both sexes: 94% for females and 85% for males (Table 3). Lower scores in males were due to some confusion between lines A and B (at F2 and F6, see above), not with line C which was correctly recognized at 94%.

4. Discussion

The geometric morphometrics of *Ae. aegypti* wings has been shown previously to be effective in correctly assigning individuals among four separate laboratory lines (Jirakanjanakit and Dujardin, 2005). It was not clear however whether the shape differences among the laboratory lines could be attributed to either genetic drift or inbreeding effects, or to different geographical origins. The use of laboratory lines founded from the same breeding site allowed us to tentatively remove the geographic or adaptive cause of metric variation. Their founding from a single pair and their following during 10 generations under the same laboratory environment allowed us to focus more on possible genetic drift effects.

The classification of successive generations of the three lines A, B and C as obtained by shape variation produced a tree corresponding to the known relationships. Since the contribution of size to shape discrimination among lines was under 5%, this separation of lines could not be explained by allometric effects, it was chiefly attributable to shape variation. In females, each generation was correctly assigned to its corresponding line, while in males there was some mixing between lines A and B during the first generations. Males of lines A and B, almost identical at the time of their founding, became progressively distinct lines, while males of B and C slightly converged with time. This was a pattern more in agreement with genetic drift than with any micro-environmental disturbance, explaining also sexual differences in the reclassification scores (94% in females and 85% in males).

Micro-environmental variation as an explanation for the observed shape variability among isofemale lines could not be excluded from our study, but was probably weak because of two reasons: (i) it could not remove the ''memory'' of the founder

Table 2
Metric disparity measured within each generation in lines A, B and C, over the three lines at each generation (row ABC) and over generations for each lines (column F2 to F10) in males and females

	F2	F4	F6	F8	F10	F2 to F10
Males						
A	856.0	866.0 ns	877.0 ns	795.0 ns	664.0	867
B	966.0 ns	781.0	761.0	790.0 ns	828.0 ns	904
C	641.0	698.0	692.0	530.0	565.0	677
ABC	951.0	894.0	883.0	833.0	844.0	908
Females						
A	989.0	996.0	837.0	800.0	738.0	921
B	907.0	1050.0	917.0	1060.0 ns	845.0	1018
C	816.0	861.0	1085.0 ns	985.0	868.0 ns	994
ABC	1122.0	1199.0	1110.0	1179.0	1027.0	1140

Metric disparity (MD) is measured within each generation in lines A, B and C, on the pooled lines at each generation (row ABC) and, for each line, on the pooled generations (rightmost column F2 to F10) in males and females. Number of specimens like in Table 1. ns, non-significantly smaller than the value for pooled lines (row ABC), after Bonferroni correction. MD multiplied by 10^6.

Table 3
Reclassification scores in male and female mosquitoes of each line

	Total	Correct	Percents
Males			
Line A	432	357	82
Line B	412	314	76
Line C	428	406	94
Females			
Line A	420	397	94
Line B	424	382	90
Line C	426	416	97

Total: number of wings in each line, pooling the 10 generations; correct: number of wings correctly assigned to the corresponding line; percents: this number divided by the total number of wings in the line.

after 10 generations and (ii) at each generation, several trays were used to produce the progeny of each isofemale line (David et al., 2005). On the other hand, according to many studies performed on *Drosophila*, such micro-environmental contribution (the ''common environmental effect'', or CEE) is likely to be low (David et al., 2005).

In agreement with these considerations, our data provided additional arguments for a genetic basis of shape variation by examining the variance of metric properties.

If individuals are added having similar morphotypes to the existing ones, even if they are many, the variance of metric properties is not expected to increase. Within a given line founded by a single parental pair, there is no reason a priori for new genetic variants to appear in 10 generations. On the contrary, due to genetic drift, genetic variation is expected to decrease. Thus, combining 10 generations of a given line into a single group, the resulting variance of centroid size or MD score was expected to show values similar as or lower than values scored for a single generation. Our results were in agreement with that expectation (Tables 1 and 2, rightmost column).

If individuals added to a given sample represent new morphotypes, even if they are very few, the metric variance of this sample will increase. We thus expected the metric variation computed from pooled lines to show significantly higher values than for single lines. This was confirmed at each generation for shape variation (with 25/36 values found significantly higher, Table 2, rows ABC), much less so for size variation (only 5/36 values significantly higher, Table 1, rows ABC).

Contrary to what was suggested previously (Jirakanjanakit and Dujardin, 2005), size did not show here any trend of decreasing with time; it actually slightly increased with time. However, the time is not comparable in both studies: size was shown to decrease after more than 17–564 generations (Jirakanjanakit and Dujardin, 2005), while we examined only 10 generations. This should be verified, as the previous results could actually be reflecting different micro-environmental conditions (different laboratories) instead of true inbreeding effect.

Females were consistently larger than males (Table 1) and they showed larger variance of size (Table 1) and also of shape (Table 2). Since the variance of size was often (12/15) larger in females, the sexual dimorphism observed for metric disparity could partially be explained by an allometric effect. Actually, a significant correlation was found on the total sample between MD scores and mean values and/or variances of size (determination coefficients 30% and 18%, respectively). Nevertheless, no such correlation could be detected within each sex.

Thus, shape showed more consistent evidence for genetic determinism than size. Notably, it was able to show patterns suggestive of genetic drift under the same laboratory environment, allowing to distinguish individuals belonging to recently isolated groups. Our data suggest that the metric disparity index could be used in natural conditions to evaluate the relative complexity of collecting sites. They also confirmed that the wing shape could be used to monitor laboratory lines (Jirakanjanakit and Dujardin, 2005). However, the present demonstration does not mean that shape can be used on natural populations to assess their current level of isolation. Natural populations are not genetically as simple as isofemale lines, and many factors – in addition to isolation and/or bottlenecks – are liable to affect the development of the insect, hence its final shape, exaggerating or blurring its expected relationship to isolation levels.

Acknowledgements

The authors gratefully thank Dr. Christian Bellec (IRD-DSF program) on the support for training, kind suggestion and encouragement. Many thanks to Napaporn Kuatrakool and Sadanun Boonsatien for their help in mosquito rearing and wing slides preparation, and to Pornpimol Rongnaparut (Mahidol University, Department of Biochemistry) for critical reading. This research work was supported by Faculty of Graduate Studies, Mahidol University, academic year 2006.

References

Bennett, K.E., Flick, D., Fleming, K.H., Jochim, R., Beaty, B.J., Black, W.C., 2005. Quantitative trait loci that control dengue-2 virus dissemination in the mosquito *Aedes aegypti*. Gene 170, 185–194.

Bookstein, F.L., 1991. Morphometric Tools for Landmark Data: Geometry and Biology. Cambridge University Press, Cambridge, 435 pp.

Brogdon, W.G., McAllister, J.C., Corwin, A.M., Cordon-Rosales, C., 1999. Independent selection of multiple mechanisms for pyrethroid resistance in Guatemalan *Anopheles albimanus* (Diptera: Culicidae). J. Econ. Entomol. 92, 298–302.

Christophers, S.R., 1960. *Aedes aegypti* (L.). The Yellow Fever Mosquito. Cambridge University Press, London, 739 pp.

Cook, S., Diallo, M., Sall, A.A., Cooper, A., Holmes, E.C., 2005. Mitochondrial markers for molecular identification of *Aedes* mosquitoes (Diptera: Culicidae) involved in transmission of arboviral disease in West Africa. J. Med. Entomol. 42, 19–28.

David, J.R., Gibert, P., Legout, H., Petavy, G., Morreteau, B., 2005. Review: isofemale line in Drosophila: an empirical approach to quantitative trait analysis in natural population. Heredity 94, 3–12.

Ganesh, K.N., Vijayan, V.A., Urmila, J., Gopalan, N., Prakash, S., 2002. Role of esterases and monooxygenase in the deltamethrin resistance in *Anopheles stephensi* Giles (1908), at Mysore. Indian J. Exp. Biol. 40, 583–588.

Good, P., 2000. Permutation Tests: A Practical Guide to Re-sampling Methods for Testing Hypotheses. Springer, New York.

Jirakanjanakit, N., Dujardin, J.P., 2005. Discrimination of *Aedes aegypti* (diptera; culicidae) laboratory lines based on wing geometry. Southeast Asian J. Trop. Med. Public Health 36, 858–861.

Miller, B.R., Mitchell, C.J., 1991. Genetic selection of a flavivirus-refractory strain of the yellow fever mosquito *Aedes aegypti*. Am. J. Trop. Med. Hyg. 45, 399–407.

Mourya, D.T., Gokhale, Basu, A., Barde, P.V., Sapkal, G.N., Padbidri, V.S., Gore, M.M., 2001. Horizontal and vertical transmission of dengue virus type 2 in highly and lowly susceptible strains of *Aedes aegypti* mosquitoes. Acta Virol. 45, 67–71.

Rodhain, F., 1996. Ecology of *Aedes aegypti* in Africa and Asia. Bull. Soc. Pathol. Exot. 89, 103–106.

Rohlf, F.J., 1990. Rotational fit (Procrustes) methods. In: Rohlf, F.J., Bookstein, F.L. (Eds.), Proceedings of the Michigan Morphometrics Workshop. Special Publication Number 2. The University of Michigan Museums, Ann Arbor, MI, pp. 227–236.

Rohlf, F.J., 1993. Morphometric spaces, shape components and the effects of linear transformations. In: Marcus, L.F., Corti, M., Loy, A., Naylor,

G.J.P., Slice, D. (Eds.), Advances in Morphometrics, Proceedings of the 1993 NATO-ASI on Morphometrics, NATO ASI, Series A, Life Sciences. Plenum Publ., New York, pp. 117–129.

Rohlf, F.J., 2003. TpsDig ver 1. 34. Available at http://morph.bio.sunysb.edu./morph/index/html. Department of Ecology and Evolution, State University of New York.

Tardieux, I., Poupel, O., Lapchin, L., Rodhain, F., 1991. Analysis of inheritance of oral susceptibility of *Aedes aegypti* (Diptera: Culicidae) to dengue-2 virus using isofemale lines. J. Med. Entomol. 28, 518–521.

Zelditch, M.L., Donald, L., Swiderski, H., Sheets, D., Fink, W.L., 2004. Geometric Morphometrics for Biologists: A Primer. Elsevier/Academic Press, Paris, 443 pp.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 422–432
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Using a climate-dependent model to predict mosquito abundance: Application to *Aedes* (*Stegomyia*) *africanus* and *Aedes* (*Diceromyia*) *furcifer* (Diptera: Culicidae)

Brigitte Schaeffer [a,*], Bernard Mondet [b], Suzanne Touzeau [a]

[a] *INRA, UR341 Mathématiques et Informatique Appliquées, F-78350 Jouy-en-Josas Cedex, France*
[b] *Institut Français de Pondichéry, CP 33, 11 Saint-Louis Street, Pondicherry 605001, India*

Received 15 February 2007; received in revised form 2 July 2007; accepted 5 July 2007
Available online 10 July 2007

Abstract

Mosquitoes, acting as vectors, are involved in the transmission of viruses. Thus, their abundances, which strongly depend on the weather and environment, are closely linked to major disease outbreaks. The aim of this paper is to provide a tool to predict vector abundance.

In order to describe the dynamics of mosquito populations, we developed a matrix model integrating climate fluctuations. The population is structured in five stages: two egg stages (immature and mature), one larval stage and two female flying stages (nulliparous and parous). The water availability in breeding sites was considered as the main environmental factor affecting the mosquito life-cycle. Thus, the model represents the evolution of the mosquito abundance in each stage over time, in connection with water availability.

The model was used to simulate the abundance trends over 3 years of two mosquito species, *Aedes africanus* (Theobald) and *Aedes furcifer* (Edwards), vectors of the yellow fever virus in Ivory Coast. As both these species breed in tree holes, the water dynamics in the tree hole was reproduced from daily rainfall data. The results we obtained showed a good match between the simulated populations and the field data over the time period considered.

Keywords: Mosquito vector population; Mathematical model; Climate dependency

1. Introduction

Yellow fever, a vector-borne disease, remains a major public health problem both in Africa and South America. It is due to infection by a Flavivirus and causes a viral hemorrhagic fever in humans responsible for 30,000 deaths annually (Tomori, 2004). Furthermore, global warming may increase the outbreak risk of such infectious diseases associated with hot weather or rainfall pattern modifications. Chevalier et al. (2004) showed that geographical distributions of several arthropod-borne zoonoses have dramatically expanded over the last years. Although one knows that climate influences the abundance of vectors, it is still difficult to appreciate the effect of climate variations on the emergence and re-emergence of arbovirus diseases. That is mainly because the epidemiological process is driven by numerous tangled mechanisms, making analyses fairly complex. Breaking up the process and focusing on its different elements, in particular on the biology of the vectors which is the key of the virus transmission, might help to apprehend the global process.

Mathematical models combined with computer simulations are powerful tools for describing and understanding complex biological phenomena. Indeed, modelling allows to test assumptions, explain some events, or compare alternate strategies in decision-making. Thus, models have been widely used for studying complex biological processes such as population dynamics (e.g. Jarry et al., 1996; Ghosh and Pugliese, 2004; Awerbuch-Friedlander et al., 2005; Mazaris and Matsinos, 2006). For mosquito populations in particular, Moon (1976) developed a dynamic model, but that model did not consider environmental factors. Focks et al. (1993) proposed a dynamic life table model for *Aedes aegypti* taking

* Corresponding author at: INRA, Unité de Mathématiques & Informatique Appliquées, Centre de Recherche de Jouy-en-Josas, 78352 Jouy-en-Josas Cedex, France. Tel.: +33 1 34 65 22 18; fax: +33 1 34 65 22 17.
E-mail address: Brigitte.Schaeffer@jouy.inra.fr (B. Schaeffer).

1567-1348/$ – see front matter
doi:10.1016/j.meegid.2007.07.002

into account numerous variables, but lacking flexibility to be generalized to other species. Fouque and Baumgartner (1996) constructed a distributed delay model taking into account only part of the life-cycle of the *Aedes vexans* mosquito, from newly hatched larvae to emerging adults. Ahumada et al. (2004) built a population dynamics matrix model, taking into account temperature and rainfall. These authors obtained encouraging results because their model captured field data quite adequately. However, they acknowledged the importance of breeding site availability which they had not explicitly modelled. Shone et al. (2006) investigated the role of weather on mosquito species using statistical regression models. From this point of view, the dynamical aspect of abundance fluctuations is difficult to appreciate. In an epidemiological context, most studies did not take into account the vector population dynamics and set this population size to a constant value (Esteva and Vargas, 1999; Ngwa and Shu, 2000; Derouich et al., 2003; Ishikawa et al., 2003; Ngwa, 2004). As the necessity of taking into account the fluctuations of the vector density has been shown (Costantino et al., 1998; Chattopadhyay et al., 2004) the influence of the seasonality was sometimes introduced as a sinusoid function in the epidemiological model to roughly describe the mosquito population dynamics (Lord, 2004; Glass, 2005). In order to reproduce as best as possible the dynamics of a vector in its own environment, simulation models were developed (Focks et al., 1993; Depinay et al., 2004). These models include numerous parameters that could be difficult to apprehend in field conditions, such as nutrient competition for example. Validating such models might prove difficult, as sensitive parameters, if not well estimated, might distort the model outputs.

In the present paper, our approach was to develop a model as simple and generic as possible to investigate the mosquito population dynamics, taking into account only the most relevant steps of the mosquito life-cycle and the most influent environmental factors. Such a model aims at capturing the major trends of the mosquito population size along time. Moreover, this model could easily be integrated in a larger modelling process which would, for example, include models of virus development and host demography.

This paper is divided into three major sections. The first one presents the generic model based on the life-cycle graph of a mosquito population. The second section shows how environmental factors are integrated into the model. The third section describes an application of the model using data from yellow fever vectors, *Aedes* (*Stegomya*) *africanus* and *Aedes* (*Dicermomya*) *furcifer*, in West Africa (Mondet, 1994, 1997); parameters were estimated and a sensitivity analysis was performed. Finally, conclusions are drawn.

2. The generic population model

To represent the dynamics of the mosquito population and the climate influence, we chose a matrix framework (Caswell, 2001) that is commonly used to study the dynamics of structured populations (Jensen, 1995; Jarry et al., 1996; Bommarco, 2001; Choi and Ryoo, 2003; Thomas et al., 2005). Moreover, this discrete time step corresponds well to the nature of the field data. In this section, we first present the mosquito life-cycle. Then we describe the death, growth and reproduction processes that will be finally merged to produce the population dynamic model.

2.1. The mosquito life-cycle

The biological cycle of the mosquito is represented in Fig. 1. The *Aedes* females lay their eggs on the limits of stagnant waters in their breeding sites. These eggs hatch when they are flooded. The aquatic larvae, resulting from eggs, pass through four stages of development. The last molt leads to a flying adult, winged, male or female. As soon as female mosquitoes emerge, they are fertilized, mating occurring only once in their lifetime. Then females need to find a blood meal host to start a gonotrophic cycle. When the cycle is over, female mosquitoes will deposit their eggs. After laying, females need a new blood meal to start a new gonotrophic cycle, and so on until they die. A female generally goes through 5–7 cycles.

The population model we built takes into account only females because males are not involved in transmission disease. The model is based on the three phases of the mosquito life-cycle: the eggs, the aquatic forms and the adults who produce the eggs. To describe the maturation of eggs, we took into account two egg stages: an immature stage denoted $w1$, followed by a mature stage denoted $w2$ from which eggs can hatch. The transition between these two stages may be controlled by climate. Then, we considered one larval stage, denoted L, corresponding to the aquatic forms that include instars (1 to 4) and pupae. Right after emergence as aerial adults females enter the nulliparous stage denoted A_1 during which they are fertilized and at the end of which they lay their first eggs. Right after this first laying, they enter the parous stage denoted A_2 during which they go through several more layings. The availability of hosts which provide blood meals and hence govern the laying, was not considered as a limiting factor and so, was not included in the model.

Consequently, our mosquito population is structured in five stages, $w1$, $w2$, L, A_1 and A_2, and we chose a discrete time

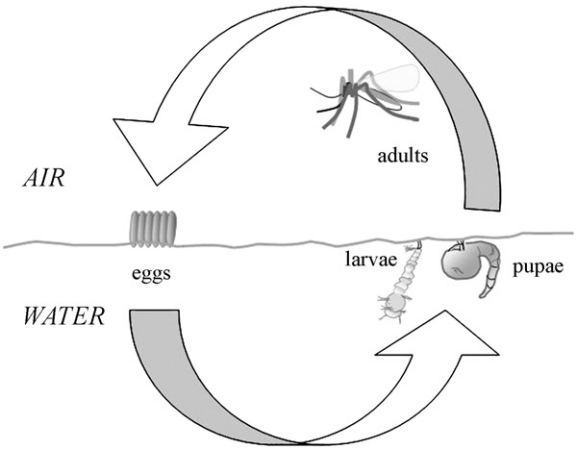


Fig. 1. Biological cycle of mosquitoes.

process to represent its dynamic over time. So at each time t, the population is described by its state vector denoted N_t,

$$N_t = \begin{pmatrix} w1_t \\ w2_t \\ L_t \\ A_{1_t} \\ A_{2_t} \end{pmatrix}.$$

To express the evolution of the population with time, we modelled the processes that affect the abundances of stages i, $i \in \{w1, w2, L, A_1, A_2\}$, during each time step. These processes are represented in the life-cycle graph in Fig. 2: birth, maturing from stage to stage, and death. The following paragraphs present these three processes (in reverse order to simplify the understanding of the model) and then, the resulting population dynamic model.

2.2. Death

Death can occur throughout the life of the mosquito. For each stage i, the mortality rate is denoted $q_{i\mathrm{D}}$ and is defined as the proportion of individuals in stage i that die during each time step.

2.3. Growth

For any stages i and j, $(i, j \in \{w1, w2, L, A_1, A_2\})$, the transition rate from stage i to stage j is denoted q_{ij}, and is defined as the proportion of individuals in stage i at time t that move (or stay) into stage j (or i) at time $t + 1$. As the population is considered to be closed (no migrations), one can write for each stage i:

$$\sum_{j \in \{w1, w2, L, A_1, A_2\}} q_{ij} + q_{i\mathrm{D}} = 1.$$

The residence times which correspond to the development time in each stage are not identical; for instance, the larval stage L can last from 6 to 20 days whereas the parous adult stage, A_2, can be up to 30 days. The time step needs to be chosen so as to be lower than the minimum residence time in any stage and the time unit is set to one time step ($\Delta t = 1$). In the following application, the time step is set to 1 day. Therefore, between two time steps, an individual can either remain in its current stage, or pass into the following stage, except for the A_2 stage, or die. In A_2 stage, an individual can only remain in its stage or die. For instance, for $i = L$, we have: $q_{Lw1} = q_{Lw2} = q_{LA_2} = 0$, hence $q_{LA_1} + q_{LL} + q_{L\mathrm{D}} = 1$. We assumed that the transition rate from stage i to next stage, in the absence of death, is equal to the inverse of the mean residence time in stage i, denoted T^i_{rsd}. It means that in the absence of death, the individuals would leave stage i with frequency $1/T^i_{\mathrm{rsd}}$. The underlying hypothesis is that the stage-duration distribution is geometric: the individual probability of moving from stage i to stage $i + 1$ is a constant, independent of the time spent in stage i. This hypothesis, though rather unrealistic in mosquito populations, is a very common simplification. Incorporating mortality, we apply this frequency to the fraction of the population in stage i that does not die during the time step. With $(1 - q_{i\mathrm{D}})$ as the survival rate for stage i, we obtain:

$$q_{ij} = \frac{1}{T^i_{\mathrm{rsd}}}[1 - q_{i\mathrm{D}}], \quad \text{if } j = i + 1, \text{ otherwise } q_{ij} = 0.$$

As $q_{ii} + \sum_{j \neq i} q_{ij} + q_{i\mathrm{D}} = 1$ we deduce that:

$$q_{ii} = \left(1 - \frac{1}{T^i_{\mathrm{rsd}}}\right)[1 - q_{i\mathrm{D}}].$$

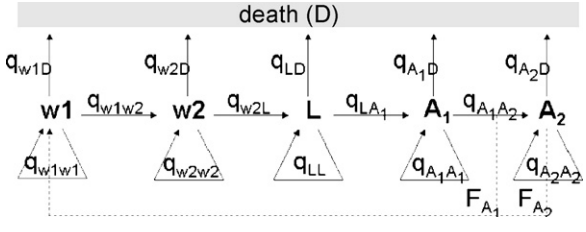


Fig. 2. Life-cycle graph of female mosquitoes. Nodes represent the stages: $w1$ = immature eggs, $w2$ = mature eggs, L = larval stages, A_1 = nulliparous adults, A_2 = parous adults. The q_{ij} $(i, j \in \{w1, w2, L, A_1, A_2\})$ are the transition rates, the $q_{i\mathrm{D}}$ are the mortality rates and the F_i, $i \in \{A_1, A_2\}$, are the fecundities. Each arrow represents a possible transition between two stages during a time step.

2.4. The reproduction process

Female mosquitoes lay eggs several times during their adult stage, each laying occurring after a blood meal. We assumed that the search for a blood meal host takes very little time compared to the model time step, so we neglected this delay. The transition from the nulliparous stage to the parous stage is triggered by the first laying. Let us denote T_{c}, with $T_{\mathrm{c}} \geq \Delta t$, the duration between the emergence and the first laying, which corresponds to the length of a gonotrophic cycle. Thus, at each time step, the transition rate from the nulliparous stage to the parous stage is expressed as:

$$q_{A_1A_2} = \frac{1}{T_{\mathrm{c}}}[1 - q_{A_1\mathrm{D}}].$$

As $q_{A_1A_1} + q_{A_1A_2} + q_{A_1\mathrm{D}} = 1$ we deduce that:

$$q_{A_1A_1} = \left(1 - \frac{1}{T_{\mathrm{c}}}\right)[1 - q_{A_1\mathrm{D}}].$$

Nulliparous females lay their eggs when they enter the parous stage. The number of females that pass into the parous stage is $q_{A_1A_2} \times A_1$. Let nwf be the mean number of female eggs produced by one female. Then the number of eggs produced at each time step by the nulliparous females is given by $\mathrm{nwf} \times q_{A_1A_2} \times A_1$. Hence the fecundity is

$$F_{A_1} = \frac{\mathrm{nwf} \times q_{A_1A_2} \times A_1}{A_1} = \mathrm{nwf} \times q_{A_1A_2}.$$

Parous females go through several gonotrophic cycles. As we neglected the delay induced by the search for a blood meal host, the time between two successive layings is equal to T_{c}, the length of a gonotrophic cycle. Thus, at each time step, the

proportion of parous females that lay their eggs and start a new cycle, denoted q_c, is given by:

$$q_c = \frac{1}{T_c}[1 - q_{A_2D}].$$

Hence, the number of eggs produced by parous females at each time step is:

$$\text{nwf} \times q_c \times A_2,$$

which gives the following expression for the fecundity of parous females:

$$F_{A_2} = \text{nwf} \times q_c.$$

2.5. The population dynamics

The q_{jD} (death), q_{ij} (growth) and F_i (reproduction) are the elements of the transition matrix, denoted Q, established from the life-cycle graph given in Fig. 2. Q is also called the population projection matrix associated to the model.

$$Q = \begin{Bmatrix} q_{w1w1} & 0 & 0 & F_{A_1} & F_{A_2} \\ q_{w1w2} & q_{w2w2} & 0 & 0 & 0 \\ 0 & q_{w2L} & q_{LL} & 0 & 0 \\ 0 & 0 & q_{LA_1} & q_{A_1A_1} & 0 \\ 0 & 0 & 0 & q_{A_1A_2} & q_{A_2A_2} \end{Bmatrix}.$$

We considered that the transition rates (each element of the matrix Q) could be affected by the climate conditions, so we say that Q is climate-dependent and we denote it Q_t. As no delay was incorporated in the model, the state of the mosquito population at time t, only depends on the population state at time $t-1$. Thus, on the current climate conditions:

$$N_{i,t} = Q_t \times N_{i,t-1}.$$

3. Climate dependency

Climatic factors acting on the development of the mosquito populations vary according to the characteristics of the weather. That is particularly true in areas were there is an alternation of seasons (cold and hot or dry and wet). Here we focused on the mosquito populations of tropical zones where dryness is the limiting factor but temperature is not. More precisely, we considered three steps in the mosquito development cycle for which the need of water is crucial: hatching, larva survival and for some species, maturation of eggs. Thus, we introduced in the model a water dependency, by considering the available water quantity in the breeding sites. Let us denote EP that quantity. EP depends on both the pluviometry and on the shape of the breeding sites. So EP needs to be determined for each species considering its specific environment.

3.1. Maturation of eggs

Once the mosquito eggs are deposited by the females, they go through a maturation period corresponding to $w1$. At the end of this period, eggs have acquired the potentiality of hatching and move to the $w2$ stage. Given that for some species, maturation requires a dry period, we introduced in the model the possibility of such a conditional transition from $w1$ to $w2$. We expressed that condition by adding to the corresponding transition rate, a steeply decreasing sigmoid function σ, that represents the egg maturation ability as a function of the trapped water ratio $\text{EP}/\text{EP}_{\max}$:

$$q_{w1w2,t} = \frac{1}{T^{w1}_{\text{rsd}}}(1 - q_{w1D})\sigma\left(\frac{\text{EP}_{t-1}}{\text{EP}_{\max}}\right).$$

As $q_{w1w1} + q_{w1w2} + q_{w1D} = 1$, the proportion of eggs that remains immature, q_{w1w1}, also depends on EP.

The sigmoid function is a simple S-shape function, with a steep slope for intermediate values, an almost flat shape for small values and a saturation level for higher values. It approximates a step function. This sigmoid function σ verifies the following conditions: when $\text{EP}_{t-1} = 0$, the surviving eggs leave stage $w1$ with frequency $1/T^{w1}_{\text{rsd}}$, and when $\text{EP}_{t-1} = \text{EP}_{\max}$, no more eggs move to stage $w2$. So:

$$\sigma\left(\frac{\text{EP}_{t-1}}{\text{EP}_{\max}}\right) = \frac{e^{-\theta_{w1.1}(\text{EP}_{t-1}/\text{EP}_{\max})^{\theta_{w1.2}}} - e^{-\theta_{w1.1}}}{1 - e^{-\theta_{w1.1}}}$$

with parameters $\theta_{w1.1} > 0$ and $\theta_{w1.2} > 0$.

3.2. Hatching

Eggs are laid by the females on the water edge. As the water level in the hole varies during the laying season, eggs are distributed along the cavity walls. Once eggs are mature, they can hatch if they are flooded. Hatching therefore depends on the trapped water quantity. Thus, we modulated the transition rate q_{w2L} by an increasing function of $\text{EP}/\text{EP}_{\max}$:

$$q_{w2L,t} = (1 - q_{w2D}), \qquad \theta_{w2.1}\left(\frac{\text{EP}_{t-1}}{\text{EP}_{\max}}\right)^{\theta_{w2.2}},$$

with parameters $\theta_{w2.1} > 0$ and $\theta_{w2.2} > 0$. So, when $\text{EP}_{t-1} = 0$, then $q_{w2L,t} = 0$ which means that no hatching occurs; when $\text{EP}_{t-1} = \text{EP}_{\max}$, then mature surviving eggs hatch at frequency $\theta_{w2.1}$ which corresponds to the inverse of the mean residence time in $w2$ stage, $1/T^{w2}_{\text{rsd}}$. We chose a power function to represent the climate dependency. Indeed, it is a quite flexible increasing function as it allows: quadratic-like curves when $\theta_{w2.2} > 1$, i.e. curves that increase more rapidly for higher values; or square-root-like curves when $0 < \theta_{w2.2} < 1$, i.e. curves that increase more rapidly for lower values.

3.3. Larva mortality

The main factor contributing to larva mortality is habitat desiccation. Because several studies showed a density-dependence of the larval survivorship (Renshaw et al., 1993; Lord, 1998; Teng and Apperson, 2000; Agnew et al., 2002), we also took into account the density effect by considering the available water by larva, EP/L. So, we expressed the larval mortality q_{LD} as a decreasing function of EP/L, $L \neq 0$, as

follows:

$$q_{LD,t} = q_{LDm} + (1 - q_{LDm})\, e^{-\theta_{L.1}(EP_{t-1}/L_{t-1})^{\theta_{L.2}}}$$

with $\theta_{L.1} > 0$, $\theta_{L.2} > 0$. When $EP_{t-1}/L_{t-1} = 0$, $L_{t-1} \neq 0$, then $q_{LD,t} = 1$ which means that all larvae die; and when EP_{t-1}/L_{t-1}, $L_{t-1} \neq 0$, increases, then $q_{LD,t}$ declines to q_{LDm} which is the minimal mortality rate for the larvae. When there are no larvae, $L = 0$, there are no larvae dying whatever the mortality rate applied. When $L \to 0$, $q_{LD} \to Q_{LDm}$. So for $L = 0$ we set $q_{LD} = q_{LDm}$. In this case, the climate dependency is based on an exponential function, as we assumed that the decrease would be rather steep.

4. Application

We applied the mosquito population model with climate dependency to two specific populations of yellow fever vectors, *Aedes africanus* and *Aedes furcifer* (Cordellier, 1978), which live in forest galleries and in the bordering savannahs in Ivory Coast. Their blood meal hosts are primates and humans. Both mosquito species are active when primates rest in the canopy. They can also leave the canopy and bite humans who come into the forest. Moreover, *A. furcifer* is capable of entering neighbouring villages. This close proximity of both species to their blood meal hosts supports our assumption that the time needed to find a host is negligible compared to the 1-day time step.

Our aim was to reproduce the major trends in the mosquito population abundances, comparing the simulations of the model with field data. To run the model, we had to determine the dynamics of the trapped water, EP_t, and the parameters values. EP_t is strongly related to the type of the habitat and so a specific habitat model was developed. Concerning the parameter values, we assumed that the biological characteristics were known, and we estimated the remaining parameters. With these parameters, simulations were produced by running the generic population model with the specific habitat model. In this section, we successively present the field data, the specific habitat modeling, the parameters, and then, simulations and results. Following that, we include a sensitivity analysis to estimate the contribution of the different parameters to the variability of the model. Then we close this part with a short discussion.

4.1. Field data

Over a 3-year period, Mondet (1994) studied the abundance of mosquito vectors of yellow fever in the Dezidougou area, Ivory Coast, after a serious outbreak of yellow fever in 1982. The village of Dezidougou is located at the limit of the forest and the savannah. In that zone, the climate is divided into two seasons: a dry season lasting from November to April and a wet season with two peaks of rain. In the rainy season, the temperature is rather uniform, with a minimum ranging between 21 and 25 °C and a maximum ranging between 25 and 27 °C. From the first rains, relative humidity reaches a maximum ranging between 95 and 100% and persisting during all the rainy season. Daily rainfall was recorded in Dezidougou from January 1989 to December 1991. Over these 3 years, female mosquitoes were captured every day during a 1-week period, using yellow fever vaccinated human baits spread over 9 sites from 5 to 8 p.m. There was a 30-day interval between the captures during the first 2 years and a 40-day interval the third year. Among the species collected, *A. africanus* and *A. furcifer*, were well represented. Each month, a mean daily abundance was calculated for these two species. These abundances are given by the bars in Fig. 7 with the simulation results.

4.2. Specific habitat modelling

Breeding sites of both populations, *A. africanus* and *A. furcifer*, are tree holes of variable capacities supplied with rainwater. During each rainfall, the rainwater that streams is trapped until saturation of the tree holes. We considered a tree hole as a cone. Let us denote EP_{max} the maximal capacity of the tree hole, $\pi \times R^2$ its basis and H its height (Fig. 3a). Then, EP_{max} is equal to:

$$EP_{max} = \pi \times R^2 \times \frac{H}{3}.$$

Let EP_{t-1} be the trapped water in the cone at time $t-1$. The radius of the water surface at time $t-1$ is:

$$r_{t-1} = \sqrt[3]{K\frac{3EP_{t-1}}{\pi}}, \quad \text{where } K = \frac{R}{H}.$$

The water surface at time $t-1$ is: $S_{t-1} = \pi(r_{t-1})^2$.

To determine the trapped water at time t, we need to consider both rainfall and evaporation that take place during the time step. We first consider evaporation (Fig. 3a and b) that is related to the surface of the water S. The larger S is, the greater the evaporation. So we express the proportion of water evaporated between time $t-1$ and t, denoted $\theta_{evap,t}$, as an increasing function of S_{t-1}:

$$\theta_{evap,t} = 1 - e^{-\rho_{evap} S_{t-1}}, \quad \text{with } \rho_{evap} \geq 0.$$

Then, we introduce the rainfall. Let P_t be the height of the rainfall between time $t-1$ and time t. The volume of the rainfall getting into the cone is approximated by a cylinder of

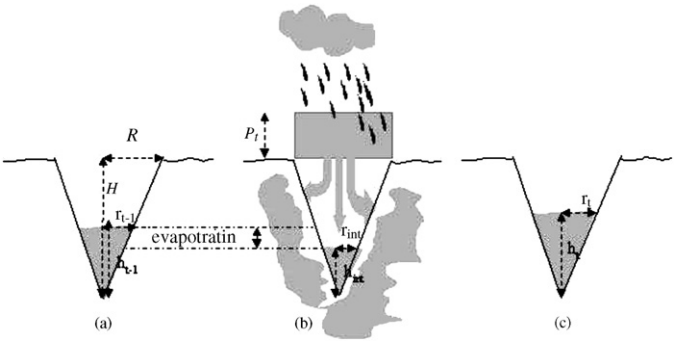


Fig. 3. Successive phases of the tree-hole filling up process: evaporation and rainwater trapping between time $t-1$ and time t. (a) Time $t-1$, (b) between evaporation and rainwater trapping and (c) time t.

basis $\pi \times R^2$ and height P_t (Fig. 3b and c). At this point, we have to take into account the part of this rainwater that seeps through the walls. The more the hole is filled, the less infiltration occurs. Thus, we express the fraction of the rain effectively trapped in the tree hole, denoted $\theta_{\text{eff},t}$, as a decreasing function of the difference between the basis of the cone and the surface of the water trapped in the cone after evaporation. We denote $\Delta S_{t-1,t}$ that difference. So we have:

$$\theta_{\text{eff},t} = \mathrm{e}^{-\rho_{\text{eff}}\,\Delta S_{t-1,t}}, \quad \text{with } \rho_{\text{eff}} > 0.$$

Taking into account the fact that EP is bounded by $\text{EP}_{\max}$, then EP_t, the trapped water at time t, is given by:

$$\text{EP}_t = \min((\text{EP}_{t-1}(1-\theta_{\text{evap},t}) + \theta_{\text{eff},t}\pi R^2 P_t), \text{EP}_{\max})$$

Combining this trapped water dynamics with the generic population model, we obtain a fully defined tree-hole breeding mosquito model.

4.3. Parameters

The biological characteristics of the two mosquito species are given in Table 1. These biological features are averages coming from references of the literature, biological surveys or expert advice. We also fixed the size of the cone-shaped tree hole, the depth being equal to 80 mm and radius to 30 mm, for both species. The resulting capacity is 75 ml which is a reasonable value for a tree hole. The remaining parameters of the model, i.e. the parameters related to the climate influence, were estimated from the field data.

4.4. Simulations and parameter estimation

To run the tree-hole breeding mosquito model we used the parameters described in Table 1 and the rainfall data from the Dezidougou area. We set the initial conditions: at time 0, we considered that there were only immature eggs in the tree hole, and we fixed that number $w1_0$ to 10 eggs. So $N_0 = (w1_0, 0, 0, 0, 0)$, with $w1_0 = 10$. We set the time step to 1 day. We then had to estimate the climate-dependent parameters. Comparing the simulated mosquito abundances with field data gathered over the first 2 years, i.e. from January 1989 to December 1990, we estimated the optimal values for these parameters by minimizing a least square criterion. A scaling was needed for this fit: abundance data represent a mean daily number of mosquitoes caught on a given area, whereas the simulations reproduce the daily number of female mosquitoes breeding in a given tree hole. We then simulated a third year using the estimated parameters. The third year of mosquito abundance data was hence used only to validate the model. All simulations were performed using the R software.

Table 2
Parameters of the model estimated from observed data over 1989–1990

Parameters of the model	*Aedes africanus*	*Aedes furcifer*
ρ_{evap}	0.00005	0.00004
ρ_{eff}	2.27	1.61
$\theta_{w1.1}$	–	15.7
$\theta_{w1.2}$	–	0.86
$\theta_{w2.1}$	0.03	0.06
$\theta_{w2.1}$	0.67	5.90
$\theta_{L.1}$	0.25	1.96
$\theta_{L.2}$	0.28	0.07

ρ_{evap} and ρ_{eff} concern, respectively, the evaporation and the filling of the tree hole. The θ_x parameters are related to the influence of climate on mosquito development at three ages: immature eggs ($x = w1$), mature eggs ($x = w2$), larvae ($x = L$).

4.5. Results

Simulations were produced by running the model over the 3 years, using the parameter values in Table 1 and the estimated parameters given in Table 2.

The first estimated parameters ρ_{evap} and ρ_{eff}, related to the filling of the tree hole are, respectively, involved in the

Table 1
Biological characteristics and simulation parameters introduced in the model

	Parameters	*Aedes africanus*	*Aedes furcifer*
Biological characteristics			
Residence time in $w1$ stage (days)	T^{w1}_{rsd}	7	7
Mortality rate in $w1$ stage	$q_{w1\text{D}}$	0.05	0.01
Mortality rate in $w2$ stage	$q_{w2\text{D}}$	0.05	0.01
Residence time in L stage (days)	T^{L}_{rsd}	12	16
Minimal mortality rate in L stage	$q_{L\text{Dm}}$	0.05	0.01
Cycle length between two layings (days)	T_{c}	7	7
Mortality rate in A_1 stage	$q_{A_1\text{D}}$	0.05	0.01
Mortality rate in A_2 stage	$q_{A_2\text{D}}$	0.15	0.06
Number of female eggs/female	nwf	40	30
Habitat characteristics			
Tree hole depth (mm)	H	80	80
Radius of the tree hole basis (mm)	R	30	30
Simulation conditions			
Initial population (at time $t = 0$)	N_0	(10, 0, 0, 0, 0)	(10, 0, 0, 0, 0)
Time step (days)	Δt	1	1

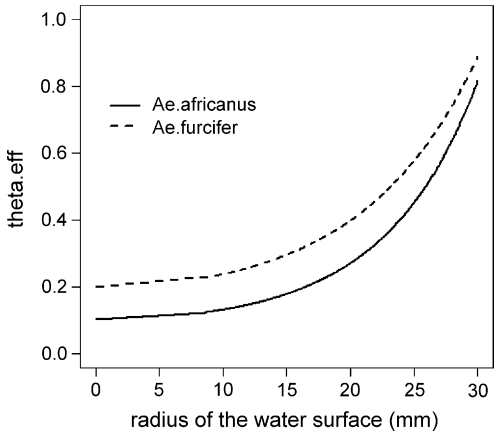


Fig. 4. Efficient fraction of rainfall θ_{eff} in function of the radius of the trapped water surface r for each species.

expression of θ_{evap} and θ_{eff}. We note that the evaporation rates θ_{evap} are very similar for both species and never exceed 0.2, which is usual in forest area. The efficient fraction of rainfall θ_{eff} expressed in function of the water surface in the holes is slightly higher for *A. furcifer* than for *A. africanus*, as shown in Fig. 4. Consequently, the trapped water dynamics represented in Fig. 5 are a little different in the two breeding sites.

The parameters in Table 2 related to the climate influence on the transition rates are different between the two mosquito species. Consequently, the impact of the trapped water EP on these transition rates also differs between the two species, as shown in Fig. 6. For the egg maturation, no trapped water effect was introduced for *A. africanus* (Fig. 6a). However, for *A. furcifer*, the maturation rate rapidly decreases when the tree hole fills in Fig. 6b. The hatching rate increases with EP for both

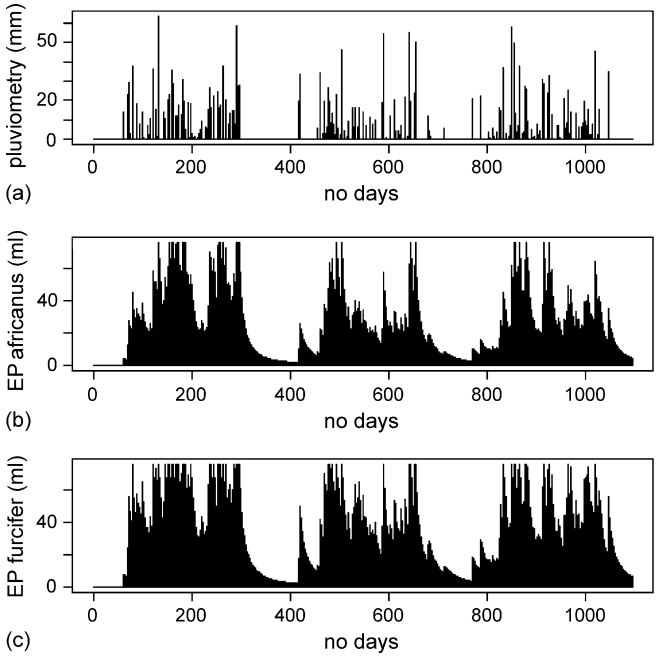


Fig. 5. (a) Pluviometry in Dezidougou from January 1989 to December 1991; and related trapped water (EP) dynamics in the cone-shaped tree hole of (b) *Aedes africanus* and (c) *Aedes furcifer*.

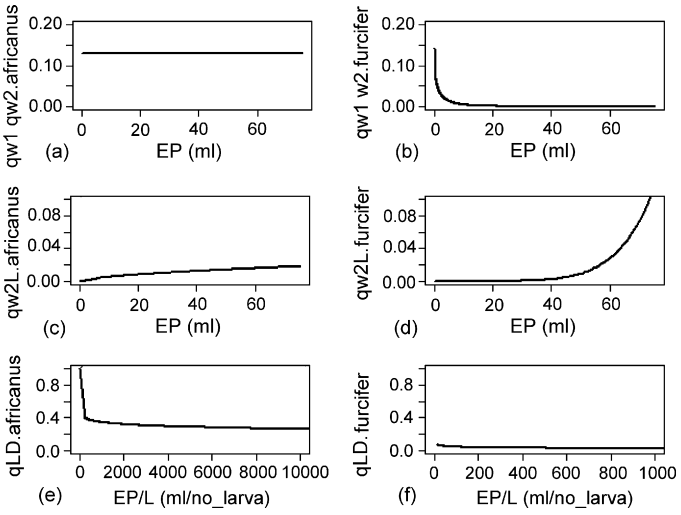


Fig. 6. Climate-dependent transition rates as functions of the tree hole trapped water EP (a–d) or trapped water available by larva EP/L (e and f): (a and b) immature to mature egg transition rates, (c and d) hatching rates, (e and f) larva mortality. The graphs on the left correspond to *Aedes africanus* (a, c and e); the graphs on the right to *Aedes furcifer* (b, d and f).

species. For *A. furcifer* the growth starts for higher levels of EP, but the increase is then steeper than for *A. africanus* (Fig. 6c and d). The *A. furcifer* larva mortality is almost always to its minimal value q_{LDm} (Fig. 6f): neither the trapped water quantity, nor the larva density have much impact on this value. The *A. africanus* larva mortality shows a stronger density dependence: the larva mortality decreases with the trapped water available by larva.

The mosquito abundance data and simulations are represented in Fig. 7. Note that the scales of the two graphics are different, as *Ae. africanus* is globally twice as abundant as *Ae. furcifer*. The estimation was performed over the first 2 years and the resulting abundances fit the field data well. Simulations over the third year correspond to predictions: for both species we notice an overestimate. Graphs in Fig. 7 show that the abundance fluctuations of both mosquito species roughly follow the fluctuations of the trapped water in the tree holes

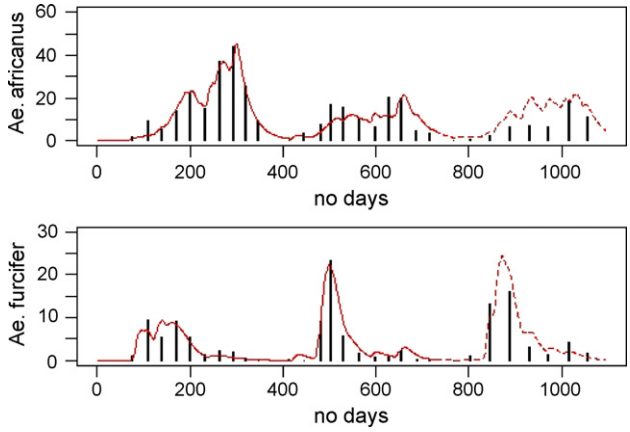


Fig. 7. Abundance of *Aedes africanus* and *Aedes furcifer* female adults ($A_1 + A_2$). The bars correspond to the number of adult mosquitoes caught in a day in Dezidougou (average over a week calculated once a month). The lines represent the simulation results: solid lines correspond to the fitted abundances and dashed lines correspond to the predictions.

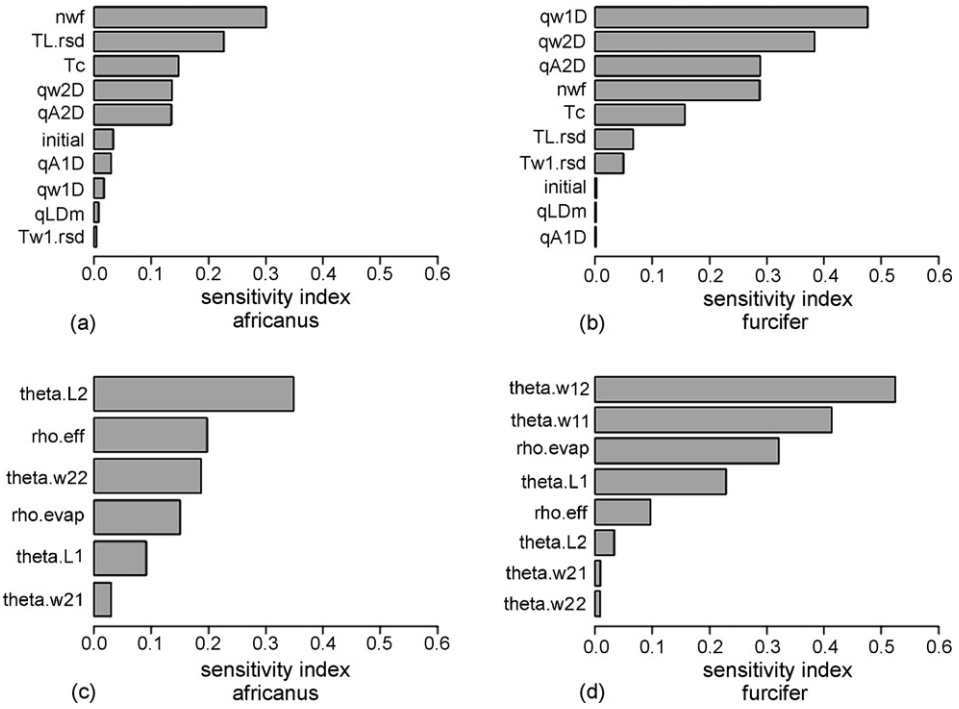


Fig. 8. Sensitivity indexes of the biological parameters and the initial egg population (a and b) and the estimated parameters (c and d). Graphics (a and c) concern *Ae. africanus*, graphics (b and d) concern *Ae. furcifer*. The sensitivity indexes were obtained after a ±20% variation of the reference values shown in Table 1 (for a and b) and in Table 2 (for c and d).

(Fig. 5). However, the abundance peaks of the two species are shifted in time. The population of *A. africanus* increases along the rainy season whereas the population of *A. furcifer* rises at the beginning of the rainy season then declines. Moreover, the amplitude of the abundance fluctuations is higher for *A. africanus* than for *A. furcifer*.

4.6. Sensitivity analysis

To determine which parameters have the greatest impact on the model, we carried out a sensitivity analysis on the biological characteristics, the initial number of eggs and the estimated parameters. Because the biological characteristics are needed to estimate the remaining parameters, we performed two separate analyses. In both cases, parameters were increased and decreased by 20% from their reference values given in Tables 1 and 2. To explore this parameter space, we used a complete experimental design. The output variable, denoted Y, corresponds to the distance between the simulated abundances $A = A_1 + A_2$ and the field data O, calculated at the observation times t_o as follows:

$$Y = \sum_{t_o} \left(\frac{A_{t_o} - O_{t_o}}{O_{t_o}} \right)^2.$$

The sensitivity index is defined for each parameter i by:

$$S_i = \frac{V(Y) - V_{-i}}{V(Y)}$$

where $V(Y)$ is the total variance and V_{-i} is the sum of all variance terms that do not include the parameter i. The higher the index, the more influence the parameter has. The results are given by the barplots in Fig. 8.

Globally, the *A. furcifer* model is more sensitive to parameter variations than the *A. africanus* model. For both species, the mean number of female eggs produced by one female, the mortality rates in the $w2$ and A_2 stages and the cycle length between two layings have a major impact on the model output. The residence time in the L stage specifically affects the *A. africanus* model, whereas the mortality in the $w1$ stage only affects the *A. furcifer* model. The other biological parameters and the initial condition do not show a great influence.

For both species, the estimated parameters related to the trapped water dynamics show a moderate influence on the model outputs. $\theta_{w1.1}$ and $\theta_{w1.2}$, involved in the maturation of *A. furcifer* eggs, have a high sensitivity index. θ_{L2} involved in the larva mortality has the most impact on the *A. africanus* model whereas it does not affect the *A. furcifer* model; in this latter model, θ_{L1} has a moderate influence.

These additional simulations and the analyses were performed using the R software.

4.7. Discussion

The model we developed for the *A. africanus* and *A. furcifer* in Ivory Coast was able to reproduce the major trends in the adult population abundances. The third simulation year, which was predicted from parameters estimated over the first 2 years, did not show such a good fit. It caught the global patterns but overestimated the abundances, both for *A. africanus* and *A. furcifer*. Many explanations may hold: one could think that because of a lower capture frequency the third year, field data

did not catch the abundance peaks, or that the mosquitoes that year experienced a mortality increase ....

For both species, the most influent parameters are those that relate to the production and survival of eggs. However, the initial condition has very little impact, which implies that the population development does not only depend on the egg survival from the previous rainy season. Each year, the abundances of each species show a distinct pattern. These different trends are explained by the fact that the eggs of *A. furcifer* need a dry season to mature: eggs laid during the rainy season become mature during the next dry season and are therefore accumulated in this stage, until new rainfalls. The sensitivity analysis showed that this maturation process is critical in the population development. The very low larva mortality rate obtained for *A. furcifer* suggests that the larvae need very little water to survive. Therefore, once the larvae hatch, which only occurs when sufficient water is available in the tree hole, they are quite resilient to environmental fluctuations. *A. africanus* larvae seem to be more affected by larva density and water availability, which is supported by the sensitivity analysis (θ_{L2} has a high index). It could explain why the adult abundance fluctuations observed seem to follow more closely the trapped water evolution.

No field data were available to define the tree holes. From expert advice we chose a conical shape as one of the simplest representation of a rather deep and narrow hole. Moreover, we were able to fix the hole capacity to a reasonable value (75 ml). Keeping the same fixed capacity and varying the shape of the cone (width and depth), the results were not modified. The only differences observed were the estimations of the parameters controlling the filling of the tree hole, ρ_{evap} and ρ_{eff}, which fluctuated slightly. These two parameters have a moderate impact on the adult population abundances. However, ρ_{eff} has different values for both species, which leads us to assume that t may not have the same type of breeding sites. The retention of rainwater in the tree hole appears higher for *A. furcifer*: it could be due to a different shape of the tree hole or a different species of tree. Indeed, various types of breeding sites are available in the wild and mosquitoes might show preferences. It is the case for *Aedes usambara*, vector of the Chikungunya virus, which was caught inside bamboo plantations in Ivory Coast (Mondet and Montange, 1993). So to simulate the abundance of *A. usambara*, we would have to consider the breeding habitat as a cylinder-shaped hole to represent the EP process.

Our results suggest that the field populations of *A. africanus* and *A. furcifer* are well simulated by a stage-structured matrix model incorporating pluviometry influence. In that area, as far as we know, pluviometry is the only environmental limiting factor in the *Aedes* population development. Other factors, such as temperature, are not considered as limiting factors for both species in that area, and thus were not introduced in the model. However, they may also interact with pluviometry and modulate the mosquito population abundance. Nevertheless, this simple pluviometry-based model allowed us to capture the major trends in the population fluctuations, which was our aim in this application.

In Dezidougou the yellow fever is not endemic. The virus migrates from the southern forests during the rainy season, carried by primate and mosquito populations along the forest galleries. So the virus usually appears in savannah at the end of the rainy season. Both data and model agree on the following trend: the population of *A. africanus* is higher at the end of the rainy season whereas the population of *A. furcifer* shows a peak at the beginning of the season. Therefore, high *A. africanus* abundances increase the risk of a yellow fever epidemic in the monkey population of the Dezidougou area. But YF virus can survive in some *A. furcifer* eggs during the dry season. These females born of contaminated eggs will be capable to generate a human YF outbreak at the beginning of the following rainy season, especially with abundant rainfalls that lead to a high peak of *A. furcifer*.

5. Conclusions

We have presented a generic mosquito population model that includes several features: various development stages, climate and density dependency. Results obtained by applying the model to field data have shown that it allows a rather good assessment of the mosquito abundance trends. Indeed, by taking into account the rain profile in the demographic process, we have reproduced over 2 years, and predicted over 1 year, the population evolution of two species of mosquitoes, that differ by their ecological and biological features. The model fitted the data quite accurately during the first 2 years; it overestimated the abundances the third year, but reproduced the global patterns. In this deterministic model, pluviometry fluctuations were sufficient to capture the major trends in the population abundances. However, we may introduce stochastic factors to account for the biological variability. As we introduced the effect of trapped water, which is the main factor under inter-tropical climates for tree-hole breeding mosquitoes, we can also introduce other factors, such as the temperature for example. We can also reduce or expand the number of stages. Although, the generic nature of the model allows the incorporation of many other factors, we have to keep in mind the parsimony rule before refining the model.

Applying the model to real abundance and rainfall data, we have shown its predictive capacity once parameters are estimated. Therefore, this model could prove useful to estimate the abundance evolution of disease vectors, according to different climate change scenarios. Thus, such a model is a valuable tool to assess the risk of an epidemiological outbreak.

Appendix A. Filling up process of the tree hole

Let us consider a tree hole as an inverted cone with πR^2 its basis and H its depth.

At each time step, the rainwater trapped in the tree hole, EP_t is updated. The filling up process is composed of two successive phases: evaporation and rainwater trapping (Fig. 3).

The volume of the water that evaporates between $t-1$ and t is $\theta_{\text{evap},t} \times \text{EP}_{t-1}$, where $\theta_{\text{evap},t}$ is the evaporation rate. This rate depends on the water surface at $t-1$, $S_{t-1} = \pi r_{t-1}^2$. The

larger S_{t-1}, the greater the evaporation. Moreover, $\theta_{\text{evap},t}$ being a fraction, it is bounded between 0 and 1. So we express the evaporation rate, $\theta_{\text{evap},t}$, between time $t-1$ and time t as $\theta_{\text{evap},t} = 1 - e^{-\rho_{\text{evap}} S_{t-1}}$, with $\rho_{\text{evap}} \geq 0$.

Let us denote EP_{int} the trapped water quantity remaining in the tree hole after the evaporation process is applied between $t-1$ and t:

$$\text{EP}_{\text{int}} = \text{EP}_{t-1}(1 - e^{-\rho_{\text{evap}} S_{t-1}}).$$

This volume can also be expressed as $\text{EP}_{\text{int}} = (1/3)\pi r_{\text{int}}^2 h_{\text{int}}$ where r_{int} and h_{int} are, respectively, the radius and the height of the water remaining in the hole.

Let us fix $K = R/H$. Using the cone property $R/H = r_{\text{int}}/h_{\text{int}}$ we can write that:

$$\text{EP}_{\text{int}} = \frac{1}{3}\pi \frac{r_{\text{int}}^3}{K}.$$

We deduce that:

$$r_{\text{int}} = \left(\frac{3K}{\pi}\text{EP}_{\text{int}}\right)^{1/3} = \left(\frac{3K}{\pi}\text{EP}_{t-1}(e^{-\rho_{\text{evap}} S_{t-1}})\right)^{1/3}$$

The second phase concerns the rainwater trapping. Let P_t be the rainfall height between time $t-1$ and time t. The rainwater that comes into the tree hole is approximated by a cylinder $\pi R^2 P_t$. Because of infiltration, only a fraction $\theta_{\text{eff},t}$ of this water is trapped. $\theta_{\text{eff},t}$ depends on $\Delta S_{t-1,t}$, the normalized difference between the basis of the cone, πR^2, and the trapped water surface in the cone after evaporation, πr_{int}^2:

$$\Delta S_{t-1,t} = \frac{\pi R^2 - \pi r_{\text{int}}^2}{\pi R^2} = \frac{R^2 - r_{\text{int}}^2}{R^2}.$$

The bigger this surface, the more infiltration there is. Moreover, $\theta_{\text{eff},t}$ being a fraction, it is bounded between 0 and 1. So we express $\theta_{\text{eff},t}$ as:

$$\theta_{\text{eff},t} = e^{-\rho_{\text{eff}} \Delta S_{t-1,t}}, \quad \text{with } \rho_{\text{eff}} \geq 0.$$

After evaporation and rainwater trapping, the volume of water in the tree hole is:

$$\begin{aligned}\text{EP}_t &= \text{EP}_{\text{int}} + (\theta_{\text{eff},t} \times \pi R^2 P_t) \\ &= \text{EP}_{t-1}(1 - \theta_{\text{evap},t}) + (\theta_{\text{eff},t} \times \pi R^2 P_t).\end{aligned}$$

Finally, EP_t, the trapped water at time t, is bounded by the tree hole capacity $\text{EP}_{\max} = \pi \times R^2 \times (H/3)$. So, we obtain:

$$\text{EP}_t = \min((\text{EP}_{t-1}(1 - \theta_{\text{evap},t}) + \theta_{\text{eff},t}\pi R^2 P_t), \text{EP}_{\max}).$$

References

Agnew, P., Hide, M., Sidobre, C., Michalakis, Y., 2002. A minimalist approach to the effects of density-dependent competition on insect life-history traits. Ecol. Entomol. 27, 396–402.

Ahumada, J.A., Lapointe, D., Samuel, M.D., 2004. Modeling the population dynamics of *Culex quinquefasciatus* (Diptera: Culicidae), along an elevation gradient in Hawaii. J. Med. Entomol. 41, 1157–1170.

Awerbuch-Friedlander, T., Levins, R., Predescu, M., 2005. The role of seasonality in the dynamics of deer tick populations. Bull. Math. Biol. 67, 467–486.

Bommarco, R., 2001. Using matrix models to explore the influence of temperature on population growth of arthropod pests. Agric. Forest Entomol. 3, 275–283.

Caswell, H., 2001. Matrix Population Models: Construction, Analysis and Interpretation, second ed. Sinauer Ass. Inc. Publishers.

Chattopadhyay, J., Sarkar, R.R., Chaki, S., Bhattacharya, S., 2004. Effects of environmental fluctuations on the occurrence of malignant malaria—a model based study. Ecol. Model. 177, 179–192.

Chevalier, V., de la Rocque, S., Baldet, T., Vial, L., Roger, F., 2004. Epidemiological processes involved in the emergence of vector-borne diseases: West Nile fever, Rift Valley fever, Japanese encephalitis and Crimean-Congo haemorrhagic fever. Revue Scientifique Et Technique De L Office International Des Epizooties 23, 535–555.

Choi, W.I., Ryoo, M.I., 2003. A matrix model for predicting seasonal fluctuations in field populations of *Paronychiurus kimi* (Collembola: Onychiuridae). Ecol. Model. 162, 259–265.

Cordellier, R., 1978. Les vecteurs potentiels sauvages dans l'épidémiologie de la fièvre jaune en Afrique de l'Ouest. Travaux et documents ORSTOM, vol. 81, Paris, 1978, 258 pp.

Costantino, R.F., Cushing, J.M., Dennis, B., Desharnais, R.A., 1998. Resonant population cycles in temporally fluctuating habitats. Bull. Math. Biol. 60, 247–273.

Depinay, J.M.O., Mbogo, C.M., Killeen, G., Knols, B., Beier, J., Carlson, J., Dushoff, J., Billingsley, P., Mwambi, H., Githure, J., Toure, A.M., McKenzie, F.E., 2004. A simulation model of African Anopheles ecology and population dynamics for the analysis of malaria transmission. Malaria J. 3.

Derouich, M., Boutayeb, A., Twizell, E.H., 2003. A model of dengue fever. BioMed. Eng. Online 2, 4.

Esteva, L., Vargas, C., 1999. A model for dengue disease with variable human population. J. Math. Biol. 38, 220–240.

Focks, D.A., Haile, D.G., Daniels, E., Mount, G.A., 1993. Dynamic life table model for *Aedes-Aegypti* (Diptera, Culcidae)—analysis of the literature and model development. J. Med. Entomol. 30, 1003–1017.

Fouque, F., Baumgartner, J., 1996. Simulating development and survival of *Aedes vexans* (Diptera: Culicidae) preimaginal stages under field conditions. J. Med. Entomol. 33, 32–38.

Ghosh, M., Pugliese, A., 2004. Seasonal population dynamics of ticks, and its influence on infection transmission: a semi-discrete approach. Bull. Math. Biol. 66, 1659–1684.

Glass, K., 2005. Ecological mechanisms that promote arbovirus survival: a mathematical model of Ross River virus transmission. Trans. R. Soc. Trop. Med. Hyg. 99, 252–260.

Ishikawa, H., Ishii, A., Nagai, N., Ohmae, H., Harada, M., Suguri, S., Leafasia, J., 2003. A mathematical model for the transmission of *Plasmodium vivax* malaria. Parasitol. Int. 52, 81–93.

Jarry, M., Khaladi, M., Gouteux, J.P., 1996. A matrix model for studying tsetse fly populations. Entomol. Exp. Appl. 78, 51–60.

Jensen, A.L., 1995. Simple density-dependent matrix model for population projection. Ecol. Model. 77, 43–48.

Lord, C.C., 1998. Density dependence in larval *Aedes albopictus* (Diptera: Culicidae). J. Med. Entomol. 35, 825–829.

Lord, C.C., 2004. Seasonal population dynamics and behaviour of insects in models of vector-borne pathogens. Physiol. Entomol. 29, 214–222.

Mazaris, A.D., Matsinos, Y.G., 2006. An individual based model of sea turtles: Investigating the effect of temporal variability on population dynamics. Ecol. Model. 194, 114–124.

Mondet, B., 1994. Age physiologique et biologie des *Aedes* (Diptera: Culicidae) vecteurs potentiels de fièvre jaune en Côte d'Ivoire. Implications épidémiologiques. Thesis. Paris-Sud University.

Mondet, B., 1997. Importance of Aedes (Diceromyia) furcifer Edwards, 1913 (Diptera: Culicidae) among the arbovirus potential vectors, in the epidemiology of the yellow fever, in the sub-Sudanese savannas of Ivory Coast. Annales de la Societe Entomologique de France 33, 47–54.

Mondet, B., Montange, L., 1993. Presence of Aedes (Stegomyia) usambara Mattingly, 1953 (Diptera: Culicidae) in the Ivory Coast, seasonal variations

and epidemiological position. Annales de la Societe Entomologique de France 29, 261–267.

Moon, T.E., 1976. Astatistical model of the dynamics of a mosquito vector (*Culex tarsalis*) population. Biometrics 32, 355–368.

Ngwa, G.A., 2004. Modelling the dynamics of endemic malaria in growing populations. Dis. Cont. Dyn. Syst. Ser. B 4, 1173–1202.

Ngwa, G.A., Shu, W.S., 2000. A mathematical model for endemic malaria with variable human and mosquito populations. Math. Comput. Model. 32, 747–763.

Renshaw, M., Service, M.W., Birley, M.H., 1993. Density-dependent regulation of *Aedes-Cantans* (Diptera, Culicidae) in natural and artificial populations. Ecol. Entomol. 18, 223–233.

Shone, S.M., Curriero, F.C., Lesser, C.R., Glass, G.E., 2006. Characterizing population dynamics of *Aedes sollicitans* (Diptera: Culicidae) using meteorological data. J. Med. Entomol. 43, 393–402.

Teng, H.J., Apperson, C.S., 2000. Development and survival of immature *Aedes albopictus* and *Aedes triseriatus* (Diptera: Culicidae) in the laboratory: effects of density, food, and competition on response to temperature. J. Med. Entomol. 37, 40–52.

Thomas, L., Buckland, S.T., Newman, K.B., Harwood, J., 2005. A unified framework for modelling wildlife population dynamics. Aust. N. Z. J. Stat. 47, 19–34.

Tomori, O., 2004. Yellow fever: the recurring plague. Crit. Rev. Clin. Lab. Sci. 41, 391–427.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 433–438
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Molecular characterization of Thai *Ehrlichia canis* and *Anaplasma platys* strains detected in dogs☆

Danai Pinyoowong [a], Sathaporn Jittapalapong [b], Fanan Suksawat [c], Roger W. Stich [d], Arinthip Thamchaipenet [a,*]

[a] *Department of Genetics, Faculty of Science, Kasetsart University, Thailand*
[b] *Department of Parasitology, Faculty of Veterinary Medicine, Kasetsart University, Thailand*
[c] *Department of Medicine, Faculty of Veterinary Medicine, Khon Kaen University, Thailand*
[d] *Department of Veterinary Pathobiology, University of Missouri, Columbia, MO, USA*

Received 31 March 2007; accepted 12 June 2007
Available online 19 June 2007

Abstract

Canine monocytic ehrlichiosis caused by *Ehrlichia canis* is of veterinary importance worldwide. In Thailand, there has been little information available on *E. canis* and its phylogeny. The objective of this study was to characterize and establish molecular structure and phylogeny of Thai *Ehrlichia* and *Anaplasma* strains. Genus-specific primers for *Ehrlichia* and *Anaplasma* were used to amplify the 16S rRNA gene from naturally infected canine blood samples, and these amplicon sequences were compared with other sequences from GenBank. Both homology and secondary structure analysis of 16S rRNA sequences indicated that they were novel *E. canis* and *A. platys* strains. Phylogenetic analysis revealed that the Thai *E. canis* strain was closely related and formed a single cluster with *E. canis* from different countries. *A. platys* found in this study showed close relationship with earlier report of *A. platys* from Thailand. To our knowledge this report represents the first molecular characterization of the nearly complete 16S rRNA gene from *E. canis* in dogs from Thailand.

Keywords: *Ehrlichia canis*; *Anaplasma platys*; 16S rDNA; Phylogenetic tree; Dogs; Bangkok; Thailand

1. Introduction

Canine monocytic ehrlichiosis (CME) and canine cyclic thrombocytopenia (CCT) are caused by *Ehrlichia canis* and *Anaplasma platys*, respectively. These bacteria are classified in the rickettsial family Anaplasmataceae, which includes obligate intracellular prokaryotic parasites that reside within a parasitophorous vacuole (Dumler et al., 2001). In canine hosts, *E. canis* is infective for monocytes while *A. platys* infect platelets (Greene and Harvey, 1990). *Rhipicephalus sanguineus* ticks are considered the primary vector of both pathogens (Groves et al., 1975; Greene and Harvey, 1990). Globally distributed, well-characterized pathogens such as *E. canis* offer unique opportunities to study coevolution and interaction between tick-borne pathogens and their vertebrate and invertebrate hosts (Stich et al., in press).

Although CME and CCT are considered enzootic throughout Thailand, these conclusions are based on diagnoses that rely on clinical signs, haematological abnormalities and microscopic examination of peripheral blood. Thus, these diagnoses are often ambiguous and may fail to identify the pathogen species involved. Molecular diagnostic methods allow direct detection of these etiologic agents and sequence analysis facilitates their comparison to geographically diverse strains. To our knowledge, detailed genetic and phylogenetic information about *E. canis* and *A. platys* in Thailand are limited to a single report of an *A. platys* 16S rRNA gene (16S rDNA) sequence (Suksawat et al., 2001).

In this study, we investigated diagnosed cases of canine ehrlichiosis from a private laboratory in Bangkok, to confirm the presence of *E. canis* and *A. platys* and to compare Thai strains to those from other regions. PCR and 16S rDNA sequence analysis were used to characterize the molecular

☆ This study was presented in part as an abstract at the MEEGID VIII Congress, Bangkok, Thailand, 30 November to 2 December 2006.

* Corresponding author. Tel.: +66 2 562 5444x4208; fax: +66 2 579 5528.
E-mail address: arinthip.t@ku.ac.th (A. Thamchaipenet).

1567-1348/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2007.06.002

features of these infections, which allowed identification of a novel strain of *A. platys* that differed from that of the previous report of Suksawat et al. (2001). Moreover, *E. canis* from this country was genetically verified for the first time.

2. Materials and methods

2.1. DNA extraction from canine blood

Blood samples were collected with EDTA from dogs diagnosed with clinical ehrlichiosis and submitted to a private laboratory in Bangkok. The blood was chilled until arrival at the laboratory and then stored at −20 °C. Total DNA was isolated from 100 μl of thawed blood using the phenol/chloroform extraction method of Sambrook and Russell (2001).

2.2. Primers designed for 16S rDNA amplification

Primers for amplification of *Ehrlichia* and *Anaplasma* 16S rDNA were designed from nucleotide sequences deposited in GenBank database (DQ342324, AF414873, AF414870, AF414869, AB211163, U23503, CR767821, U96436, AB196302 and AF318946). All the sequences were aligned for the maximum homology by ClustalW Version 8.1 (Thompson et al., 1994). Conserved regions were selected and specific oligonucleotide primers named ATT062F (5′CCTGGCTCA–GAACGAACGCT3′) and ATT062R (5′GATCCAGCCGCA-GGTTCACC3′) were derived.

2.3. 16S rDNA amplification and sequencing

DNA isolated from dog blood was used as a template to amplify the majority of the 16S rRNA gene from *Ehrlichia* and *Anaplasma* spp. by PCR. Amplifications were generated using 3–6 μl of genomic DNA with primers ATT062F and ATT062R for each 20 μl PCR reaction mixture in a Peltier thermal cycler (MJ Research, Watertown, MA, USA), by 30 cycles of 30 s at 94.0 °C, 30 s at 64.0 °C, and 1 min at 72.0 °C, preceded by 4 min at 94.0 °C and followed by 4 min at 72.0 °C. PCR products were examined by 0.8% agarose gel electrophoresis and ethidium bromide staining. Amplicons were purified with QIAquick PCR purification kits (QIAGEN) according to the manufacturer's protocol. Direct sequencing was carried on the PCR product using the same PCR primers, ATT062F and ATT062R. In order to cover both strands of the PCR fragment completely, another set of primers, ATT066F (5′CCCTGGTAGTCCACGCTG3′) and ATT067R (5′CAGC-GTGGACTACCAGGG3′), were designed for annealing in the middle of the 16S rDNA amplicons and used for sequencing towards 3′ and 5′ ends, respectively.

2.4. Phylogenetic analysis

Multiple sequence alignments of amplicons and 16S rDNA sequences from GenBank were performed using the ClustalW Version 1.8 (Thompson et al., 1994). Phylogenetic trees were inferred using neighbor-joining (NJ) analysis by MEGA software Version 3.1 (Kumar et al., 2004). The distance matrix of nucleotide divergences was calculated according to Kimura's two-parameter model furnished by MEGA. A bootstrap re-sampling technique of 1000 replications was performed to statistically support the reliabilities of the nodes on the trees. The differences between nucleotide positions were confirmed by DnaSP Version 4.10 (Rozas et al., 2003). *Neorickettsia sennetsu* was used as an outgroup.

2.5. Nucleotide sequences accession numbers

The 16S rDNA sequences from this study were deposited at GenBank database under accession numbers EF139458 for *E. canis*-Bangkok and EF139459 for *A. platys*-Bangkok. Other Anaplasmataceae 16S rDNA sequences from Genbank (and their accession numbers) included 10 *E. canis* strains reported from China (AF162860), Japan (AF536827), Peru (DQ915970), South Africa (U54805), Spain (AY394465), Venezuela (AF373613, AF373612) and the USA (M73226, M73221), eight *A. platys* strains from China (AF156784), France (AF303467), Japan (AY077619, AF536828), Spain (AY530806), Thailand (AF286699), Venezuela (AF287153) and the USA (M82801), *E. chaffeensis* (U23503), *E. ewingii* (U96436), *E. muris* (NC_007354, AB196302), *E. ruminantium* (CR767821), *A. bovis* (AB211163), *A. centrale* (AF414869), *A. marginale* (AF414873), *A. ovis* (AF414870), *A. phagocytophilum* (DQ342324) and *N. sennetsu* (M73219).

3. Results

3.1. Amplification and sequencing of 16S rDNA

The identity of agents associated with diagnosed CME and CCT are often assumed based on clinical diagnosis and microscopic examination of peripheral blood. Blood samples from dogs diagnosed with clinical ehrlichiosis were screened to compare these agents to previously reported strains. Template was prepared and assayed by PCR with primers ATT062F and ATT062R, which were designed for specific amplification of *Ehrlichia* and *Anaplasma* 16S rDNA. Approximately 1.5 kb amplicons corresponding to the expected size of targeted 16S rRNA gene fragments were obtained (data not shown). Amplicons from two individual canine blood samples were purified and directly sequenced with the same primers. The nearly complete 16S rRNA gene sequences of 1478 (51% AT) and 1481 (52% AT) bp were identical to consensus 16S rDNA sequences for *E. canis* and *A. platys*, respectively. These sequences were named *E. canis*-Bangkok and *A. platys*-Bangkok and deposited as new 16S rDNA sequences in GenBank.

3.2. Molecular characterization of E. canis from Thai dogs

E. canis and *A. platys* are distributed globally, and it is generally assumed that all strains primarily utilize dogs and rhipicephaline ticks as vertebrate and invertebrate hosts. However, closely related anaplasmal pathogens, including

Table 1
Comparison of *E. canis*-Bangkok 16S rDNA sequence to geographically dispersed *E. canis* strains

E. canis strain	GenBank accession number	Identity (%)[a]	Nucleotide differences at position[b]												
			133	289	452	594	685	783	810	817	888	915	948	1174	1200
Bangkok	EF139458	100	G	C	A	A	A	A	–	G	C	A	T	C	C
VDE	AF373613	100	●	●	●	●	●	●	–	●	●	●	●	●	●
VHE	AF373612	100	●	●	●	●	●	●	–	●	●	●	●	●	●
Germishuys	U54805	99.92	●	–	●	●	●	●	–	●	●	●	●	●	●
Jake	NC_007354	99.92	●	●	●	●	●	●	A	●	●	●	●	●	●
Kagoshima1	AF536827	99.92	●	●	●	●	●	●	–	●	●	C	●	●	●
Gzh982	AF162860	99.84	●	●	●	●	●	–	–	●	●	●	●	T	●
Oklahoma	M73221	99.84	A	●	●	●	●	●	–	–	●	●	●	●	●
Madrid	AY394465	99.84	●	●	–	●	C	●	–	●	●	●	●	●	●
Florida	M73226	99.76	A	●	●	●	●	●	–	–	●	●	●	●	T
Lima	DQ915970	99.76	●	●	●	G	●	●	–	●	T	●	C	●	●

[a] The values are percentage of nucleotide sequence identities for 1247 bp determined from pairwise alignment.
[b] Positions based on the sequence of *E. coli* J01695 numbering system. The symbols (●) and (—) indicate conserved nucleotide and deletion, respectively.

other monocytotropic species, utilize much different mammalian hosts and tick vectors. Thus, the *E. canis*-Bangkok 16S rDNA sequence was compared to 10 other *E. canis* strains reported from China, Japan, Peru, South Africa, Spain, Venezuela and the USA, to confirm the identity of this Thai strain. All sequences were adjusted to the same length of 1247 base pairs prior to alignment. *E. canis*-Bangkok 16S rDNA was 100% identical to *E. canis*-VDE and *E. canis*-VHE strains from Venezuelan canine and human hosts, respectively (Unver et al., 2001). Remaining the *E. canis* strain sequences showed very close identity ranging from 99.76 to 99.92%. The most polymorphisms were observed between *E. canis* Bangkok and Lima strains. Four different 16S rDNA sequence patterns were found among the 11 *E. canis* strains, with polymorphisms at 13 positions that included 8 substitutions, 1 insertion and 4 deletions (Table 1). Substitutions consisted of six transitions and two transversions. Compared to *E. canis*-Bangkok (1) Germishuys, Jake and Kagoshima1 strains showed single nucleotide differences that were a deletion, an insertion and a substitution, respectively; (2) Gzh982, Oklahoma and Madrid strains had two positions with polymorphisms (one deletion and one substitution at different positions); (3) Florida and Lima strains both had three positions with polymorphisms, Florida with one deletion and two substitutions while Lima had three substitutions.

3.3. Molecular characterization of *A. platys* from Thai dogs

The same corresponding 16S rDNA sequences of *A. platys*-Bangkok and eight other *A. platys* strains reported from China, France, Japan, Spain, Thailand, Venezuela and the USA were aligned. *A. platys*-Bangkok was 100% identical to those from France and Okinawa, but was different from *A. platys* previously isolated from Thailand. Other closely related sequences of *A. platys* strains showed 99.60–99.92% sequence identity. Five sequence patterns were found among the 16S rDNA sequence alignment of 9 *A. platys* strains, with polymorphisms at 13 positions, 7 of which were substitutions, 3 were insertions and 3 were deletions (Table 2). Substitutions consisted of four transitions and three transversions. Compared to *A. platys*-Bangkok (1) Okinawa1 and Spain strains had single nucleotide additions at different positions; (2) Thailand and Venezuela strains had two nucleotide substitutions at different positions; (3) Gzh981 had two insertions and one deletion; (4) the most polymorphisms were between Bangkok and USA strains, which were three substitutions and two deletions.

Table 2
Comparison of *A. platys*-Bangkok 16S rDNA sequence to other *A. platys* strains

A. platys strain	GenBank accession number	Identity (%)[a]	Nucleotide differences at position[b]											
			152	181	393	678	766	818	820	871	961	1025	1181	1192	1233
Bangkok	EF139459	100	T	A	—	G	—	C	G	—	G	G	A	T	T
Sommieres	AF303467	100	●	●	—	●	—	●	●	—	●	●	●	●	●
Okinawa	AY077619	100	●	●	—	●	—	●	●	—	●	●	●	●	●
Okinawa1	AF536828	99.92	●	●	C	●	—	●	●	—	●	●	●	●	●
Spain	AY530806	99.92	●	●	—	●	G	●	●	—	●	●	●	●	●
Thailand	AF286699	99.84	●	●	—	●	—	●	●	—	●	A	●	C	●
Venezuela	AF287153	99.84	C	●	—	●	—	●	●	—	●	●	G	●	●
Gzh981	AF156784	99.76	●	●	C	●	—	●	●	T	—	●	●	●	●
USA	M82801	99.60	●	—	—	T	—	—	C	—	●	●	●	●	G

[a] The values are percentage of nucleotide sequence identities for 1249 bp determined from pairwise alignment.
[b] Positions based on the sequence of *E. coli* J01695 numbering system. The symbols (●) and (—) indicate conserved nucleotide and deletion, respectively.

3.4. 16S rRNA secondary structures

16S rRNA is not subject to the same selective influences on function as mRNA (i.e., 16S rRNA function relies on structure rather than codon usage), thus the effects of nucleotide changes on predicted secondary structures could be more informative than primary sequence variation alone. The positions of 16S rDNA sequences defined in Tables 1 and 2 were correlated with the *E. coli* J01695 numbering system (Konings and Gutell, 1995). Comparison of *E. canis*-Bangkok and *A. platys*-Bangkok 16S rRNA predicted secondary structures to that of the *E. coli* J01695 indicated that both had conserved tetra loops that generally constrained the rRNA architecture (Woese et al., 1990). Nucleotides at positions 289, 452, 594, 888, 915, 948 and 1200 of *E. canis*-Bangkok were common among bacteria while positions 133, 685, 783, 810, 817 and 1174 were different from other eubacteria. At position 948, we observed T (U) in most samples except *E. canis*-Lima that carried C, the latter of which is similar to those of alphaproteobacteria (Woese, 1987). In *A. platys*-Bangkok, nucleotide differences at positions 181, 678, 871 and 1025 were within the common structure of eubacteria. At position 393, eubacteria generally carried A (Woese, 1987), but most of *A. platys* had a single deletion except for *A. platys*-Okinawa1 and *A. platys*-Gzh981 that carried a C. At position 1233, T was observed in most samples except for *A. platys* USA, which had a G that is similar to those of alphaproteobacteria (Woese, 1987).

3.5. Phylogenetic analysis of Ehrlichia and Anaplasma

Ehrlichia and *Anaplasma* 16S rDNA sequences were used to generate a phylogenetic tree using the neighbor-joining method by MEGA software (Version 3.1). In addition to *E. canis* and *A. platys* strains, closely related species included the tick-borne anaplasmal parasites *E. chaffeensis*, *E. ewingii*, *E. muris*, *E. ruminantium*, *A. bovis*, *A. centrale*, *A. marginale*, *A. ovis* and *A. phagocytophilum*. A biologically divergent member of the Anaplasmataceae, *N. sennetsu*, was used as the outgroup. The resultant phylogenetic tree revealed that *E. canis*-Bangkok and *A. platys*-Bangkok were grouped tightly within the other *E. canis* and *A. platys* strains, respectively (Fig. 1). This analysis revealed that (1) *Ehrlichia* and *Anaplasma* were divided into clearly defined clades; (2) *E. canis* strains from different geographic regions were always grouped in a clade independent from *E. chaffeensis*, *E. ewingii*, *E. muris* and *E. ruminantium*; (3) *E. ewingii* showed the closet relationship to *E. canis* while *E. ruminantium* was the most distant; (4) *A. platys* strains from different countries constantly grouped in a clade independent from *A. bovis*, *A. centrale*, *A. marginale*, *A. ovis* and *A. phagocytophilum*; (5) *A. phagocytophilum* had the closet relationship to *A. platys* while *A. centrale*, *A. marginale* and *A. ovis* were the most distant; (6) *A. marginale* clustered in a branch linked to *A. centrale* and *A. ovis*.

4. Discussion

Understanding the epidemiology of these diseases and infectious cycles of anaplasmal organisms is expected to facilitate understanding of similar diseases, including zoonoses that may share similar etiologic agents or vectors. These results support previous studies that indicated *E. canis* and *A. platys* are uniform worldwide, suggesting that biological observations with different strains and from different regions could be laterally applicable. However, it is important to note that these observations are based on the highly conserved 16S rRNA gene and that further work with more divergent sequences is needed to confirm whether these strains are indeed uniform. Moreover, it is important to remember that potential heterogeneity of vertebrate and invertebrate hosts from different locations is likely to influence the biology of these organisms. Thus, global analyses of evolutionary patterns among *E. canis* and *A. platys* will likely require parallel analyses of such patterns among their reservoirs and vectors. For example, *R. sanguineus*, which is considered an important vector of both pathogens described in this report, is considered highly host-specific in some regions but reportedly feeds on diverse medium to large mammalian hosts (including humans) in other areas. Susceptibility of these tick strains to various pathogens, including *E. canis* and *A. platys*, could explain reports suggesting *E. canis* infections of ruminants and humans (Stich et al., in press; Yu et al., 2007). Alternatively, similarities among geographically diverse *E. canis* strains could be reflective of relatively recent global spread of this pathogen with its vertebrate and invertebrate hosts (Yu et al., 2007).

This report describes molecular analysis of nearly complete 16S rDNA sequences of *E. canis* and *A. platys* from Bangkok, Thailand. The sequence alignments and phylogenetic tree suggested low diversity within *E. canis* and *A. platys* strains based on the close similarity amongst their 16S rDNA sequences from all geographic regions tested, and these conclusions are consistent with other reports (de la Fuente et al., 2006; Aguirre et al., 2006; Yu et al., 2007). Genetic polymorphisms from comparison of 16S rRNA sequences based on the *E. coli* J01695 numbering system (Konings and Gutell, 1995) also indicated that *E. canis*-Bangkok and *A. platys*-Bangkok were structurally conserved in 16S rRNA architecture. Therefore, we are confident that all polymorphisms observed in these experiments are consistent with these two species. However, although nucleotide differences at many positions indicated that *E. canis*-Bangkok and *A. platys*-Bangkok shared some structure with other bacteria, other nucleotides were different from most eubacteria.

Close similarity among 16S rDNA of *A. platys* from different locations worldwide supported the hypothesis that *A. platys* strains are not geographically segregated (Huang et al., 2005). Interestingly, the sequence of *A. platys*-Bangkok reported from this study had two nucleotide substitutions compared to a previously characterized *A. platys* strain from Thailand (Suksawat et al., 2001) at positions 1025 and 1192 (Table 2) that appeared through G/A and T/C transitions, respectively, suggesting that there might be at least two *A. platys* strains enzootic to Thailand.

The phylogenetic tree separated two major clusters of *Ehrlichia* spp. and *Anaplasma* spp. As expected, *E. canis*-Bangkok was within the *Ehrlichia* clade and *A. platys*-Bangkok

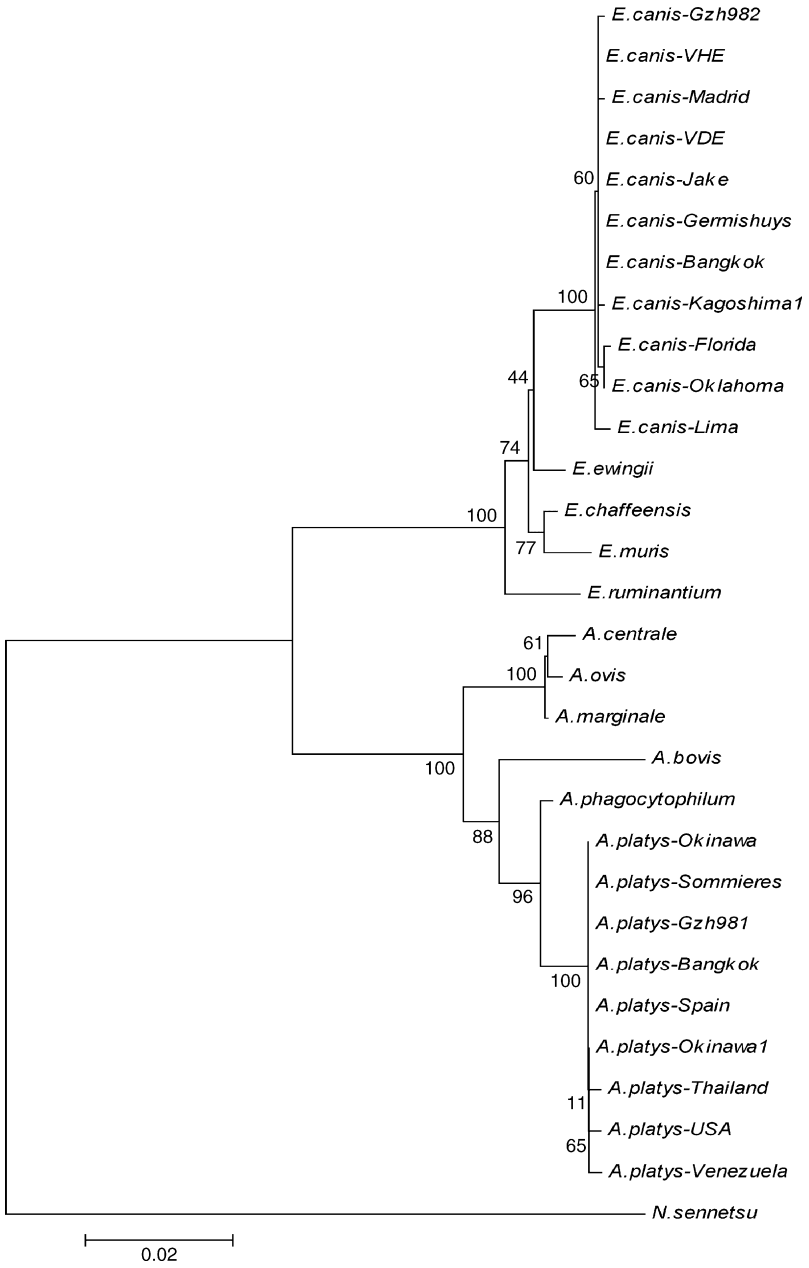


Fig. 1. Phylogenetic tree based on *Ehrlichia* and *Anaplasma* 16S rDNA. Sequences from the *Ehrlichia* and *Anaplasma* genera were compared with the neighbor-joining method with distance matrix calculation by Kimura-two parameters, operated by MEGA software (Version 3.1), using *N. sennetsu* as the outgroup. Scale bar indicates the number of mutations per sequence position. The numbers at the nodes represent the percentage of 1000 bootstrap re-samplings.

was within its respective clade. Within these two clusters, *E. canis* and *A. platys* strains grouped mostly in multiple connected branches. Although they were from geographically diverse regions, little genetic diversity was observed, suggesting slow and homogeneous evolution (Keysary et al., 1996). These results were in concordance with previous reports of slight genetic variation among 16S rDNA from different *E. canis* strains (Unver et al., 2003; Parola et al., 2003; Aguirre et al., 2004) and *A. platys* strains (Unver et al., 2003; Huang et al., 2005; Martin et al., 2005; de la Fuente et al., 2006). Notably, the *E. canis*-Bangkok 16S rDNA was identical to VDE and VHE that were respectively isolated from a dog and a human in Venezuela (Unver et al., 2001), which suggests little differentiation among *E. canis* between these geographic locations and tempts speculation about potential similarities in the epidemiology of these strains, but further analyses of less conserved sequences are needed to test this idea. This phylogenetic analysis also suggested that the *E. canis* was more closely related to *E. ewingii* than to *E. chaffeensis*, *E. muris* and *E. ruminantium*, and that *A. platys* was more closely related to *A. phagocytophilum* than to *A. bovis*, *A. centrale*, *A. marginale* and *A. ovis*, which both corroborate an earlier report (Yu et al., 2001).

In conclusion, our primers detected both *Ehrlichia* and *Anaplasma* in canine blood, which resulted in new 16S rDNA sequences from *E. canis* and *A. platys* infections of Thai dogs.

Although the 16S rDNA sequences were highly conserved among geographically diverse strains of these organisms, additional analyses of genes more subject to selective pressure from host environments (e.g., outer membrane protein gene families) could help elucidate the diversity and evolution of strains from different geographic areas. Current efforts include examination of additional canine blood samples to determine the presence and genetic diversity of *Ehrlichia* and *Anaplasma* spp. found in Thailand.

Acknowledgements

We give special thanks to Veterinary Diagnostic Laboratory (VDL) center, Bangkok for dog blood samples. This research was supported by the Graduate School, Kasetsart University.

References

Aguirre, E., Sainz, A., Dunner, S., Amusategui, I., Lopez, L., Rodriguez-Franco, F., Luaces, I., Cortes, O., Tesouro, M.A., 2004. First isolation and molecular characterization of *Ehrlichia canis* in Spain. Vet. Parasitol. 125, 365–372.

Aguirre, E., Tesouro, M.A., Ruiz, L., Amusategui, I., Sainz, A., 2006. Genetic characterization of *Anaplasma* (*Ehrlichia*) *platys* in dogs in Spain. J. Vet. Med. B Infect. Dis. Vet. Public Health 53, 197–200.

de la Fuente, J., Torina, A., Naranjo, V., Nicosia, S., Alongi, A., Mantia, F.L., Kocan, K.M., 2006. Molecular characterization of *Anaplasma platys* strains from dogs in Sicily. Italy Vet. Res. 2, 24–29.

Dumler, J.S., Barbet, A.F., Bekker, C.P., Dasch, G.A., Palmer, G.H., Ray, S.C., Rikihisa, Y., Rurangirwa, F.R., 2001. Reorganization of genera in the families *Rickettsiaceae* and *Anaplasmataceae* in the order *Rickettsiales*: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of *Ehrlichia equi* and 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. Int. J. Syst. Evol. Microbiol. 51, 2145–2165.

Greene, C.E., Harvey, J.W., 1990. Canine ehrlichiosis. In: Greene, C.E. (Ed.), Clinical Microbiology and Infectious Diseases of the Dog and Cat. W.B. Saunders, Philadelphia, pp. 137–148.

Groves, M.G., Dennis, G.L., Amyx, H.L., Huxsoll, D.L., 1975. Transmission of *Ehrlichia canis* to dogs by tick (*Rhipicephalus sanguineus*). Am. J. Vet. Res. 36, 937–940.

Huang, H., Unver, A., Perez, M.J., Orellana, N.G., Rikihisa, Y., 2005. Prevalence and molecular analysis of *Anaplasma platys* in dogs in Lara, Venezuela. Braz. J. Microbiol. 36, 211–216.

Keysary, A., Waner, T., Rosner, M., Warner, C.K., Dawson, J.E., Zass, R., Biggie, K.L., Harrus, S., 1996. The first isolation, in vitro propagation, and genetic characterization of *Ehrlichia canis* in Israel. Vet. Parasitol. 62, 331–340.

Konings, D.A.M., Gutell, R.R., 1995. A comparison of thermodynamic foldings with comparatively derived structures of 16S and 16S-like rRNAs. RNA 1, 559–574.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.

Martin, A.R., Brown, G.K., Dunstan, R.H., Roberts, T.K., 2005. *Anaplasma platys*: an improved PCR for its detection in dogs. Exp. Parasitol. 109, 176–180.

Parola, P., Cornet, J.P., Sanogo, Y.O., Miller, R.S., Thien, H.V., Gonzalez, J.P., Raoult, D., Telford III, S.R., Wongsrichanali, C., 2003. Detection of *Ehrlichia* spp., *Anaplasma* spp., *Rickettsia* spp., and other eubacteria in ticks from the Thai-Myanmar border and Vietnam. J. Clin. Microbiol. 41, 1600–1608.

Rozas, J., Sanchez-Delbarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analysis by the coalescent and other methods. Bioinformatics 19, 2496–2497.

Sambrook, J., Russell, D.W., 2001. In: Molecular Cloning: A Laboratory Manual. third ed. Cold Spring Harbor Laboratory Press, New York.

Stich, R.W., Schaefer, J.J., Bremer, W.G., Needham, G.R., Jittapalapong, S., in press. Host surveys, ixodid tick biology and transmission scenarios as related to the tick-borne pathogen, *Ehrlichia canis*. Vet. Parasitol., in press.

Suksawat, J., Pitulie, C., Arraga-Alvarado, C., Madrigal, K., Hancock, S.I., Breitschwerdt, E.B., 2001. Coinfection with three *Ehrlichia* species in dogs from Thailand and Venezuela with emphasis on consideration of the 16S ribosomal DNA secondary structure. J. Clin. Microbiol. 39, 90–93.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Unver, A., Perez, M., Orellana, N., Huang, H., Rikihisa, Y., 2001. Molecular and antigenic comparison of *Ehrlichia canis* isolates from dogs, ticks, and a human in Venezuela. J. Clin. Microbiol. 39, 2788–2793.

Unver, A., Rikihisa, Y., Kawahara, M., Yamamoto, S., 2003. Analysis of 16S rRNA gene sequence of *Ehrlichia canis*, *Anaplasma platys*, and *Wolbachia* species from canine blood in Japan. Ann. N.Y. Acad. Sci. 990, 692–698.

Woese, C.R., 1987. Bacterial evolution. Microbiol. Rev. 51, 221–271.

Woese, C.R., Winker, S., Gutell, R.R., 1990. Architecture of ribosomal RNA: constraints on the sequence of ''tetra-loops''. Proc. Natl. Acad. Sci. U.S.A. 87, 8467–8471.

Yu, X.J., Zhang, X.F., McBride, J.W., Zhang, Y., Walker, D.H., 2001. Phylogenetic relationships of *Anaplasma marginale* and '*Ehrlichia platys*' to other *Ehrlichia* species determined by GroEL amino acid sequences. Int. J. Syst. Evol. Microbiol. 51, 1143–1146.

Yu, X.J., McBride, J.W., Walker, D.H., 2007. Restriction and expansion of *Ehrlichia* strain diversity. Vet. Parasitol. 143, 337–346.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 439–444
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

The identification, diversity and prevalence of trypanosomes in field caught tsetse in Tanzania using ITS-1 primers and fluorescent fragment length barcoding

E.R. Adams [a,*], P.B. Hamilton [b], I.I. Malele [c], W.C. Gibson [a]

[a] School of Biological Sciences, University of Bristol, Bristol BS8 1UG, UK
[b] School of Biological Sciences, University of Exeter, Exeter EX4 4PS, UK
[c] Tsetse and Trypanosomiasis Research Institute, P.O. Box 1026, Tanga, Tanzania

Received 15 April 2007; received in revised form 26 July 2007; accepted 30 July 2007
Available online 6 August 2007

Abstract

We report on the development of two generic, PCR-based methods, which replace the multiple species-specific PCR tests used previously to identify the trypanosome species carried by individual tsetse flies. The first method is based on interspecies size variation in the PCR product of the ITS-1 region of the ribosomal RNA (rRNA) locus. In the second approach, length variation of multiple fragments within the 18S and 28S rRNA genes is assayed by PCR amplification with fluorescent primers; products are subsequently sized accurately and rapidly by the use of an automated DNA sequencer. Both methods were used to identify samples collected during large-scale field studies of trypanosome-infected tsetse in Tanzania in the National Parks of Tarangire and Serengeti, and the coastal forest reserve of Msubugwe. The fluctuations of trypanosome prevalence over time and two different field seasons are discussed. As well as facilitating the identification of trypanosome species with increased speed, precision and sensitivity, these generic systems have enabled us to identify two new species of trypanosome.

Keywords: *Trypanosoma*; *Glossina*; SSU ribosomal RNA gene; LSU ribosomal RNA gene; PCR; Fluorescent primers; ITS-1 region; Generic primers; FTA® cards

1. Introduction

Tsetse-transmitted trypanosomes are the causative agents of Human African Trypanosomiasis (HAT) and Animal African Trypanosomiasis (AAT), diseases of major medical and socio-economic importance in sub-Saharan Africa. *Trypanosoma brucei* is a human pathogen and also a livestock pathogen, along with several other species, including *T. congolense*, *T. simiae*, *T. godfreyi* and *T. vivax*. These tsetse-transmitted parasites collectively cause an annual estimated loss to Africa of US$ 4.75 billion in potential agricultural production (FAO, http://www.fao.org/ag/againfo/programmes/en/paat/disease.html). In order to target limited resources effectively, it is essential to assess the risk posed by particular tsetse species and populations by identifying the trypanosome species they carry.

Trypanosomes infecting tsetse have been traditionally identified by dissection and microscopic examination (Lloyd and Johnson, 1924). However, this method is inaccurate as it fails to discriminate between species that use the same developmental sites for infection, e.g., species within subgenus *Nannomonas*; it also fails to identify immature or mixed infections. The application of DNA-based methods, first species-specific DNA probes (Gibson et al., 1988; Kukla et al., 1987) and then species-specific PCR tests (Majiwa et al., 1994; Masiga et al., 1992; Moser et al., 1989), greatly improved the accuracy of identification and our understanding and knowledge of trypanosome diversity. In particular, the high prevalence of mixed infections with two, three or even four different trypanosome species was recognised for the first time (Majiwa and Otieno, 1990; McNamara et al., 1989, 1995; Solano et al., 1995; Woolhouse et al., 1996). With the use of species-specific PCR tests, it is now possible to identify the 11 tsetse-transmitted trypanosome species and subgroups for which there are available primers. However, even with these tests, some studies have failed to identify a high proportion of infections (e.g., Lehane et al.,

* Corresponding author. Tel.: +44 117 928 7574; fax: +44 117 331 7985.
E-mail address: emily.adams@bristol.ac.uk (E.R. Adams).

1567-1348/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2007.07.013

2000; Malele et al., 2003), either because of low quality or quantity of the DNA such that amplification is not possible, or, more interestingly, because these infections represent unknown trypanosomes which may be of veterinary importance. Performing multiple PCR reactions is also time-consuming and expensive. Therefore, it is desirable to develop generic PCR tests that are intended to amplify the DNA from all African trypanosomes. Desquesnes et al. (2001) developed a method based upon amplifying the Internal Transcribed Spacer (ITS-1) region of the ribosomal RNA locus. Species are subsequently identified by the inter-species size variation found in this region. This region was found suitable as it is flanked by highly conserved regions with an internal variable region and has a high copy number, around 100–200. This method was modified to increase the sensitivity of the primers to *T. vivax*, an important cattle trypanosome (Njiru et al., 2004). Nested PCR was used by Cox et al. (2005) to amplify all cattle trypanosomes and primers were refined again by Adams et al. (2006) to increase the specificity of the test for the diverse range of trypanosomes that are found in tsetse flies. This method has proved a useful tool for identifying trypanosomes quickly, accurately and cheaply. However, Adams et al. (2006) were still only able to identify 56% of infections in dissection-positive tsetse caught in Tanzania. Furthermore, an additional species-specific PCR was frequently needed when the PCR products of some species were too similar in size to distinguish, for example *T. simiae* and *T. simiae* Tsavo. This increased the expense and time taken to identify infections.

To address these problems we developed a novel method for identification of trypanosome species based on variable length regions within the 18S and 28S rRNA genes (Hamilton et al., 2007). Generic primers were designed in conserved regions to amplify across internal variable regions. One primer in a pair is labelled with a fluorescent tag; therefore, the resultant PCR can be run through an automated sequencer to measure the size of the PCR product with great accuracy. Four sets of primers within the 18S and 28S rRNA genes allow species identification based on the sizes of all four products run in a single well. This technique has the ability to distinguish all trypanosome species tested, and in addition has recognised putative new species. All species can be identified without the need for any further reaction as it is highly unlikely that two different species will have identical length regions for all four primer sets. The system shows a high level of accuracy and sensitivity, allowing a high throughput of samples with typically 96 samples genotyped within 24 h.

Here we compare the results obtained with the two methods – ITS-1 generic primers and fluorescent fragment length barcoding (FFLB) – on two sample sets of trypanosome infections from the midguts of tsetse collected in Tanzania. We also compare two methods of collection and storage of DNA from infected tsetse midgut samples.

2. Materials and methods

2.1. Field collection

Field work was organised with the Tsetse and Trypanosomiasis Research Institute (TTRI), Tanga, Tanzania. Samples were collected in two field seasons, between 21.01.2005–22.02.2005, and 12.06.2006–19.07.2006. Three field sites were used: Msubugwe, on the northern periphery of Saadani National Park; Serengeti National Park, and Tarangire National Park both in the north of Tanzania (Fig. 1). Msubugwe is a forest reserve on the east coast of Tanzania; the vegetation is mainly thick bush with some forest. The area supports a moderate population of wild animals and is close to Mivumoni and Mkwaja, two old cattle ranches closed down due to severe trypanosomiasis. In Serengeti and Tarangire National Parks the vegetation is more open woodland and savannah, with thicker bush in Tarangire. Both areas hold a wide variety of wild animals.

Fig. 1. Sketch map of Tanzania showing field collection sites. National parks are shaded pale grey.

Biconical traps were used to catch tsetse in Msubugwe; all were baited with chemical attractants acetone and 3-*n*-proplyphenol 4-methyl-phenol and octenol (Hargrove and Langley, 1990). In Serengeti and Tarangire mobile scoop traps were used, as it was found that the species of tsetse in these areas, typically *Glossina swynnertoni*, did not frequent the biconical traps. These traps were also baited with the same attractants. Flies were collected every morning for dissection later the same day.

2.2. Fly dissection and sample collection

For both 2005 and 2006 field collections, all non-teneral flies were dissected on the same day of collection and examined for the presence of trypanosomes using a light microscope. For the

2005 field collection, positive midgut samples were spotted onto Whatman® FTA cards, air-dried and stored at room temperature (Adams et al., 2006). For the 2006 field collection, positive samples were stored in 100 μl of 100% ethanol in individual 1.5 ml microcentrifuge tubes. Proboscis samples were also taken, examined for trypanosomes and held in 50 μl of 100% ethanol.

2.3. DNA preparation and PCR analysis

Whatman® FTA cards were prepared for PCR as per manufacturer's instructions. Samples in 100% ethanol were prepared by ammonium acetate precipitation (Bruford et al., 1998). Briefly, samples were digested in Digsol buffer (50 mM Tris, 20 mM EDTA, 117 mM NaCl and 1% SDS) with Proteinase K (final concentration 10 mg/ml) at 55 °C for 3 h, and then precipitated using ammonium acetate at a final concentration of 2.5 M, followed by ethanol precipitation and washing with 70% ethanol. Pellets were air-dried for 30 min, DNA was re-suspended in 50 μl sterile water, samples were stored at −20 °C until use.

Standard and nested PCR were performed as in (Adams et al., 2006), except that for the analysis of the 2006 field samples, the annealing temperature of the primers TRYP 3 and 4 was raised to 58 °C and 30 cycles were carried out. In general, nested PCR was not required for the 2006 samples, as the DNA was stored in sufficient quantity and of high enough quality that standard PCR was able to identify the majority of samples. Fluorescent fragment length barcoding (FFLB) was carried out as described in (Hamilton et al., 2007).

3. Results and discussion

3.1. Sample preparation and storage

During the 2005 field collection, infected midguts were spotted onto Whatman® FTA cards. This was a simple and convenient method to use in the field, as no further processing was needed and they could be held at room temperature for transportation back to the UK. However, the procedure for preparation of the discs for PCR and for washing them for re-use was tedious and labour intensive. Samples could only be used in PCR for a total of three times, limiting the use of the sample. Most importantly, we were only able to identify 56.9% of the infected midgut samples, even after nested PCR (Adams et al., 2006) or 58.3% with FFLB. The high proportion of unidentified samples (43.1% nested PCR with ITS-1 primers and 41.7% FFLB) may have been due to sub-optimum binding of trypanosome DNA to the FTA cards. There were many occasions when the ITS-1 primers and FFLB did not identify the same species from an individual fly midgut sample. This may have been because trypanosome DNA was unable to bind evenly throughout the card due to competition with midgut material; as a result, one species of trypanosome may be found in one section of the card, whilst another may have bound to a different area. Thus, multiple discs from the same FTA card may not contain the same DNA for an individual midgut sample. In addition, since midguts contain high levels of PCR inhibitors such as proteases and haem moeties, it is possible that the DNA extraction method was not capable of cleaning the DNA sufficiently for PCR.

The identification rate improved considerably for the 2006 collection, when samples were stored in 100% ethanol and DNA purified using the ammonium acetate precipitation (Bruford et al., 1998). This gave 78% identification with generic ITS-1 primers, and 97% using FFLB (Hamilton et al., 2007). It is possible that some mixed infections may be missed due to over-representation of one species of trypanosome in the midgut, however these identification rates are impressive.

Although collection and storage of field samples by this method is less convenient than for FTA cards, the superior identification rate now makes this our method of choice. Furthermore, although still labour intensive, purification of DNA samples needs to be carried out once only, not every time samples are re-used as for FTA cards, and 50 μl of homogeneous sample is available for use in many PCR reactions.

3.2. Comparison of tsetse infection rates from 2005 to 2006

In 2005 during the dry season, 1692 non-teneral tsetse flies were dissected from Msubugwe, of which 102 were positive for trypanosome infection by microscopy; a further 700 flies from Serengeti National Park were dissected, of which 42 were positive for trypanosomes. This gives an infection rate of exactly 6% for both areas. The majority of the flies caught in Msubugwe were *Glossina pallidipes*, with few *G. brevipalpis*; in Serengeti only *G. swynnertoni* were found, a smaller species of tsetse the range of which is restricted to Northern Tanzania.

In 2006 at the end of the wet season, 1331 flies were dissected from Msubugwe, of which 77 (5.8%) had an infected midgut; proboscides were also examined for infection from 519 flies, of which 19 were infected (3.6%). In Tarangire National Park 1523 flies were dissected, with 70 (4.6%) infected midguts, and 49 infections from 720 dissected proboscides (6.8%). Again, *G. pallidipes* and a few *G. brevipalpis* were found in Msubugwe. However, in 2006 no *G. brevipalpis* flies were found to be positive for trypanosomes, perhaps because of the low numbers of this species of fly dissected. In Tarangire only *G. swynnertoni* were found and dissected. These figures suggest that infection rate of tsetse is relatively consistent throughout the dry and wet seasons.

3.3. Diversity and prevalence of midgut infections from 2005 to 2006 Msubugwe

The identification of midgut infections by the ITS-1 and FFLB methods allows comparison of the relative distribution and abundance of trypanosome species.

Fig. 2 shows a comparison of trypanosome species identified by the ITS-1 and FFLB methods from Msubugwe for the 2006 field collection. The chart shows a good level of agreement between the species identified by both methods. As expected, the FFLB system is more sensitive and detects more infections, but there are four occasions when the ITS-1 method has identified an

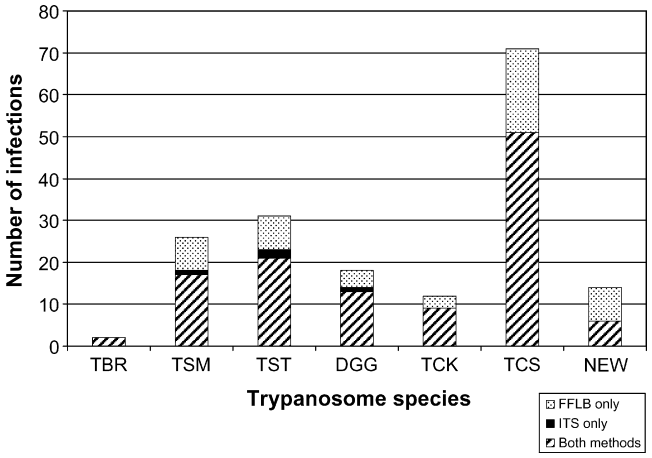


Fig. 2. Comparison of the species of trypanosome identified by ITS-1 and FFLB in Msubugwe in 2006. The graph shows the degree of congruency between the methods. TBR, *T. brucei*; TSM, *T. simiae*; TST, *T. simiae* Tsavo; DGG, *T. godfreyi*; TCK, *T. congolense* Kilifi; TCS, *T. congolense* savannah; NEW, novel genotype found in 2006 field trip. Each species in a mixed infection is counted separately.

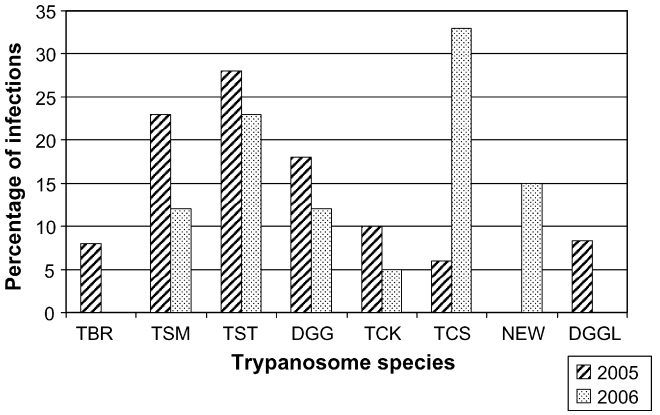


Fig. 3. Comparison of the percentage of infections found in Msubugwe from 2005 and 2006 field collections. TBR, *T. brucei*; TSM, *T. simiae*; TST, *T. simiae* Tsavo; DGG, *T. godfreyi*; TCK, *T. congolense* Kilifi; TCS, *T. congolense* savannah; NEW, novel genotype found in 2006 field trip; DGGL, *T. godfreyi*-like. Each species in a mixed infection is counted separately.

infection that the FFLB system has missed. The congruency between the results shows that both methods are accurate in their identification of infection, the results are reproducible and the methods can work in combination. Each method has advantages and disadvantages: FFLB is accurate, quick and has an extremely high level of identification; however, it requires expensive equipment and trained technical staff. The ITS-1 method, whilst not quite as sensitive as the FFLB, is again quick and accurate, with the advantage that it is significantly cheaper to run and less expertise and equipment are required.

Fig. 3 compares the total percentage of species identified in the 2005 and 2006 Msubugwe field collections using both FFLB and ITS-1. The percentage of infections remains similar for *T. godfreyi*, *T. simiae* Tsavo and *T. congolense* Kilifi, with a drop of infections of *T. simiae* in 2006. Notably, no *T. brucei* infections were identified from the area in 2006, whereas in the 2005 field collection there were a high number of mixed infections of *T. brucei* with *T. simiae*, indicating that the *T. brucei* infections may have originated from suids. Pigs are a common reservoir of *T. brucei* in East Africa, e.g., (Waiswa et al., 2006). It is also possible that this is a result of cattle being removed from the area in the last decade, thus removing a reservoir of *T. brucei* infection. In contrast, the percentage of infections containing *T. congolense* savannah increased dramatically from 6% to 33% between 2005 and 2006 (Fig. 3). This may be due to seasonal variation in the mammalian host population. The 2005 field collection was during the dry season when many animals migrate between the areas of Mivumoni and Saadani National Park in search of food and water (Malele et al., 2003); in 2006 the collection was taken at the end of the long rains in June. It is therefore possible that hosts of *T. congolense* moved back into the area as conditions become more favourable. However, as many hosts of *T. congolense* savannah are also hosts of *T. brucei*, it is difficult to explain the increase in *T. congolense* savannah infections at the same time as the absence of *T. brucei*.

In 2005 a new trypanosome named *T. godfreyi*-like was found in Msubugwe (Adams et al., 2006) that had previously been discovered in nearby field sites at Mivumoni and Pangani, and found to be most closely related to *T. godfreyi* by phylogenetic analysis (Malele et al., 2003). In 2005 this trypanosome showed relatively high prevalence in this area (7%, 6 of 87 infections), but surprisingly was not found at all in the 2006 field collection. Again, this may be due to seasonal migration of presumed mammalian hosts.

The FFLB method revealed the existence of a putative new species of trypanosome from Msubugwe in the 2006 field collection that is most closely related to *T. brucei* by phylogenetic analysis (Hamilton et al., 2007). The amplified ITS-1 fragment of this trypanosome was the same size as that of *T. simiae*, making it difficult to identify using the ITS-1 primers, and, of course, no species-specific PCR test for it yet exists. By FFLB, it was evident that several fly samples produced a unique and consistent pattern of amplified rRNA fragments, and sequence analysis of the genes for 18S rRNA and glycosomal glyceraldehyde phosphate dehydrogenase, confirmed that this was a new trypanosome (Hamilton et al., 2007). To find a species so closely related to *T. brucei* is highly unexpected, and it could provide the evolutionary link between the *T. brucei* clade and the other subgenera of tsetse-transmitted trypanosomes, *Nannomonas* and *Duttonella*. This is an extremely interesting result and we hope to be able to isolate the living specimen in the near future for further biological characterisation. The new species was found in six single infections and in seven mixed infections (Table 1), and has a relatively high prevalence at 15% of all infections in Msubugwe from 2006 (Fig. 3).

In Msubugwe in 2005 the suid-associated trypanosomes (*T. simiae*, *T. simiae* Tsavo, *T. godfreyi* and possibly also *T. godfreyi*-like) were found to be more prevalent than *T. congolense* savannah, *T. congolense* Kilifi and *T. brucei*, with 75.8% total prevalence. However in 2006, with the increase in prevalence of *T. congolense* savannah, the difference is less marked, with 47.0% of trypanosomes from suid-associated species. However, the new trypanosome was mainly found in

Table 1
Prevalence of new trypanosome species in single and mixed infections in tsetse midguts

Single infections	6
Mixed infections	
New + TCS	1
New + TST	2
New + DGG	3
New + TSM	1

mixed infections with suid-associated trypanosomes (Table 1), placing it in the same category. This would take the total prevalence of suid-associated trypanosomes to 61.7% in 2006. Thus it appears that the tsetse in Msubugwe feed mainly on suids, but this may fluctuate between the wet and dry seasons as animals migrate to find food and water.

3.4. Diversity and prevalence of midgut infections from 2005 to 2006 Tarangire and Serengeti

Tarangire and Serengeti National parks are approximately 50 miles apart, with animals migrating between the two parks (Fig. 1). Both 2005 and 2006 collections were made at the end of the rainy season, with *G. swynnertoni* the only tsetse species sampled. In both Tarangire and Serengeti, the main trypanosome species found were *T. brucei*, *T. simiae*, *T. simiae* Tsavo, *T. godfreyi*, *T. congolense* savannah and Kilifi. The *T. godfreyi*-like trypanosome was not found in 2005 and 2006, but has been previously reported in these areas (Malele et al., 2003). We have also not yet found the new species of trypanosome discovered in Msubugwe in 2006 in these areas, perhaps because it is geographically restricted to the coastal areas or its natural animal hosts are not present in these National Parks. It is also possible that the species of fly *G. swynnertoni*, is not a suitable insect host.

In Tarangire *T. congolense* savannah predominates, at 50% of all infections (Fig. 4), with only a few *T. brucei* and *T. godfreyi* infections present. In Serengeti *T. congolense* savannah is still the most prevalent trypanosome (24%), but other trypanosome species are also frequently found. Only low numbers of *T. congolense* Kilifi infections were found in either region (Fig. 4).

Considering the relative proportion of infections with suid-associated trypanosomes found in the two areas, it appears that the *G. swynnertoni* population is more dependent on suids in Serengeti than Tarangire. In Tarangire, bovid-associated trypanosomes (*T. congolense* savannah and Kilifi, *T. brucei*) are more prevalent, 59% (47 of 80) infections, whilst in Serengeti the prevalence is 48% (34 of 70) infections. Tsetse in these National parks may be expected to have a wide variety of animals to feed upon, and we may hypothesise that, at least in Tarangire, they feed on bovids predominantly. In order to confirm these hypotheses, it is necessary to complete blood-meal analysis on samples of tsetse from these areas.

T. vivax was not found in this study as this species colonises the fly mouthparts only and so far only midgut samples have been analysed. We expect that *T. vivax* will be found among the proboscis samples collected from the same areas.

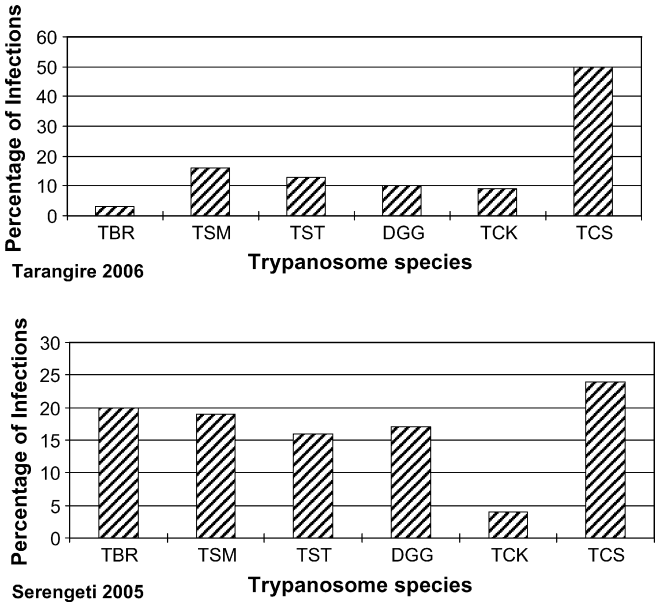


Fig. 4. Comparison of the percentage of infections found in Serengeti (2005) and Tarangire (2006) field collections. TBR, *T. brucei*; TSM, *T. simiae*; TST, *T. simiae* Tsavo; DGG, *T. godfreyi*; TCK, *T. congolense* Kilifi; TCS, *T. congolense* savannah. Each species in a mixed infection is counted separately.

4. Conclusions

We have developed two generic systems to identify trypanosome infections from wild-caught tsetse based upon PCR analysis. We have demonstrated the benefits of both systems and any disadvantages, including cost, that they have. Encouragingly both methods show high levels of congruency when looking at individual infections, thus showing the systems to be reliable and the results reproducible. The systems have also allowed the identification of two new species of trypanosome by their unique sizes in variable regions of DNA. We now have the opportunity to study the DNA, and hopefully living specimens, of these trypanosomes in detail for full biological characterisation.

The identification systems now enable us to study the changing relationships of trypanosomes in different areas, and how land use, species of fly, host prevalence and seasonal migration patterns can help us predict disease risk in these areas. We can now see that the prevalence of trypanosomes in Msubugwe is a dynamic situation, possibly as a consequence of migration of host animals. We have found a putative new species of trypanosome most closely related to *T. brucei* in this area, which may be of veterinary importance. This putative species of trypanosome could be the evolutionary link between *T. brucei* and the other species of tsetse-transmitted trypanosomes, and its characterisation will greatly benefit the understanding of these parasites.

Acknowledgements

We would like to thank NERC for the studentship for Emily Adams and The Genetics Society for funding part of the field work in 2006. We would also like to thank everyone at the

TTRI, especially Dr. Msangi and Fussah, Matembo and Mashenga for help with fieldwork and dissection, TANAPA HQ in Arusha and Tarangire National Park Management for the permit to trap tsetse in the park. Thanks to Tim Colborn for assistance with artwork.

References

Adams, E.R., Malele, I.I., Msangi, A.R., Gibson, W.C., 2006. Trypanosome identification in wild tsetse populations in Tanzania using generic primers to amplify the ribosomal RNA ITS-1 region. Acta Trop. 100, 103–109.

Bruford, M.W., Hanotte, O., Brookfield, J.F.Y., Burke, T., 1998. Multilocus and single-locus DNA fingerprinting. In: Hoelzel, A.R. (Ed.), Molecular Genetic Analysis of Populations: A Practical Approach. IRL Press, Oxford, pp. 287–336.

Cox, A., Tilley, A., McOdimba, F., Fyfe, J., Eisler, M., Hide, G., Welburn, S., 2005. A PCR based assay for detection and differentiation of African trypanosome species in blood. Exp. Parasitol. 111, 24–29.

Desquesnes, M., McLaughlin, G., Zoungrana, A., Davila, A.M.R., 2001. Detection and identification of *Trypanosoma* of African livestock through a single PCR based on internal transcribed spacer 1 of rDNA. Int. J. Parasitol. 31, 610–614.

Gibson, W.C., Dukes, P., Gashumba, J.K., 1988. Species-specific DNA probes for the identification of trypanosomes in tsetse. Parasitology 97, 63–73.

Hamilton, P.B., Adams, E.R., Malele, I.I., Gibson, W.C., 2007. A novel high throughput technique for species identification, reveals a new species of tsetse-transmitted trypanosome related to *Trypanosoma brucei*. Infect. Genet. Evol., submitted.

Hargrove, J.W., Langley, P.A., 1990. Sterilizing tsetse in the field: a successful field trial. Bull. Entomol. Res. 80, 397–403.

Kukla, B.A., Majiwa, P.A.O., Young, C.J., Moloo, S.K., Ole-Moiyoi, O.K., 1987. Use of species-specific DNA probes for the detection and identification of trypanosome infections in tsetse flies. Parasitology 95, 1–26.

Lehane, M.J., Msangi, A.R., Whitaker, C.J., Lehane, S.M., 2000. Grouping of trypanosome species in mixed infections in *Glossina pallidipes*. Parasitology 120, 583–592.

Lloyd, L., Johnson, W.B., 1924. The trypanosome infections of tsetse flies in Northern Nigeria and a new method of estimation. Bull. Entomol. Res. 14, 265–288.

Majiwa, P.A.O., Otieno, L.H., 1990. Recombinant DNA probes reveal simultaneous infection of tsetse flies with different trypanosome species. Mol. Biochem. Parasitol. 40, 245–254.

Majiwa, P.A.O., Thatthi, R., Moloo, S.K., Nyeko, J.H.P., Otieno, L.H., Maloo, S., 1994. Detection of trypanosome infections in the saliva of tsetse flies and buffy coat samples from antigenaemic but aparasitaemic cattle. Parasitology 108, 313–322.

Malele, I., Craske, L., Knight, C., Ferris, V., Njiru, Z., Hamilton, P., Lehane, S., Lehane, M., Gibson, W., 2003. Identification of new trypanosome species from wild tsetse flies in Tanzania. Infect. Genet. Evol. 3, 271–279.

Masiga, D.K., Smyth, A.J., Hayes, P.J., Bromidge, T.J., Gibson, W.C., 1992. Sensitive detection of trypanosomes in tsetse flies by DNA amplification. Int. J. Parasitol. 22, 909–918.

McNamara, J., Dukes, P., Snow, W.F., Gibson, W.C., 1989. Use of DNA probes to identify *Trypanosoma congolense* and *T. simiae* in tsetse from The Gambia. Acta Trop. 46, 55–61.

McNamara, J.J., Laveissiere, C., Masiga, D.K., 1995. Multiple trypanosome infections in wild tsetse in Cote d'Ivoire detected by PCR analysis. Acta Trop. 59, 85–92.

Moser, D.R., Cook, G.A., Ochs, D.E., Bailey, C.P., McKane, M.R., Donelson, J.E., 1989. Detection of *Trypanosoma congolense* and *T. brucei* subspecies by DNA amplification using the polymerase chain reaction. Parasitology 99, 57–66.

Njiru, Z.K., Makumi, J.N., Okoth, S., Ndungu, J.M., Gibson, W.C., 2004. Identification of trypanosomes in *Glossina pallidipes* and *G. longipennis* in Kenya. Infect. Genet. Evol. 4, 29–35.

Solano, P., Argiro, L., Reifenberg, J.M., Yao, Y., Duvallet, G., 1995. Field application of the polymerase chain reaction (PCR) to the detection and characterization of trypanosomes in *Glossina longipalpis* (Diptera: Glossinidae) in Cote d'Ivoire. Mol. Ecol. 4, 781–785.

Waiswa, C., Olaho-Mukani, W., Katunguka-Rwakishaya, E., 2006. Domestic animals as reservoirs for sleeping sickness in three endemic foci in southeastern Uganda. Ann. Trop. Med. Parasitol. 97, 149–155.

Woolhouse, M.E.J., McNamara, J.J., Hargrove, J.W., Bealby, K.A., 1996. Distribution and abundance of trypanosome (subgenus *Nannomonas*) infections of the tsetse fly *Glossina pallidipes* in southern Africa. Mol. Ecol. 5, 11–18.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 445–449
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Molecular detection of divergent trypanosomes among rodents of Thailand

S. Jittapalapong [a], T. Inpankaew [a], N. Sarataphan [b], V. Herbreteau [c], J.P. Hugot [d], S. Morand [e,*], R.W. Stich [f]

[a] *Department of Parasitology, Kasetsart University, Bangkok, Thailand*
[b] *National Institute of Animal Health, Pathumthani, Thailand*
[c] *Laboratoire Espace, Santé, Territoire, Université Paris X, Nanterre, France*
[d] *OSEB, UMR 5202 CNRS, MNHN, Paris, France*
[e] *Institut des Sciences de l'Evolution, CNRS-UM2, Université Montpellier 2, 34095 Montpellier, France*
[f] *Department of Veterinary Pathobiology, University of Missouri, Columbia, MO, USA*

Received 19 April 2007; received in revised form 17 August 2007; accepted 21 August 2007
Available online 25 August 2007

Abstract

Herpetosoma is a homogenous subgenus of several dozen named species that are often described as morphologically indistinguishable *T. lewisi*-like parasites. These trypanosomes normally infect rodents and utilize fleas as vectors. Although this trypanosome subgenus is considered non-pathogenic to normal hosts, some of them are on rare occasion reported in association with human disease. Recently, a *T. lewisi*-like infection was detected in a sick Thai infant, thus the objective of this study was to investigate the prevalence of *T. lewisi* infections among different rodents indigenous to Thailand in order to identify possible sources of human cases. Blood was collected from a total of 276 rodents trapped from urban and rural areas of three Thai provinces between 2006 and 2007. These samples were processed for DNA isolation and tested with a PCR assay universal for the genus *Trypanosoma*, followed by internal transcribed spacer 1 (ITS-1) sequence analysis to identify infections in positive samples. *Herpetosoma* known as *T. lewisi*-like trypanosomes were present among *Rattus* (14.3%) and *Bandicota* (18.0%) rodent species and salivarian trypanosomes closely related to *T. evansi* were detected in *Leopoldamys* (20%) and *Rattus* (2.0%) species. *Herpetosoma* were prevalent among rodents associated with both human and sylvatic habitats, while three of the four salivaria-positive rodents were from a forest biotope. A *Herpetosoma* ITS-1 sequence amplified from one of these samples was 97.9% identical to that reported for *T. lewisi* in an experimentally infected rat and 96.4% identical to the sequence amplified from blood from a Thai infant. Habitats where rodents were collected significantly affect rodent infection, at least for *T. lewisi*, suggesting that the degree of anthropization may influence the transmission of *Trypanosoma* spp. These results suggest that multiple *Herpetosoma* species or strains are enzootic to Thailand, and that *Rattus* and *Bandicota* species are possible sources of human exposure to these parasites.

Keywords: *Trypanosoma*; *Herpetosoma*; Salivaria; Rodents; Habitat; Internal transcribed spacer 1 (ITS-1)

1. Introduction

The etiologic agents of many serious infectious diseases utilize invertebrate hosts during a portion of their life cycle. Most of these agents are adapted to hematophagous arthropods that share their vertebrate hosts, and identification of arthropod vectors and vertebrate reservoirs is usually key to sustained control of vector-borne diseases. Trypanosomes are flagellated protozoan parasites, some of which can cause distinct zoonoses, the most notable of which include the Leishmaniases, American trypanosomiasis (Chagas' disease) and African trypanosomiasis (sleeping sickness). Like most vector-borne zoonotic agents, trypanosomes that infect humans utilize a broad vertebrate host range and a relatively narrow range of invertebrate vectors. However, there are rare yet notable exceptions to the latter generalization.

The genus *Trypanosoma* can be divided into two major groups that infect vertebrates, the salivaria and the stercoraria

* Corresponding author.
E-mail address: morand@isem.univ-montp2.fr (S. Morand).

1567-1348/$ – see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.meegid.2007.08.002

(Hoare, 1972). Members of the salivaria or 'anterior station' group are frequently pathogenic to vertebrate hosts. These organisms usually undergo cyclical development in the anterior part of the insect, midgut, prior to biological transmission via vector salivary glands. Some members of this group are mechanically transmitted by inoculation during vector feeding and some others are completely adapted to vertebrates without the need of an invertebrate vector. Conversely, most members of the stercoraria, or 'posterior station' group, are non-pathogenic to natural vertebrate hosts. These parasites undergo cyclical development in the arthropod hindgut before transmission to the vertebrate hosts through vector feces. Both *Trypanosoma* groups are enzootic to Thailand.

Salivarian trypanosome enzootic to Thailand is *T. evansi*, a salivaria, which is considered the primary agent of trypanosomiasis among domestic animals in Asia and India. *T. evansi* is mechanically transmitted among bovids, camelids, cervids, equids and canids by biting flies in the suborder Brachycera (Shrivastava and Shrivastava, 1974; Joshi et al., 2005). However, recently *T. evansi* was isolated from a case of human trypanosomiasis in India (Joshi et al., 2005). Stercorarian trypanosomes found in Thailand include *T. lewisi*-like species of the subgenus *Herpetosoma*, which are generally vertebrate-specific, non-pathogenic flea-borne parasites of rodents. Although *Herpetosoma* species are considered specific to a single vertebrate host genus, they infect a relatively broad range of flea vectors (Molyneux, 1969; Linardi and Botelho, 2002; Desquesnes et al., 2002). Fleas are often opportunistic ectoparasites of available mammalian hosts. Thus, a rodent reservoir and flea vector appeared to be the most likely source of a *T. lewisi*-like (*Herpetosoma*) infection recently reported in a sick infant from Thailand (Sarataphan et al., 2007).

In Thailand, different rodent species have partially overlapping distributions in various habitats, thus the natural vertebrate reservoir(s) of the recent human infection cannot be identified using current information. Although rodents found in human habitats may be suspected, given the rarity of human trypanosomiasis in Thailand, it is also plausible that the aforementioned human infection was a *Trypanosoma* sp. naturally infecting rodents spp. not normally associated with humans. Thus, the objective of this study was to identify the rodent reservoirs of trypanosomes among different biotopes in Thailand. This was accomplished with the same PCR assay used to characterize the infant infection (Sarataphan et al., 2007). Within 12 examined rodent species, 38 included positive individuals for stercorarian trypanosomes, and 4 were positive for salivarian trypanosomes.

2. Materials and methods

2.1. Samples

Blood samples were collected from rodents in Kalasin, Loei and Phrae provinces. Rodents were trapped either from city or rural areas, and habitats were classified as rice fields, upland fields, secondary forests, and domestic habitat (including houses and their gardens, or markets). Rodents were examined in the field and identified to the species. Blood samples were collected via heart puncture, preserved in citrate salt tubes, and sent to the Department of Parasitology, Faculty of Veterinary Medicine, Kasetsart University, where they were stored at −20 °C.

2.2. Molecular diagnostics

The blood samples were thawed and DNA was isolated with a commercial kit (FlexiGene, QIAGEN, USA) according to the manufacturer's protocol. Control DNA samples from *T. evansi* (Tansui, Taiwan isolate, D89527) were kindly provided by Prof. Ikuo Igarashi, Director of National Research Center for Protozoan Disease, Obihiro University of Agriculture and Veterinary Medicine, Japan. PCR was carried out in 50 μl reaction mixture containing final concentrations of 10 mM Tris–HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM of each of the four deoxyribonucleotide triphosphates, 1 μM of each primer, 0.25 units of *Taq* polymerase (QIAGEN, USA), 5% (v/v) DMSO, and 5 μL of DNA template. Primers TRYP1S (CGTCCCTGCCATTTGTACACAC) and TRYP1R (GGAAGCCAAGTCATCCATCG) were used for amplification of the ITS-1 fragment (Desquesnes et al., 2002). PCR thermal cycling parameters consisted of an initial denaturation step at 94 °C for 2 min, followed by 35 cycles of 30 s at 94 °C, 30 s at 55 °C and 30 s at 72 °C, with a final extension step at 72 °C for 10 min. PCR products were electrophoresed on a 1.2% agarose gel, stained with ethidium bromide and observed under ultraviolet light.

Amplicons were purified with a NucleoSpin Extract II purification kit (Clonetech, Mountain View, CA, USA), and cloned into pGEM-T (Promega, Madison, WI, USA) according to the manufacturers' protocols. Cloned amplicon sequences were determined with *Taq* DyeDeoxy Cycle Sequencing System reagents (Applied Biosystems, Foster City, CA, USA) and a 310 Genetic Analyzer automated DNA sequencer (Applied Biosystems). Homologous sequences in the GenBank database were searched using basic local alignment search tool (BLAST) provided by NCBI (USA). Clustal W (V. 1.82) provided by EBI (UK) was used for multiple-sequence alignments with ITS-1 sequences of *T. lewisi* from an infected rat and *T. evansi* from Taiwan.

2.3. Statistical analysis

A chi-square test was used to examine differences in prevalence of trypanosome infections among the different species with a theoretical prevalence over 5%. Habitats where rodent were trapped were classified according to a gradient of anthropization: from (1) secondary forests, (2) upland fields and (3) rice fields to (4) domestic habitat (i.e. houses, markets). We performed a logistic regression on rodent infection, either by *T. lewisi* or by *T. evansi*, in relation to habitat.

3. Results

Different rodent species indigenous to Thailand are potential hosts for different *Herpetosoma* species. In the present study,

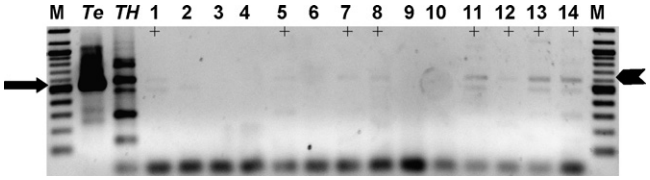


Fig. 1. PCR assays for trypanosomes rodent blood. Amplicons that included ITS-1 of trypanosome-specific rDNA were generated with primers TRYP1S and TRYP1R and visualized on an agarose gel stained with ethidium bromide. The molecular size standards (M) were 100 bp ladders, and the positive controls were *T. evansi* (*Te*) that contained the 520 bp amplicon (arrow) and a *T. Herpetosoma* species (*TH*) that contained the 623 bp amplicon (chevron). Lanes 1–14 are rodent blood samples. Samples interpreted as PCR-positive are indicated with a "+" sign.

blood samples were collected from 276 rodents of 12 different species. *Bandicota savilei* and *Rattus exulans* were the dominant species in our sampling.

Molecular tests to differentiate trypanosomes of the genus *Trypanosoma* were preferably used because morphological discrimination of trypanosomes is difficult. The problem is further compounded in the subgenus *Herpetosoma* which has more than one trypanosome species. Blood samples were, therefore, screened using universal primers in the PCR assay developed Desquesnes et al. (2002) to detect *T. lewisi* and *T. vivax* in rats. The PCR assay amplified the region that includes the internal transcribed spacer 1 (IST-1) between the 18S and 5.8S ribosomal RNA genes (rDNA) for differential diagnosis. The advantage of assay is that the relatively polymorphic ITS-1 is flanked by highly conserved rDNA, allowing detection of divergent species followed by more specific identification through amplicon sequence analysis. This assay can also be used to distinguish salivarian and stercorarian infections with 520 and 623 bp amplicons, respectively (Fig. 1).

Out of 276 rodent blood samples, 42 (15.2%) were PCR-positive for trypanosome infections (Tables 1 and 2). The prevalence of trypanosome infections in rodents was statistically different, using a chi-square test (chi-square = 123.1, $p < 0.0001$), among the six species of rodents with a theoretical prevalence over 5%: *B. indica, B. savilei, L. edwardsi, M. cookii, R. andamanensis* and *R. exulans*. The 36 PCR-positive rodents for trypanosomes belonged to three of the 12 rodent species collected (*B. indica, B.sarilei, B exulans*) and had stercorarian infections. The majority of the rodents were from Phrae province, the remainder from Loei province. Furthermore, the numbers of infected rodents collected from different environments were 14 from domestic, 12 from upland, 10 from rice field and 2 from forest. Interestingly, four rodents belonging to three species (*L. edwardsi, R. andamanensis* and *R. exulans*) were PCR-positive for salivarian trypanosomes, with a mixed infection in one *R. exulans*. Three of the salivarian-positive specimens were collected from forests, one was collected from a house suggesting that salivarian trypanosome infections are less frequent in rodents from domestic environment, Furthermore, all but one positive specimen was from Loei province while the remaining specimen was from Phrae province. None of the five specimens collected in Kalasin province were PCR-positive for trypanosomes suggesting that there is probably no trypanosome infection in rodents within the area. None of the 31 tested *Mus* spp. were found positive for trypanosomes.

Amplicons were sequenced to confirm the identity of the parasites detected with PCR. The ITS-1 portion of the sequence representative of the stercorarian amplicon was 97.9% identical to that reported for *T. lewisi* in an experimentally infected rat (Desquesnes et al., 2002) and 96.4% identical to the sequence amplified from trypanosomes isolated from rodents from the Thai infant (Sarataphan et al., 2007). The ITS-1 portion of the salivarian amplicon from trypanosomes isolated from these rodents was 96.5% identical to *T. evansi* and 41.1% identical to the sequence trypanosomes isolated from the Thai infant.

There was no effect of rodent habitat on *T. lewisi* infection (Likelihood ratio test, $P = 0.46$). However, a significant effect of habitat on *T. evansi* infection was observed (Likelihood ratio test, $P = 0.02$; analysis of deviance: model deviance = 5.4, $P = 0.02$). The relationship between *T. evansi* infection and habitat is well described by the following model (Fig. 2):

Table 1
Species and number of rodents found to be positive to *Trypanosoma lewisi* and *T. evansi* in several biotopes of Thailand

Rodent species	Biotope	Number examined	Number of positive *T. lewisi* (%)	Number of positive *T. evansi* (%)	Number of mixed positive *T. lewisi and T. evansi*
Bandicota indica	Ricefields, wet	58	10 (17.2)	0	0
Bandicota savilei	Fields, dry	64	12 (18.8)	0	1 (1.6)
Berylmys bowersi	Fields, proximity to forests, uplands	3	0	0	0
Berylmys berdmorei	Forests	1	0	0	0
Leopoldamys edwardsi	Forests	10	2 (20.0)	2 (20.0)	
Maxomys surifer	Forests	12	0	0	
Mus cervicolor	Ricefields	13	0	0	
Mus cookii	Grasslands, upland ricefields	17	0	0	
Rattus andamanensis	Forests	3	0	1 (33.3)	
Rattus exulans	Domestic	80	14 (17.5)	1 (1.3)	
Rattus losea	Ricefields, other fields	2	0	0	
Rattus tanezumi	Peri-domestic	13	0	0	
Total		276	38 (13.8)	4 (1.5)	1 (0.4)

Table 2
Trypanosoma spp. infections of rodents according to rodent habitat in 3 provinces of Thailand

	Category	Number examined	Number of positive *T. lewisi* (%)	Number of positive *T. evansi* (%)	Number of mixed positive *T. lewisi and T. evansi*
Habitat	Forests	26	2 (7.7)	3 (11.5)	
	Upland fields	67	12 (17.9)	0	1 (1.5)
	Ricefields	90	10 (11.1)	0	
	Domestic	93	14 (15.1)	1 (1.1)	
Total		276	38	4	
Province	Loei	71	3 (8.3)	3 (12.5)	
	Phrae	200	35 (25)	1 (1)	
	Kalasin	5	0	0	
Total		276	38	4	1

· Infection by *T. evansi* = exp(*X*)/(1+exp(*X*)), where *X* = −1.37–1.15 * habitat.
· The model explained 10.7% of *T. evansi* infection as dependent of habitat.

4. Discussion

Trypanosoma genus generally utilizes different vertebrate and invertebrate hosts. Thus accurate identification of these parasites and their natural hosts is needed for appropriate control of their respective zoonoses. ITS-1 sequence analysis suggests that the salivarian trypanosomes detected in this study are *T. evansi*, whose genetic diversity remains to be investigated, and the stercorarian trypanosomes in the subgenus *Herpetosoma* consists of closely related species including *T. lewisi* which recently was observed in a Thai infant. However, discrepancies among ITS-1 sequences suggest that multiple *Herpetosoma* species or strains are enzootic to Thailand. These observations need to be confirmed by future sequencing or the PCR products followed by phylogenetic reconstruction, including trypanosomes detected in domestic animals.

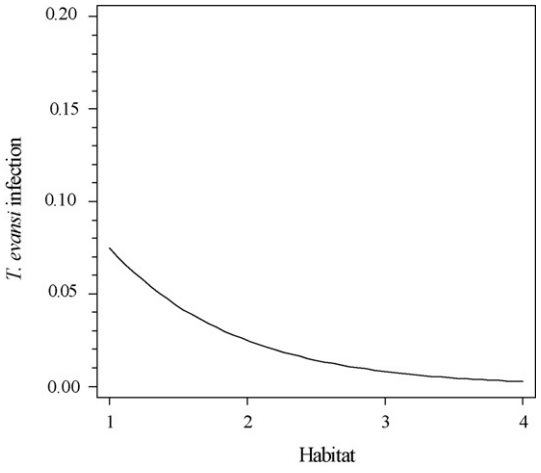


Fig. 2. Fitted logisitic regression of rodent infection by *T. evansi* in relation to the degree of habitat anthropization: (1) secondary forests, (2) upland fields, (3) rice fields and (4) domestic habitat (i.e. houses, markets).

Detection of salivarian trypanosomes in these rodents was unexpected, and, although only a few rodents are positive, it is noteworthy that all but one were collected from a forest habitat and that two out of five *L. edwardsi* collected were positive were positive for salivarian trypanosomes. This could indicate that *L. edwardsi* rodents are more susceptible to salivarian trypanosomes infection and that forest habitat may increase fly-rodent contact resulting in insect bites through which transmission is effected. We found no effect of rodent habitat on *T. lewisi* infection, whereas a significant effect of habitat on *T. evansi* infection was observed with *T. evansi* infection decreasing with habitat anthropization (i.e. from secondary forests to houses or markets). Although more investigations are needed, particularly on rodent ecology and habitat preferences, this suggests that secondary forests and upland field are places of potential risks for both humans and domestic animals. Habitat fragmentation and new land uses in uplands may favor the spread and emergence of trypanosomes.

Human trypanosomiasis associated with *Trypanosoma* spp. enzootic to Thailand is considered extremely rare, thus the recent finding of a *T. lewisi*-like (*Herpetosoma*) infection in an infant compelled us to survey different rodent populations for similar infections. In this study only three rodent species were identified as *Herpetosoma* reservoirs, one of which, *R. exulans*, is closely associated with humans. Although fleas that feed on *R. exulans* are more likely to come into contact with human hosts, the rarity of presumably flea-borne human trypanosomiasis indicates that the field rat hosts identified in this study, *B. indica* and *B. savilei*, also warrant consideration as possible sources of human infection with the *T. lewisi*-like parasite detected in the Thai infant. We observed a 13.8% prevalence of stercoraria among the apparently susceptible rodent populations.

Trypanosomes, presumably *T. evansi*, are routinely detected in buffaloes, cats, dairy cows, dogs, elephants, pigs and horses all over Thailand (Chauchanapunpol et al., 1987; Nishigawa et al., 1990; Rodthian et al., 2004), and *T. lewisi*-like parasites were reported in rats from Chiang Mai province (Natheewattana et al., 1973). *T. lewisi* and other *Herpetosoma* are commonly found in the blood of rats

worldwide, and members of this subgenus are generally considered to be non-pathogenic and rarely found in humans (Hoare, 1972). The results of this survey confirmed that some wild rodents could represent highly prevalent reservoirs of *T. lewisi*-like parasites. Our results show that habitats where we collected rodents significantly affect rodent infection, at least by *T. lewisi*, suggesting that the degree of anthropization may influence the transmission of *Trypanosoma* spp.

In summary, various rodents from different habitats appear to present risks for human exposure to trypanosomes in Thailand. The results of this study suggest that *Rattus* and *Bandicota* species are possible sources of human exposure to stercoraria, and that multiple trypanosomes species or strains within the *Herpetosoma* subgenus is enzootic to Thailand. Remaining questions include whether trypanosomes detected in these rodents are infective to flea and human populations in the same region. Further work is also needed to isolate and characterize these parasites in an experimental system and to test the vector competence of indigenous fleas.

Acknowledgements

The authors gratefully acknowledge Kasetsart University Research Development Institution (KURDI) for partial funding for publication and presentation of this work and the French program ANR-Santé-Environnement (no. 00121 0505) for supporting the fieldwork. We also thank Arthorn Boonsaner (National Park, Wildlife and Plant Conservation Department, Thailand, Bangkok), Khomson Satchasataporn and Kittiponk Chaisiri (Faculty of Veterinary Technology, Kasetsart University) for their precious assistance.

References

Chauchanapunpol, I., Withuragoon, C., Suvarnvasi, P., Nopphakhun, W., Panphan, S., Sakdeechumphol, K., 1987. Study on Trypanosoma evansi infection in dairy cattle. In: The proceeding of the Sixth Annual Conference of Livestock Development in Thailand. Department of Livestock development, Ministry of Agriculture, 18–20 January, 1986, pp. 13–20.

Desquesnes, M., Ravel, S., Cuny, G., 2002. PCR identification of *Trypanosoma lewisi*, a common parasite of laboratory rats. Kinetoplastid Biol. Dis. 1, 2.

Hoare, C.A., 1972. The Trypanosomes of Mammals. A Zoological Monograph. Blackwell Scientific Publications, Oxford and Edinburgh, p. 749.

Joshi, P.P., Shegokar, V.R., Powar, R.M., Herder, S., Katti, R., Salkar, H.R., Dani, V.S., Bhargava, A., Jannin, J., Truc, P., 2005. Human trypanosomiasis caused by *Trypanosoma evansi* in India: the first case report. Am. J. Trop. Med. Hyg. 73, 491–495.

Linardi, P.M., Botelho, J.R., 2002. Prevalence of *Trypanosoma lewisi* in *Rattus novegicus* from Belo Horizonte, State of Minas Gerais. Br. Mem. Inst. Oswaldo Cruz. Rio de Janeiro 97, 411–414.

Molyneux, D.H., 1969. Intracellular stages of *Trypanosoma lewisi* in fleas and attempts to find such stages in other trypanosome species. Parasitology 59, 737–744.

Natheewattana, N., Hongsbhanich, L., Khamboonruang, C., Thitasut, P., 1973. Preliminary study of a *Trypanosoma lewisi*-like parasite of rats in Chiang Mai. Thailand. Southeast Asian J. Trop. Med. Public Health 4, 322–323.

Nishigawa, H., Sarataphan, N., Tuntasuvan, D., Neramitmansuk, P., 1990. Serological survey of Trypansomiasis and babesiosis in cattle and buffaloes in Thailand. In: The proceedings of the seventh FAVA Congress, Pattaya, Thailand, November 4–7, p. 1990.

Rodthian, P., Hinon, W., Uthaiwan, W., Withuragoon, P., Taritsanarom, U., Chaisert, S., Khaihong, S., Sarataphan, N., Muangyai, M., 2004. A case report: *Trypanosoma evansi* in draft elephants in Lamphang. In: The 19th Annual Conference of Livestock Development, Pathumthani. Thailand, May 26–30.

Sarataphan, N., Vongpakorn, M., Nuansrichay, B., Autarkool, N., Keowkarnkah, T., Rodtian, P., Stich, R.W., Jittapalapong, S., 2007. Diagnosis of a *Trypanosoma lewisi*-like infection in a sick infant from Thailand. J. Med. Microbiol. 56, 1118–1121.

Shrivastava, K.K., Shrivastava, G.P., 1974. Two cases of *Trypanosoma* (*Herpetosoma*) species infection of man in India. Trans. Royal Soc. Trop. Med. Hyg. 68, 143–144.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 450–458
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Association between the DQA MHC class II gene and Puumala virus infection in *Myodes glareolus*, the bank vole

Julie Deter [a,*], Josef Bryja [b], Yannick Chaval [a], Maxime Galan [a], Heikki Henttonen [c], Juha Laakkonen [c,d], Liina Voutilainen [c], Olli Vapalahti [d], Antti Vaheri [d], Alexis Ribas Salvador [e], Serge Morand [f], Jean-François Cosson [a], Nathalie Charbonnel [a]

[a] *Centre de Biologie et de Gestion et des Populations (CBGP), Département INRA-EFPA 1062, Campus International de Baillarguet, CS 30016, 34988 Montferrier sur lez, France*
[b] *Department of Population Biology, Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, 67502 Studenec 122, Czech Republic*
[c] *Vantaa Research Unit, Finnish Forest Research Institute, PO Box 18, FIN-01301 Vantaa, Finland*
[d] *Department of Virology, Haartman Institute, PO Box 21, FIN-00014 University of Helsinki, Finland*
[e] *Laboratori de Parasitologia, Departament de Microbiologia i Parasitologia Sanitàries, Facultat de Farmàcia, Universitat de Barcelona, Av. Joan XXIII s/n, E-08028 Barcelona, Spain*
[f] *Institut des Sciences de l'Evolution, CNRS, CC064, Université Montpellier 2, 34095, Montpellier Cedex 05, France*

Received 25 April 2007; received in revised form 2 July 2007; accepted 5 July 2007
Available online 10 July 2007

Abstract

Puumala virus (PUUV) is a hantavirus specifically harboured by the bank vole, *Myodes* (earlier *Clethrionomys*) *glareolus*. It causes a mild form of hemorrhagic fever with renal syndrome (HFRS) in humans, called *Nephropathia epidemica* (NE). The clinical severity of NE is variable among patients and depends on their major histocompatibility complex (MHC) genetic background. In this study we investigated the potential role of class II MHC gene polymorphism in the susceptibility/resistance to PUUV in the wild reservoir *M. glareolus*. We performed an association study between the exon 2 of the DQA gene and PUUV antibodies considering a natural population of bank voles. Because immune gene polymorphism is likely to be driven by multiple parasites in the wild, we also screened bank voles for other potential viral and parasitic infections. We used multivariate analyses to explore *DQA* polymorphism/PUUV associations while considering the potential antagonist and/or synergistic effects of the whole parasite community. Our study suggests links between class II MHC characteristics and viral infections including PUUV and Cowpox virus. Several alleles are likely to be involved in the susceptibility or in the resistance of bank voles to these infections. Alternatively, heterozygosity does not seem to be associated with PUUV or any other parasite infections. This result thus provides no evidence in favour of the hypothesis of selection through overdominance. Finally this multivariate approach reveals a strong antagonism between ectoparasitic mites and PUUV, suggesting direct or indirect immunogenetic links between infections by these parasites. Other datasets are now required to confirm these results and to test whether the associations vary in space and/or time.

Keywords: *Myodes* (*Clethrionomys*) *glareolus*; Hantavirus; SSCP; Coinertia; Helminths; Mites; Cowpox virus

1. Introduction

1.1. Immunogenetics and zoonoses

Polymorphism at immune genes may be involved in specific recognition of pathogens as variation allows the recognition of spectra of epitopes (Doherty and Zinkernagel, 1975). In this context, one of the leading goals of immunogenetics, i.e. the analysis of genetic polymorphisms in specific recognition and immune regulation, has been to understand the genetic basis of susceptibility to complex diseases (Geraghty, 2002). The profound influence of the host genetics on resistance to infections has been established in numerous studies, which mainly concerned human infections such as malaria, HIV and hepatitis (review in Cooke and Hill, 2001; Hill, 2001). In wild

* Corresponding author. Tel.: +33 4 99 62 33 46; fax: +33 4 99 62 33 45.
E-mail address: deter@supagro.inra.fr (J. Deter).

1567-1348/$ – see front matter
doi:10.1016/j.meegid.2007.07.003

animal populations, immunogenetics provides key insight into the relative influence of genetic variation and environmental factors on host-pathogen interactions. A number of zoonoses, i.e. infections transmitted from animals to humans (Taylor et al., 2001), has (re-)emerged during the past 15 years (Kallio-Kokko et al., 2005). Therefore, an important application of immunogenetics concerns the assessment of emergent or re-emergent disease risks in natural populations. (i) The study of immune genes may explain why hosts differ in their susceptibility to different parasites. The reasons why hosts differ in their susceptibility/resistance to different parasites could rely on the degree of matching between immune genes and parasite antigens. Immunogenetics may thus contribute to the identification of unknown zoonotic agents, which is essential for understanding zoonoses epidemiology (Mills and Childs, 1998). (ii) Natural population studies of immune gene polymorphism may provide key insight into the factors determining the appearance, spread and distribution of resistance/immuno-modulating alleles within populations and across geographical landscapes. Such information is essential to study the spatial and temporal variations in disease risk or incidence.

1.2. Puumala virus and hemorrhagic fever with renal syndrome

Hantaviruses are one of the three main emerging infectious agents in Europe (Kallio-Kokko et al., 2005). They are specific rodent-borne viruses belonging to the Bunyaviridae family, and are transmitted to humans primarily from aerosols of rodent excreta. Puumala virus (PUUV) is a hantavirus, which causes a mild form of haemorrhagic fever with renal syndrome (HFRS) in humans, called *Nephropathia epidemica* (NE). The mortality rate ranges between 0.1 and 0.4% (Vapalahti et al., 2003). Infections are frequent in Northern Europe, European Russia and parts of Central Western Europe (Vapalahti et al., 2003), with about 6000 cases reported per year. In France, it has been classified as an important emerging disease by the Institut de Veille Sanitaire (INVS, Capek, 2006).

The specific reservoir host of PUUV is the bank vole *Myodes* (earlier *Clethrionomys*) *glareolus*. In this rodent, PUUV produces a chronic, long-lasting infection. Host and virus have geographical ranges that do not completely overlap. Many parts of Europe, such as the Mediterranean peninsulas and Britain remain blank of PUUV infections although *M. glareolus* is present. Two recent studies provided arguments in favour of selection acting on bank voles and mediated by PUUV. First, Kallio et al. (2006a,b) observed that PUUV infection could affect bank vole fecundity (earlier breeding). Second, experimental infections of *M. glareolus* laboratory colonies with PUUV revealed that some bank voles could not be successfully infected. Neither antibodies, nor PUUV RNA could be detected in these individuals reared in PUUV-contaminated beddings, providing evidence of variability in vole susceptibility to PUUV. This variability was not significantly explained by age, sex, maturation status or individual weight (Kallio et al., 2006a,b). This suggests that immunogenetic factors could be involved not only in the disease progression (Klein et al., 2004a) but also in controlling infection itself. The absence of HFRS in some parts of *M. glareolus* geographic distribution could therefore be partly explained by selected vole immunogenetic factors associated with resistance to PUUV infections.

1.3. Immunogenetic factors associated with Puumala virus infections

Association studies have provided evidence for the role of immunogenetics in the course of hantavirus infections in rodents. Using the hantavirus *Seoul* and the rat *Rattus norvegicus*, Klein et al. (2004a,b) characterized immunologic pathways that differ between males and females in response to infection. They showed that most of the 192 identified immune-related genes that encode for these immune responses were differentially expressed in infected females compared with infected males. Several MHC genes were associated with these sex differences in gene expression. Alternatively, the considerable variability in the clinical severity of NE among humans depends on MHC genes (Mustonen et al., 1996, 1998, 2004; Plyusnin et al., 1997). The HLA-B227 haplotype is associated with a mild course of NE, whereas the extended haplotypes HLA-B8-DR3 and HLA-DRB1-0301 are associated with severe clinical courses. The complete MHC class II region has been sequenced in rats (Günther and Walter, 2001; Hurt et al., 2004) and mice (Blake et al., 2003). It has a similar regional organization as the human MHC, and orthologous relationships exist between class II regions in all mammals (Takahata and Nei, 1990). Moreover, in laboratory mouse and rats, the DQA gene (called RT1.B in rats, H2-A in mice) is tightly linked to DRB genes (Blake et al., 2003; Hurt et al., 2004). As DRB is at least quadriduplicated in bank voles (Axtner and Sommer, 2007), what makes genotyping difficult to perform, DQA was a relevant candidate to look for associations between MHC class II genes and PUUV infection.

In this study we investigate the role of the class II MHC *DQA* gene polymorphism on susceptibility/resistance to PUUV by the mean of an association study conducted in a natural population of *Myodes glareolus*, the specific reservoir of this hantavirus. Vole sampling was performed within the second most important French endemic area of HFRS (Jura, Franche Comté), which has had an unusually high number of human cases in 2005 (230 in France, 38 in the Jura, INVS).

2. Materials and methods

2.1. Trapping design

The main study area is located around Mignovillard (46°8′N, 6°13′E and elevation 850 m) in Jura, France. It consists of a 4-km^2 site, half of which is composed of wooded meadows and the other half of forests (man-made spruce or semi-natural forests). Bank voles were sampled in September 2004, July and September 2005 using French Agricultural Research Institute (INRA) live traps, which were fitted out with dormitory boxes,

baited with hay and a piece of apple. Twenty 100-m trap-lines composed of 34 traps placed at 3-m intervals were equally distributed in open (meadows and hedges, nine lines) and wooded (five lines in forests and six lines in meadows in border of forest) areas. For each trapping session, these lines were checked twice a day during three consecutive nights.

2.2. Infection screening

Once trapped, rodents were killed by cervical dislocation as recommended by Mills et al. (1995). For each individual, blood samples were taken from the heart or the thoracic cavity on a 1-cm^2 Whatman® blotting paper. In the laboratory, each piece of 1-cm^2 paper with dried blood was placed in 1 ml phosphate-buffered saline. These diluted blood samples were screened for IgG antibodies to Puumala virus (PUUV) using immunofluorescence assays (IFA) as described in Lundkvist et al. (1991) (see also Klingström et al., 2004; Kallio-Kokko et al., 2005). As PUUV produces a long-lasting, chronic infection in bank voles, the detection of antibodies for PUUV is strongly related with the presence of the virus except for very young voles because of possible maternal antibodies.

Because immune gene polymorphism is likely to be driven by multiple parasites in the wild (Chantrey et al., 1999; Wegner et al., 2003), we also screened bank voles for other potential viral and parasitic infections. *M. glareolus* is considered as one of the main reservoir of Cowpox virus (CPXV, Chantrey et al., 1999). Moreover, recent serological surveys have revealed the potential role of *M. glareolus* as a reservoir of the Lymphocytic Choriomeningitis virus (LCMV) (Kallio-Kokko et al., 2006). We thus used vole blood samples to screen for IgG antibodies against CPXV and LCMV by IFA as described in (Kallio-Kokko et al., 2005, 2006; Laakkonen et al., 2006). On the field, the presence of ectoparasitic mites was visually inspected and noted. Whole fresh brains were squashed and checked microscopically for the presence of *Frenkelia* cysts (Coccidia, Apicomplexa: Sarcocystidae). Finally, we looked for the presence of larval cestodes in the liver and in the body cavity of voles. The digestive tracts (stomach, intestines) were stored in 95% ethanol before being dissected in the laboratory. All the helminths detected were carefully counted under the microscope and then identified in Barcelona.

2.3. MHC class II gene genotyping

Genomic DNA was extracted from toe tissues by Puregene DNA purification kit® (Gentra Systems) according to manufacturer's instructions and finally eluted in 400-μl of water. We amplified the complete exon 2 of *DQA* (called *RT1.Ba* in rat) homologue of the MHC class II as described in Bryja et al. (2005, 2006). A single PCR (PCR1 in Bryja et al. (2006)) was performed on an Epgradient® thermocycler (Eppendorf) as it allowed the amplification of all bank vole alleles. Capillary electophoreses-single strand conformation polymorphism (CE-SSCP) were performed on a MegaBACE 1000 DNA Analysis system® (Amersham Biosciences) as described in Bryja et al. (2006). All alleles detected following the CE-SSCP analyses were cloned and sequenced using a minimal subset of animals exhibiting the diverse SSCP patterns observed. Briefly, the *DQA* gene was amplified as described above, but using non-labelled primers. The PCR products were cloned and purified by the QIAquick PCR purification kit® (Promega) as described in Bryja et al. (2006). Eight clones were subsequently isolated for each individual selected. The plasmid DNA was extracted using QIAprep Spin Miniprep kit® (Qiagen). The *DQA* insert was sequenced using DYEnamic ET terminator cycle sequencing kit® (Amersham Biosciences) and primer SP6. Sequences were performed on MegaBACE 1000 DNA analysis system® (Amersham Biosciences). They were edited and aligned using BioEdit Sequence Alignment Editor 7.0.5.2 (Hall, 1999). A new sequence variant was considered as a new allele when it met the criteria summarised by Kennedy et al. (2002). Paup version 4.0b10 for Microsoft Windows 95/NT was employed to construct a phylogenetic tree of the *DQA* amino acid sequences based on 247 bp of the exon 2 (base pairs no. 3-249 of the complete exon), applying the neighbour-joining (NJ) algorithm with Kimura's two-parameter distances for nucleotides, and the Poisson correction for amino acids. Bootstrap analyses (5000 replicates) were performed to determine the reliability of the branching. *R. norvegicus* (GenBank accession no. AJ554214) and *Mus musculus* (no. K01924) sequences were used as outgroups in the phylogenetic analyses.

2.4. Statistical analyses

Seroprevalence (percentage of seropositive animals) for each virus and prevalence (percentage of infected animals) for each parasite were calculated and 95% confidence intervals were estimated by bootstraps using the software Quantitative Parasitology 3.0 (Rozsa et al., 2000).

We explored the immunologic factors (*DQA* alleles and heterozygosis) potentially associated with PUUV infection status using multivariate analyses, which affords the integration of other viruses, ectoparasites, coccidia and helminths. This gave the opportunity to explore at the same time the antagonist and/or synergistic effects of other parasites on *DQA* polymorphism/PUUV associations. We performed a co-inertia analysis (ACO) between (i) a genetic matrix including the presence/absence of each *DQA* allele and the *DQA* heterozygosis for each individual, and (ii) a parasitological matrix including the abundance (when available) or presence of each parasite species and the presence of antibodies against each virus for each individual. The ACO links the independent multivariate analyses of the genetic matrix (correspondence analysis, COA) and of the parasitological matrix (normalised principal component analysis, PCA). It considers all types of variables and it is robust to correlations between the variables of a matrix (Dray et al., 2003). The ACO gives co-inertia axes that maximize the covariance between the row coordinates of the two matrices (Dray et al., 2003; Thioulouse et al., 2004). The genetic and parasitological results can be compared by superimposing both kinds of variables on the ACO factor map. Therefore, particular associations were visually detected

using the distribution of genetic and parasitological variables on the F1 × F2 ACO factor map. When the two first axes explained a large part of the co-inertia, the variables located in a given direction relative to the origin were considered positively associated, whereas the variables located in opposite directions were considered antagonistic. Variables plotted near the origin could not be interpreted. The ACO was performed using ADE-4 software (Thioulouse et al., 1997). The global relationship between the two matrices was investigated by comparing the ACO previously estimated to the distribution of co-inertia estimated from 1000 permutations of the parasitological matrix rows. Considering PUUV, the relative risk (RR) associated with the alleles of interest detected using the ACO was estimated following Haldane (1956).

Besides revealing associations between immunogenetics and PUUV infection, this multivariate analysis allows to discriminate two mechanisms underlying selection mediated by pathogens/parasites and acting on *DQA* gene. First, under the hypothesis of frequency-dependent selection (Takahata and Nei, 1990), we expect positive or negative associations between the presence of specific allele and parasite species. This association could be interpreted in terms of resistance (negative association) or susceptibility (positive) alleles to parasite species. Second, under the hypothesis of heterozygosis advantage, i.e. overdominance (Doherty and Zinkernagel, 1975; Hughes and Nei, 1990), heterozygous voles are expected to be less parasitized than homozygous ones.

3. Results

The 98 bank voles, 53 males and 45 females, trapped between September 2004 and 2005 in Mignovillard were included in the analyses.

3.1. Infection screening

The IFA analyses revealed nine (9.2%) PUUV-seropositive bank voles. The sequencing of the hantavirus involved has been described elsewhere on the basis of the S-segment (Plyusnina et al., in press) and confirmed the presence of PUUV in bank voles. Fifteen (15%) of the bank voles were CPXV seropositive. No antibodies against LCMV were detected. Orange-coloured mites were found in the ear lobes of 48 bank voles. These were identified as larvae of the trombiculid *Neotrombicula inopinata* (Acari, Actinotrichida, Prostigmata). Sporocysts of *Frenkelia glareoli* (Apicomplexa: Sarcocystidae) were found in the brain of a single bank vole. Larvae of two cestodes, *Echinococcus multilocularis* (Cestoda: Cyclophyllidae) and *Taenia taeniaeformis* (Cestoda: Cyclophyllidae) were found encysted in the liver of respectively 11 and 29 voles. In gastro-intestinal tracts, we found the following helminths *Paranoplocephala gracilis* (Cestoda: Anoplocephalidae), *Catenotaenia henttoneni* (Cestoda: Catenotaeniidae), *Mastophorus muris* (Nematoda: Spirocercidae), *Hymenolepis horrida* (Cestoda: Hymenolepididae), *Trichuris arvicolae* (Nematoda: Trichuridae) and *Syphacia petrusewiczi* (Nematoda: Oxyuridae). Details concerning the prevalence levels are presented in Table 1.

Table 1
Parasitological data concerning the 98 bank voles sampled

Parasites	Code	*N*	Prevalence (CI)
Viruses			
Puumala virus	PUUV	9	9.2 (4.87–15.48)
Cowpox virus	CPXV	15	15.3 (9.67–22.59)
Macroparasites and coccidia			
Neotrombicula inopinata	Mites	48	49.0 (40.27–50.74)
Frenkelia glareoli	Cocc-Fg	1	1.0 (0.05–4.75)
Echinococcus multilocularis	Cest-Em	11	11.2 (6.42–17.90)
Taenia taeniaeformis	Cest-Tt	29	29.6 (22.04–38.09)
Paranoplocephala gracilis	Cest-Pg	9	9.2 (4.87–15.48)
Catenotaenia henttoneni	Cest-Ch	16	16.3 (10.51–23.74)
Mastophorus muris	Nem-Mm	13	13.3 (8.02–20.27)
Hymenolepis horrida	Nem-Hh	5	5.1 (2.03–10.43)
Trichuris arvicolae	Nem-Ta	3	3.1 (0.83–7.73)
Syphacia petrusewiczi	Nem-Sn	4	4.1 (1.40–9.10)

Cocc, Cest and Nem, respectively, refer to coccidia, cestodes and nematodes. *N* is the number of seropositive (viruses) or infected (macroparasites, protozoans) animals. Prevalence estimates are presented in percentage. CI corresponds to the 95% confidence interval of prevalence estimates.

3.2. DQA genotyping

The screening of DQA polymorphism using CE-SSCP for the 98 bank voles sampled revealed nine different SSCP patterns. Seventeen individuals were cloned and 136 clones were sequenced to constitute the DNA bank. After applying the criteria described by Kennedy et al. (2002), we defined nine SSCP patterns (alleles) corresponding to nine different sequences of the complete *DQA* exon 2. Nomenclature is Clgl-*DQA* as already used in Bryja et al. (2006) and proposed by Klein et al. (1990). The sequences of the *DQA* exon 2 are available in GenBank under the following accession nos. EU008329, EU008330, EU008331 and EU008332. Phylogenetic analyses resulted in a single shortest tree for both nucleotides (not shown) and amino acids (Fig. 1). All exon 2 sequences identified translated in unique amino acid sequences. This reflects the high level of non-synonymous substitutions at this gene. As previously described in Bryja et al. (2006), one to four different alleles could be observed for a single bank vole,

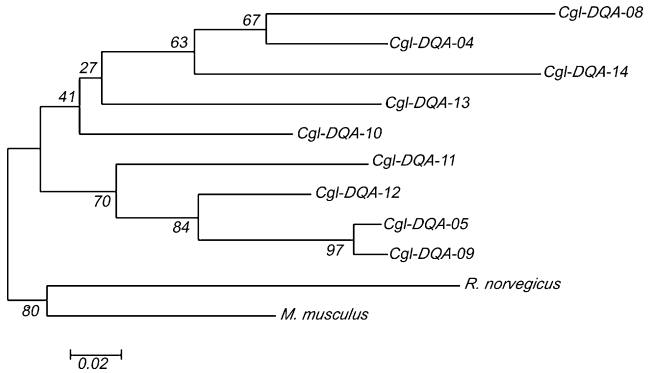


Fig. 1. Neighbour-joining phylogenies for the nine *M. glareolus* amino acid sequences using *Rattus norvegicus* and *Mus musculus* sequences as outgroups. The scale bar represents the genetic distance and the numbers are bootstrap values (5000 replicates).

Table 2
Clgl-*DQA* allele's frequencies in bank voles trapped in Mignovillard

Clgl-*DQA* alleles	Frequencies (%)
4	3
5	16
8	83
9	11
10	16
11	21
12	7
13	1
14	14

$N = 98$. Frequencies are calculated in terms of number of individuals presenting alleles compared to the total dataset.

indicating both the duplication of the *DQA* gene in *M. glareolus* and the intra-population polymorphism of this duplication. Based on the phylogenetic analyses, it was not possible to assign the nine different alleles to a particular copy of the *DQA* gene. The duplication event in *M. glareolus* MHC probably took place after the separation of Cricetidae and Muridae (see also Bryja et al., 2006). Therefore, the duplication polymorphism was not included as a measure of *DQA* diversity in the following analyses. Allelic frequencies, estimated from the presence/absence of each allele per individual, are detailed in Table 2.

3.3. Statistical analyses

The coccidium *Frenkelia glareoli* and the allele Clgl-*DQA*13 were removed from the association analysis because they were only found in one bank vole.

The first two dimensions of the COA explained 39% of the variance of genetic data (first factor F1: 21%; second factor F2: 18%). The projection of the genetic variables on the F1/F2 plane revealed that heterozygosis, Cgl-*DQA*-08 and Cgl-*DQA*-10 were located close to the origin, and thus did not discriminate bank voles (Fig. 2A). The alleles Cgl-*DQA*-05 and Cgl-*DQA*-11 or Cgl-*DQA*-04 were located in opposite directions on F1 and exhibited the highest factorial values. They have therefore antagonist effects. The alleles Cgl-*DQA*-05 and Cgl-*DQA*-14 were opposed on F2 (Fig. 2A). Fig. 2B shows the PCA correlation circle for the parasitological variables projected on the F1/F2 plane. These first two axes accounted for 32% of the variance, with 17% explained on the first factor F1 and 15% on the second one F2. The first factor opposed bank voles infected with different helminths while the second factor (F2) divided voles in a group including PUUV seropositive ones and another group including voles infected with mites. Note that most of the variables were located near the centre of the factor map, which means that they have low factorial values.

These independent results were coupled using ACO. The global co-inertia between the genetic and parasitological matrices reached 32%. After 10,000 permutations, 2547 co-inertia estimates were larger than the one observed in our data. This indicates a poor global relationship between the two matrices. We further analysed particular associations between PUUV and *DQA* gene polymorphism. The two first axes of the

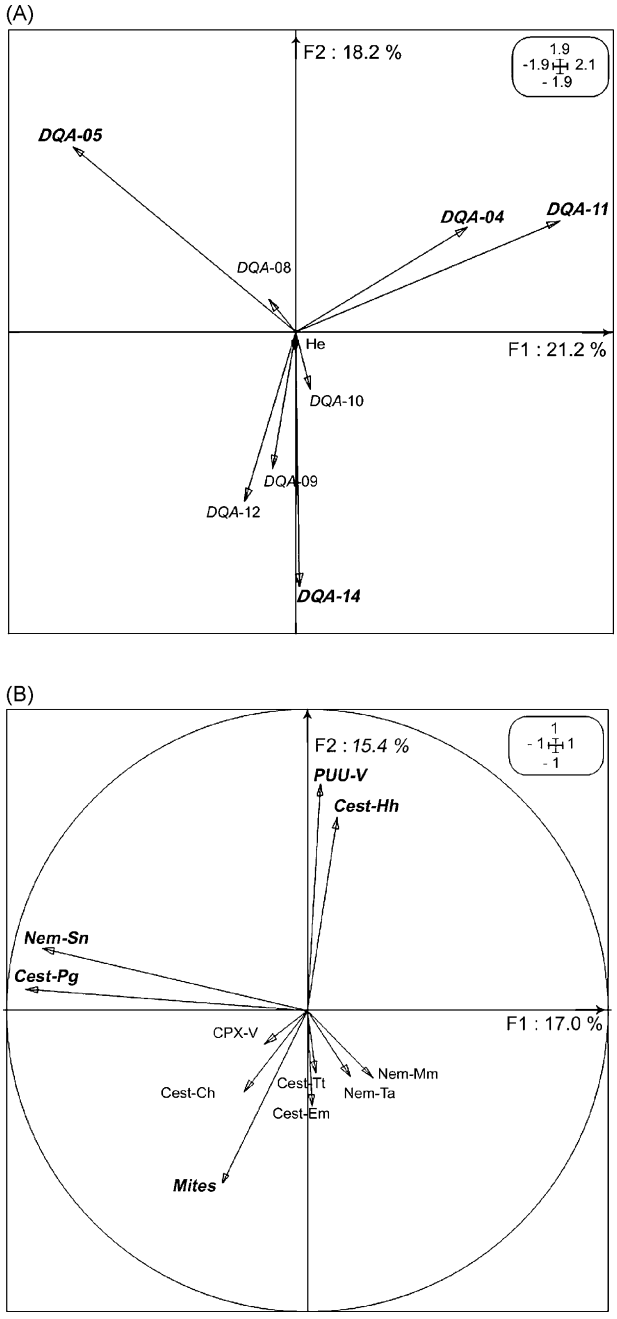


Fig. 2. Multivariate analyses of the genetic and parasitological matrices analysed independently. Variables are projected on the factor map defined by factor 1 (F1) and factor 2 (F2). F1 explains most of the variance and F2, which is orthogonal to F1, explains the most of the remaining variance. For each analysis, the variance explained by each factor is indicated in percentage. For each pair of axes, the variables located in a given direction relative to the origin can be considered positively associated, whereas the variables located in opposite directions can be considered antagonistic. The main variables explaining the total variance are edited in bold. Variables plotted near the origin cannot be interpreted. (A) Correspondence analysis (COA) of *DQA* gene polymorphism in the 98 bank voles. The *DQA* alleles, coded Cgl-*DQA*-XX, and heterozygosity (Het), are projected on the F1/F2 map. (B) Correlation circle (F1 × F2) of the principal component analysis (PCA) performed on the parasitological matrix. Nem and Cest, respectively, indicate nematode or cestode. The two following letters refer to the code used to identify the species (see Table 1).

co-inertia represented 58% of the variance (31% F1 and 27% F2). Fig. 3A shows the ACO factor map for the genetic variables. The distribution of genetic variables on this map was almost similar to the one observed on the previous COA factor map (Fig. 2A). Two changes could be noticed. First, although the Cgl-*DQA*-10 allele has not been important in the COA, it strongly structured the second factor of the ACO. Second, the Cgl-*DQA*-12 and Cgl-*DQA*-05 have been negatively associated on the second factor of the COA, but they were positively associated on the first factor of the ACO. Fig. 3B shows the distribution of the parasitological variables on the F1 × F2 ACO factor map. PUUV infection played a role in the co-inertia between genetic and parasitological matrices as PUUV serological status structured the F2 axis, with CPXV serological status. A surprising result was the opposition between PUUV infection and the presence of mites. This opposition was previously observed from the PCA analysis. Major changes between the PCA and ACO factor maps can be noted. CPXV antibodies appeared important in the co-inertia analysis although they did not structure the PCA factor map. Inversely the cestode *H. horrida*, which was important in the PCA and positively associated with PUUV infection, was no more important in the co-inertia analysis.

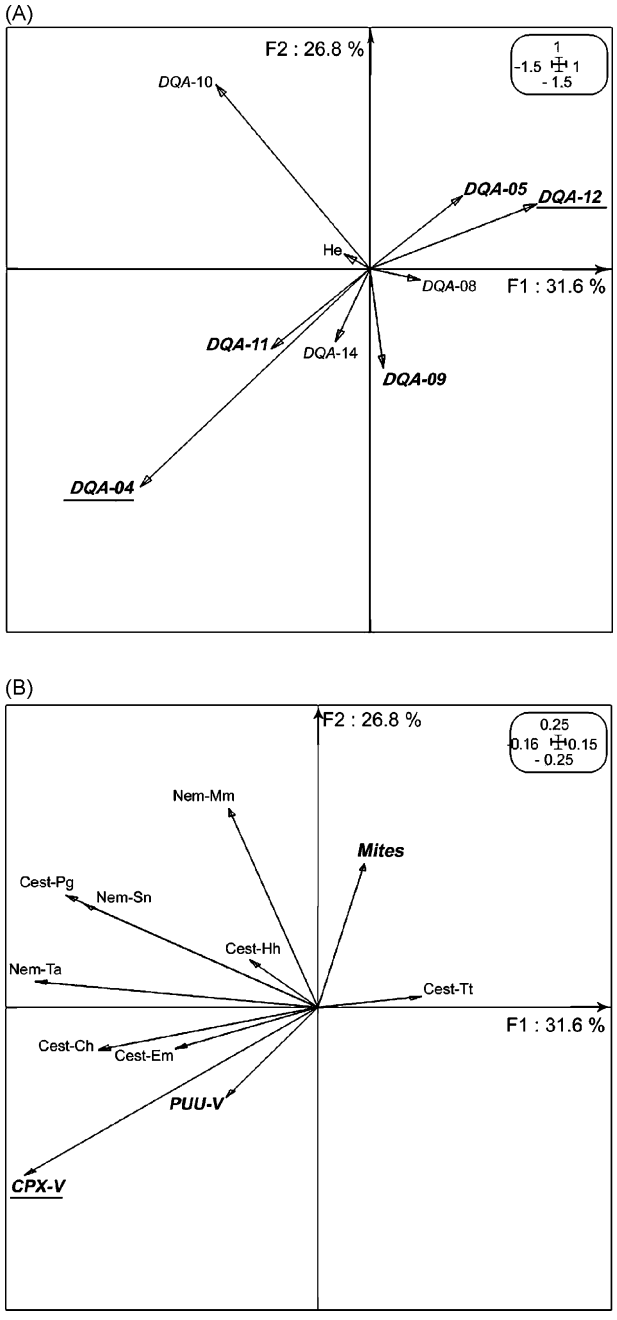


Fig. 3. Distribution of (A) the genetic variables encoded as in Fig. 2A, and (B) the parasitological variables encoded as in Fig. 2B on the first factor map (F1 × F2, defined as in Fig. 2) of the ACO. Under strong co-inertia, the variables located in a given direction relative to the origin can be considered positively associated, whereas the variables located in opposite directions can be considered antagonistic. The genetic and parasitologic variables associated with PUUV in the ACO are edited in bold style. Those associated with CPXV are also underlined.

The comparison between Figs. 3A and B revealed that the main genetic variables associated with PUUV antibodies are Cgl-*DQA*-05 and Cgl-*DQA*-12 (negative associations, RR = 0.21 and 0.57, respectively: individuals carrying these alleles have respectively a 0.21 and 0.57 times lower risk to get infected), and in a lesser extent Cgl-*DQA*-09 and Cgl-*DQA*-11 (positive association, RR = 2.83 and 2.07). This comparison also stressed the strong genetic–parasitological associations between CPXV antibodies and both Cgl-*DQA*-04 (positive association, RR = 10.18) and Cgl-*DQA*-12 (negative association, RR = 0.32). This supports the important role of CPXV serological status in the co-inertia between parasitism and MHC polymorphism. Details concerning the number of bank voles within each of these categories are provided in Table 3.

Interestingly, all four alleles mainly implicated in the association with PUUV antibodies, Cgl-*DQA*-05, Cgl-*DQA*-12 and Cgl-*DQA*-09 clustered in a well-sustained clade (bootstrap value >85, Fig. 1). Despite their phylogenetic proximity, the three alleles differed by ten amino acid sites (over a total of 81 sites). Moreover, 5 amino acid substitutions occurred within the 11 sites, which were previously found to be under positive selection in rodents (see Bryja et al., 2006). The two closest alleles, Cgl-*DQA*-05 and Cgl-*DQA*-09, which were involved in antagonistic associations with regard to PUUV, were separated by only two amino acid changes, one of them belonging to a site under positive selection.

4. Discussion

In humans, the clinical severity of NE depends on the MHC genetic background of the patients, with haplotypes associated to a benign clinical course and other with a severe one (Mustonen et al., 1996, 1998, 2004; Plyusnin et al., 1997). Investigating the existence of such associations in reservoirs of hantaviruses is thus particularly important to better understand the risks of HFRS emergence related to wild rodents populations.

4.1. Associations between PUUV and DQA alleles

Our study revealed associations between a class II MHC gene and PUUV infection in bank voles. More precisely, we

Table 3
Correspondence table between genetic factors selected from the co-inertia analysis and (A) Puumala virus antibodies or (B) Cowpox virus antibodies

	Cgl-*DQA*-09		Cgl-*DQA*-05		Cgl-*DQA*-12		Cgl-*DQA*-11	
	Presence	Absence	Presence	Absence	Presence	Absence	Presence	Absence
(A) Puumala virus antibodies								
Presence (%)	2 (18.19)	7 (8.04)	0 (0)	9 (10.98)	0 (0)	9 (9.89)	3 (14.28)	6 (7.79)
Absence (%)	9 (81.81)	80 (91.96)	16 (100)	73 (89.02)	7 (100)	82 (90.11)	18 (85.72)	71 (92.21)
RR	2.83		0.21		0.57		2.07	

	Cgl-*DQA*-04		Cgl-*DQA*-12	
	Presence	Absence	Presence	Absence
(B) Cowpox virus antibodies				
Presence (%)	2 (66.67)	13 (13.68)	0 (0)	15 (16.48)
Absence (%)	1 (33.33)	82 (86.32)	7 (100)	76 (83.51)
RR	10.18		0.32	

The number in each cell shows the sample size. Seroprevalence distribution (percentage of seronegative and seropositive individuals) is given both for individuals with and without each genetic factor identified. RR corresponds to the relative risk associated with each allele, estimated from Haldane (1956). Alleles are grouped in the table depending on the sign of the association (positive: Cgl-*DQA*-09 and 04 or negative: Cgl-*DQA*-05 and 12); Cgl-*DQA*-11 is also positively associated with PUUV antibodies but in a lesser extent (see ACO, Fig. 3).

detected a positive association between PUUV antibodies and the allele Cgl-*DQA*-09, and negative associations with the alleles Cgl-*DQA*-05 and Cgl-*DQA*-12. Considering the allele Cgl-*DQA*-09, two individuals exhibit both the alleles and PUUV antibodies. There was a high proportion of seronegative voles that did not carry this allele, as well as a low proportion of voles carrying this allele and being seronegative. This suggests that voles carrying the Cgl-*DQA*-09 allele could be more susceptible than those which did not carry it. Alternatively, none of the PUUV seropositive voles carried either the alleles Cgl-*DQA*-05 (nine individuals) or Cgl-*DQA*-12 (nine individuals). This could be interpreted as a ''resistance'' role of these alleles in PUUV infection. However, this interpretation requires confirmation through experimental infection as PUUV seronegative individuals might be resistant voles or voles which have not yet met the virus.

In humans, the class I HLA-B8, the class II *DRB**0301 and *DQ2* alleles are associated with a severe course of NE (Mustonen et al., 1996). The precise mechanisms underlying these associations are yet not well described but different hypotheses have been developed (review in Terajima et al., 2004). Individuals that carried this extended haplotype have impaired lymphocyte activation and an imbalance in cytokine production (Mustonen et al., 1998). It has been proposed that the antiviral response either may be dysfunctional and fail to clear the virus, or may also function with inappropriate efficacy: the HLA-B8-DRB3 haplotype is associated with autoimmune diseases that may have a viral etiology. Molecular mimicry could explain the association of this haplotype with the renal disease caused by Puumala virus infection (Plyusnin et al., 1997).

However, these associations might be indirect, if this polymorphism is in linkage disequilibrium with a functional polymorphism in other genes located within the MHC or that predispose voles to PUUV infection. A potential candidate is the TNF-α gene, located within the class III region of the MHC. Increased levels of TNF-α are known to induce pathophysiological and clinical changes similar to those seen in NE. Polymorphism in the promoter region of this gene is thus likely to be associated with the severity of NE. A recent study showed that patients carrying the TNF-2 allele, which is associated with enhanced TNF-α transcriptional activity, suffered from a more severe NE (Mäkelä et al., 2002). However, it is likely that this result is due to strong linkage disequilibrium with the HLA haplotype previously mentioned. The HLA polymorphism seems to be a more important risk factor for severe NE than the TNF-α polymorphism (Mäkelä et al., 2002).

Experimental infections and immunological surveys need to be investigated in the bank vole to confirm the existence of resistant/susceptible individuals or the variability in the outcome of the infection with PUUV, and to study the relative influence of DQA polymorphism versus other gene polymorphism on PUUV infection.

4.2. No associations between PUUV and DQA heterozygosis

The alternative hypothesis potentially underlying associations between MHC polymorphism and PUUV infection is overdominance. Doherty and Zinkernagel (1975) suggested that heterozygous individuals at MHC loci could have an immunologic advantage because of their ability to present a wider range of foreign peptides to T cells. Under this scenario, we expected heterozygous voles to be less parasitized than homozygous ones. Our results do not provide evidence in favour of this heterozygote advantage hypothesis. Nevertheless, we have to be cautious with these results. Our dataset concerns a single site and might not be large enough to cover the whole pathogen community. The absence of associations between heterozygosity and parasitism might thus be confirmed before concluding about the importance of overdominance as a mechanism of selection acting on DQA gene in bank voles.

4.3. Antagonist infections with mites and PUUV

Analysing PUUV infections jointly with the whole helminth community, viruses and protozoan infections gave us the opportunity to detect antagonist and/or synergistic effects of parasites on the PUUV/MHC associations observed. In this study, the multivariate approach developed on parasitological variables revealed a strong opposition between mites and PUUV infections. Among the nine PUUV seropositive voles, only two were infected by mites. Alternatively, among the PUUV seronegative voles, the proportions of individuals infected and non-infected by mites were similar (respectively, 51% and 48%). This contrast between mites and PUUV infections was still observed when analysing the co-structure between parasitological and genetic variables. This suggests an antagonist role of immunogenetics in resistance/susceptibility to mites and PUUV. To our knowledge, this is the first time that this antagonism is highlighted. It is quite surprising as a role of trombiculid mites in hantavirus transmission has even been suggested (Song, 1999; Houck et al., 2001). Several explanations might be envisaged and tested experimentally: (i) one of the immune response developed against PUUV infection and associated with MHC susceptibility factor in voles might prevent from further mite infestations, (ii) one of the immune pathway associated with MHC factor of resistance to PUUV infection might not prevent from further mite infestations, and (iii) trombiculid mites are vectors of microparasites, which were not detected in this study. The immune response developed against these microparasites might prevent further infestations with PUUV. This hypothesis is worth investigating as trombiculid mites are known vectors of bacteria such as *Borrelia* sp. (Kampen et al., 2004) or *Rickettsia* sp. (Frances et al., 2000).

4.4. Associations between Cowpox virus and DQA polymorphism

Surprisingly, our results show that CPXV plays an important role in the co-structure between parasitological and MHC data in bank voles. The two main associations detected in this study concerned the allele Cgl-*DQA*-04, which is rare but found in over 60% of CPXV-seropositive individuals, and the allele Cgl-*DQA*-12, which was never found among the 15 CPXV-seropositive individuals. This might suggest that this allele is associated with resistance to CPXV.

Orthopoxviruses, along with a number of other viruses, share mechanisms capable of modulating the host immune response. Many orthopoxviruses possess a protein that is capable of down-regulating MHC class I, which is responsible for presenting all intracellular antigens (Dasgupta et al., 2007). Biochemical and functional analyses have also revealed that poxviruses infection directly interfered with ligand binding to class II molecules (Ping et al., 2005). Therefore, down-regulation of MHC class I trafficking and disruption of MHC class II-mediated antigen presentation may be some of the multiple strategies orthopoxviruses have evolved to escape host immune surveillance. This modulation of the host defence mechanisms seems to rely on homologues of immune molecules encoded by orthopoxviruses (Johnston and McFadden, 2003). Combining immunogenetics of genes involved in host immunity and virus evasion could help us understanding these associations.

5. Conclusions

Our study provides evidence for the role of immunogenetics in viral infections. Several alleles are suggested for either susceptibility or resistance to these infections. However, these results have to be taken cautiously as they concern few seropositive individuals. Moreover, as demonstrated for passerines and malaria parasites, links between host immunogenetics and resistance can result from local adaptation processes (Bonneaud et al., 2006). The associations detected might thus involve population-specific alleles and consequently, vary in space and/or time. Further investigations are needed to confirm the associations observed and to test the existence of such spatio-temporal fluctuations.

Acknowledgements

We thank Patrice Cadet for help in the statistical analyses, Audrey Scala and Florian Holon for help in field work and Stanislas Kalùz for mite identification. This research was partially funded by EU grant GOCE-2003-010284 EDEN and the paper is catalogued by the EDEN Steering Committee as EDEN 044 (http://www.eden-fp6project.net/).

References

Axtner, J., Sommer, S., 2007. Gene duplication, allelic diversity, selection processes and adaptive value of MHC class II *DRB* genes of the bank vole, *Clethrionomys glareolus*. Immunogenetics 59, 417–426.

Blake, J.A., Richardson, J.E., Bult, C.J., Kadin, J.A., Eppig, J.T., Mouse Genome Database Group, 2003. MGD: the mouse genome database. Nucleic Acids Res. 31, 193–195.

Bonneaud, C., Perez-Tris, J., Federici, P., Chastel, O., Sorci, G., 2006. Major histocompatibility alleles associated with local resistance to malaria in a passerine. Evolution 60, 383–389.

Bryja, J., Galan, M., Charbonnel, N., Cosson, J.F., 2005. Analysis of major histocompatibility complex class II gene in water voles using capillary electrophoresis-single stranded conformation polymorphism. Mol. Ecol. Notes 5, 173–176.

Bryja, J., Galan, M., Charbonnel, N., Cosson, J.F., 2006. Duplication, balancing selection and trans-species evolution explain the high levels of polymorphism of the *DQA* MHC class II gene in voles (Arvicolinae). Immunogenetics 58, 191–202.

Capek, I., 2006. Définition de priorités et actions réalisées dans le domaine des zoonoses non alimentaires, 2000–2005. Bull. Epid. Hebdo. 27–28, 196–199.

Chantrey, J., Meyer, H., Baxby, D., Begon, M., Bown, K., Feore, S.M., Jones, T., Montgomery, W.I., Bennett, M., 1999. Cowpox: reservoir hosts and geographic range. Epidemiol. Infect. 122, 455–460.

Cooke, G.S., Hill, A.V.S., 2001. Genetics of susceptibility to human infectious disease. Nat. Rev. Genet. 2, 967–977.

Dasgupta, A., Hammarlund, E., Slifka, M.K., Frueh, K., 2007. Cowpox virus evades CTL recognition and inhibits the intracellular transport of MHC class I molecules. J. Immunol. 178, 1654–1661.

Doherty, P.C., Zinkernagel, R.M., 1975. A biological role for the major histocompatibility antigens. Lancet 1, 1406–1409.

Dray, S., Chessel, D., Thioulouse, J., 2003. Procrustean co-inertia analysis for the linking of multivariate datasets. Ecoscience 10, 110–119.

Frances, S.P., Watcharapichat, P., Phulsuksombati, D., Tanskul, P., 2000. Transmission of *Orientia tsutsugamushi*, the aetiological agent for scrub typhus, to co-feeding mites. Parasitology 120, 601–607.

Geraghty, D.E., 2002. Genetic diversity and genomics of the immune response. Immunol. Rev. 190, 5.

Günther, E., Walter, L., 2001. The major histocompatibility complex of the rat (*Rattus norvegicus*). Immunogenetics 53, 520–542.

Haldane, S., 1956. The estimation and significance of the logarithm of a ratio of frequencies. Ann. Hum. Genet. 20, 309–311.

Hall, T., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Hill, A.V.S., 2001. The genomics and genetics of human infectious disease susceptibility. Annu. Rev. Genomics Hum. Genet. 2, 373–400.

Houck, M.A., Qin, H., Roberts, H.R., 2001. Hantavirus transmission: potential role of ectoparasites. Vector-Borne Zoonotic Dis. 1, 75–79.

Hughes, A.L., Nei, M., 1990. Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. Mol. Biol. Evol. 7, 191–514.

Hurt, P., Walter, L., Sudbrak, R., Klages, S., Müller, I., Shiina, T., Inoko, H., Lehrach, H., Günther, E., Reinhardt, R., Himmelbauer, H., 2004. The genomic sequence and comparative analysis of the rat major histocompatibility complex. Gen. Res. 14, 631–639.

Johnston, J.B., McFadden, G., 2003. Poxvirus immunomodulatory strategies: current perspectives. J. Virol. 77, 6093–6100.

Kallio-Kokko, H., Laakkonen, J., Rizzoli, A., Tagliapietra, V., Cattadori, I., Perkins, S.E., Hudson, P.J., Cristofolini, A., Versini, W., Vapalahti, O., Vaheri, A., Henttonen, H., 2006. Hantavirus and arenavirus antibody prevalence in rodents and humans in Trentino, Northern Italy. Epidemiol. Infect. 134, 830–836.

Kallio-Kokko, H., Uzcategui, N., Vapalahti, O., Vaheri, A., 2005. Viral zoonoses in Europe. FEMS Microbiol. Rev. 29, 1051–1077.

Kallio, E.R., Klingström, J., Gustafsson, E., Manni, T., Vaheri, A., Henttonen, H., Vapalahti, O., Lundkvist, A., 2006a. Prolonged survival of Puumala hanatvirus outside the host: evidence for indirect transmission via the environment. J. Gen. Virol. 87, 2127–2134.

Kallio, E.R., Poikonen, A., Vakeri, A., Vapalahti, O., Henttonen, H., Koskela, E., Mappes, T., 2006b. Maternal antibodies postpone hantavirus infection and enhance individual breeding success. Proc. Roy. Soc. B 273, 2771–2776.

Kampen, H., Scholer, A., Metzen, M., Oehme, R., Hartelt, K., Kimmig, P., Maier, W.A., 2004. *Neotrombicula autumnalis* (Acari, Trombiculidae) as a vector for *Borrelia burgdorferi* sensu lato? Exp. Appl. Acarol. 33, 93–102.

Kennedy, L.J., Ryvar, R., Gaskell, R.M., Addie, D.D., Willoughby, K., Carter, S.D., Thomson, W., Ollier, W.E.R., Radford, A.D., 2002. Sequence analysis of MHC *DRB* alleles in domestic cats from the United Kingdom. Immunogenetics 54, 348–352.

Klein, J., Bontrop, R.E., Dawkins, R.L., Erlich, H.A., Gyllensten, U.B., Heise, E.R., Jones, P.P., Parham, P., Wakeland, E.K., Watkins, D.I., 1990. Nomenclature for the major histocompatibility complex of different species: a proposal. Immunogenetics 31, 217–219.

Klein, S.L., Cernetich, A., Hilmer, S., Hoffman, E.P., Scott, A.L., Glass, G.E., 2004a. Differential expression of immunoregulatory genes in male and female Norway rats following infection with *Seoul virus*. J. Med. Virol. 74, 180–190.

Klein, S.L., Zink, M.C., Glass, G.E., 2004b. Seoul virus infection increases aggressive behaviour in male Norway rats. Anim. Behav. 67, 421–429.

Klingström, J., Maljkovic, I., Zuber, B., Rollman, E., Kjerrstrom, A., Lundkvist, A., 2004. Vaccination of C57/BL6 mice with Dobrava hantavirus nucleocapsid protein in Freund's adjuvant induced partial protection against challenge. Vaccine 22, 4029–4034.

Laakkonen, J., Kallio-Kokko, H., Oktem, M.A., Blasdell, K., Plyusnina, A., Niernimaa, J., Karatas, A., Plyusnin, A., Vaheri, A., Henttonen, H., 2006. Serological survey for viral pathogens in Turkish rodents. J. Wildl. Dis. 42, 672–676.

Lundkvist, A., Fatouros, A., Niklasson, B., 1991. Antigenic variation of European haemorrhagic fever with renal syndrome virus strains characterized using bank vole monoclonal antibodies. J. Gen. Virol. 72, 2097–2103.

Mäkelä, S., Mustonen, J., Ala-Houhala, I., Hurme, M., Partanen, J., Vapalahti, O., Vaheri, A., Pasternack, A., 2002. Human leukocyte antigen-B8-DR3 is a more important risk factor for severe puumala hantavirus infection than the tumor necrosis factor-alpha(-308) G/A polymorphism. J. Infect. Dis. 843–846.

Mills, J.N., Childs, J., Ksiazek, T.G., Peters, C.J., Velleca, W.M., 1995. Methods for Trapping and Sampling Small Mammals for Virologic Testing. Centers for Disease Control and Prevention, Atlanta.

Mills, J.N., Childs, J.E., 1998. Ecologic studies of rodent reservoirs: their relevance for human health. Emerg. Infect. Dis. 4, 529–537.

Mustonen, J., Huttunen, N.P., Partanen, J., Baer, M., Paakkala, A., Vapalahti, O., Uhari, M., 2004. Human leukocyte antigens B8-DRB1*03 in pediatric patients with *Nephropathia epidemica* caused by puumala hantavirus. Pediatr. Infect. Dis. J. 23, 959–961.

Mustonen, J., Vaheri, A., Clement, J., 1996. Third international conference on haemorrhagic fever with renal syndrome (HFRS) and hantaviruses. Nephrol. Dialysis Transpl. 11, 730–733.

Mustonen, J., Vapalahti, I., Henttonen, H., Pasternack, A., Vaheri, A., 1998. Epidemiology of hantavirus infections in Europe. Nephrol. Dialysis Transpl. 13, 2729–2731.

Ping, L., Wang, N., Zhou, D., Yee, C.S.K., Chang, C.-H., Brutkiewicz, R.R., Blum, J.S., 2005. Disruption of MHC class II-restricted antigen presentation by vaccinia virus. J. Immunol. 175, 6481–6488.

Plyusnin, A., Horling, J., Kanerva, M., Mustonen, J., Cheng, Y., Partanen, J., Vapalahti, O., Kukkonen, S.K.J., Niemimaa, J., Henttonen, H., Niklasson, B., Lundkvist, A., Vaheri, A., 1997. Puumala hantavirus genome in patients with *Nephropathia epidemica*: correlation of PCR positivity with HLA haplotype and link to viral sequences in local rodents. J. Clin. Microbiol. 35, 1090–1096.

Plyusnina A., Deter J., Charbonnel N., Cosson J.-F., and Plyusnin A., in press. Puumala and Tula hantaviruses in France. Virus Research.

Rozsa, L., Reiczigel, J., Majoros, G., 2000. Quantifying parasites in samples of hosts. J. Parasitol. 86, 228–232.

Song, G., 1999. Epidemiological progresses of hemorrhagic fever with renal syndrome in China. Chin. Med. J. 112, 472–477.

Takahata, N., Nei, M., 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics 124, 967–978.

Taylor, L.H., Latham, S.M., Woolhouse, M.E.J., 2001. Risk factors for human disease emergence. Philos. Trans. Roy. Soc. Lond. B: Biol. Sci. 356, 983–989.

Terajima, M., Vapalahti, O., Van Epps, H.L., Vaheri, A., Ennis, F.A., 2004. Immune responses to Puumala virus infection and the pathogenesis of *Nephropathia epidemica*. Microbes Infect. 6, 238–245.

Thioulouse, J., Chessel, D., Doledec, S., Olivier, J.M., 1997. ADE-4: a multivariate analysis and graphical display software. Stat. Comput. 7, 75–83.

Thioulouse, J., Simier, M., Chessel, D., 2004. Simultaneous analysis of a sequence of paired ecological tables. Ecology 85, 272–283.

Vapalahti, O., Mustonen, J., Lundkvist, A., Henttonen, H., Plyusnin, A., Vaheri, A., 2003. Hantavirus infections in Europe. Lancet Infect. Dis. 3, 753–754.

Wegner, K.M., Reusch, T.B.H., Kalbe, M., 2003. Multiple parasites are driving major histocompatibility complex polymorphism in the wild. J. Evol. Biol. 224–232.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 459–466
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Differential strain-specific diagnosis of the heartwater agent: *Ehrlichia ruminantium*

Nathalie Vachiery [a], Gael Maganga [a], Thierry Lefrançois [a], Yane Kandassamy [a], Frédéric Stachurski [a], Hassane Adakal [b], Conchita Ferraz [c], Anne Morgat [d,e], Albert Bensaid [a,1], Eric Coissac [d], Frédéric Boyer [d], Jacques Demaille [c], Alain Viari [d], Dominique Martinez [a], Roger Frutos [a,*]

[a] *CIRAD TA A-15/G, Campus International de Baillarguet, 34398, Montpellier Cedex 05, France*
[b] *Cirdes, 01, BP. 454, Bobo-Dioulasso 01, Burkina-Faso*
[c] *Centre de Séquençage Génomique, IGH-CNRS-UPR 1142, 141 rue de la Cardonille, 34396 Montpellier Cedex 05, France*
[d] *Inria Rhône-Alpes—Projet HELIX, 655, Av. de l'Europe, 38330 Montbonnot-Saint Martin, France*
[e] *Swiss Institute of Bioinformatics, Swiss-Prot Group, 1 rue Michel, Servet, CH-1211 Geneva 4, Switzerland*

Received 2 May 2007; received in revised form 8 June 2007; accepted 11 June 2007
Available online 19 June 2007

Abstract

Ehrlichia ruminantium is the causative agent of heartwater, a major tick-borne disease of livestock in Africa introduced in the Caribbean and threatening to emerge and spread in the American mainland. Complete genome sequencing was done for two isolates of *E. ruminantium* of differing phenotype, isolates Gardel (*Erga*) from Guadeloupe Island and Welgevonden (*Erwe*) originating from South Africa and maintained in Guadeloupe. The type strain of *E. ruminantium* (*Erwo*), previously isolated and sequenced in South Africa; is identical to *Erwe* with respect to target genes. They make the *Erwe/Erwo* complex. Comparative analysis of the genomes shows the presence of 49 unique CDS and 28 truncated CDS differentiating *Erga* from *Erwe/Erwo*. Three regions of accumulated differences (RAD) acting as mutational hot spots were identified in *E. ruminantium*. Ten CDS, six unique CDS and four truncated CDS corresponding to major genomic changes (deletions or extensive mutations) were considered as targets for differential diagnosis on four isolates of *E. ruminantium*: *Erga*, *Erwe/Erwo*, Senegal and Umpala. Pairs of PCR primers were developed for each target gene. PCR analysis of the target genes generated strain-specific patterns on *Erga* and *Erwe/Erwo* as predicted by comparative genomics, but also for isolates Senegal and Umpala. The target genes identified by bacterial comparative genomics are shown to be highly efficient for strain-specific PCR diagnosis of *E. ruminantium* and further vaccine management tools.

Keywords: *Ehrlichia ruminantium*; Heartwater; Comparative genomics; Diagnosis

1. Introduction

Ehrlichia ruminantium is a tick-borne *Rickettsiale* responsible for cowdriosis or heartwater in ruminants all over Sub-Saharan Africa and in the Caribbean (Provost and Bezuidenhout, 1987). Heartwater is not only responsible for high economical losses in Africa (Mukhebi et al., 1999), but also threatens the American mainland because of the presence of tick species capable of transmitting the disease (Barré et al., 1987; Burridge et al., 2002). The current commercial vaccine is based on the high-risk method of injection of infected blood followed by treatment with antibiotics (Bezuidenhout, 1989). Attenuated and DNA vaccines were developed with good protection in challenged experiments (Collins et al., 2003a,b; Jongejan, 1991; Mahan et al., 2003; Martinez et al., 1994, 1996), they had a limited efficacy in the field (Collins et al., 2003a,b; Jongejan, 1991; Mahan et al., 2003; Martinez et al., 1994, 1996). However, a more recent vaccinal strain has not yet been tested in the wild (Zweygarth et al., 2005). *map1*

* Corresponding author at: CIRAD TA30/G, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France. Tel.: +33 4 67 59 39 62; fax: +33 4 67 59 39 60.
E-mail address: roger.frutos@cirad.fr (R. Frutos).
[1] Current address: Centre de Recerca en Sanitat Animal, Campus de Bellaterra, Edifici V, 08193 Bellaterra, Barcelona, Spain.

1567-1348/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2007.06.001

nested-PCR revealed the presence of various *E. ruminantium* genotypes simultaneously circulating in limited geographical areas, highlighting thus the need for extensive studies on *E. ruminantium* genetic diversity if vaccines have to be efficiently applied in the field (Martinez et al., 1990). Furthermore, comparative genomic analysis of *E. ruminantium* isolates showed the presence of active mechanisms of genome plasticity possibly involved in the limited field-efficacy of vaccines (Frutos et al., 2006). Strain-specific diagnosis is thus essential. However, serodiagnosis of heartwater has long been limited by a lack of specificity and sensitivity (Mondry et al., 1998). ELISA tests based on Map1-related membrane proteins improved the diagnostic but cross reactions with other *Ehrlichia* species still occur (Jongejan et al., 1993; Katz et al., 1997; Van Vliet et al., 1995). PCR-based method also proved efficient for detecting *E. ruminantium* in hosts and ticks (Peter et al., 1995) and to characterize isolates of *E. ruminantium* from related species (Allsopp et al., 1997; Bekker et al., 2002). However, routine strain-specific diagnosis is not yet achieved and additional diagnostic targets are still to be identified.

The genomes of two different isolates of *E. ruminantium*, Gardel and Welgevonden, were completely sequenced and annotated (Frutos et al., 2006). The genome of the type strain of *E. ruminantium*, another isolate of the Welgevonden genotype, was sequenced separately (Collins et al., 2005). A comparative genomic analysis of these three strains was conducted to identify additional diagnostic targets in which *Erwe* and *Erwo* were found identical with respect to the target genes reported here (Frutos et al., 2006). We report here the identification of six unique CDS and four truncated CDS allowing for strain-specific diagnosis of *E. ruminantium*. PCR amplification of these 10 target genes yielded differential patterns discriminating *Erga* from *Erwe* and *Erwo*. Furthermore, PCR screening on these target genes was also capable of discriminating strains of *E. ruminantium* for which no genomic data were available.

2. Material and methods

2.1. Bacterial isolates and growth conditions

The isolate Gardel of *E. ruminantium* (*Erga*) was isolated in Guadeloupe Island in 1982 from a goat injected with a homogenate of a female individual of *A. variegatum* collected on cows (Uilenberg et al., 1985). The isolate Welgevonden of *E. ruminantium* (*Erwe*) was isolated in South Africa in 1985 from mice injected with individually homogenized infected field-collected *A. hebraeum* ticks (Du Plessis, 1985). DNA from Welgevonden passage 14 was used in this study. The isolate Senegal was isolated in Senegal 1981 from sheep inoculated with infected bovine blood (Jongejan et al., 1988). The isolate Umpala was isolated in Mozambique from ticks fed on infected bovines (Jongejan et al., 1988). The isolates Lamba 479, Bekuy 255, Bankouma 421, Banan 112 and Banan 455 were all field-isolated in Burkina-Faso. *E. ruminantium* isolates were multiplied on bovine umbilical endothelial cell (BUEC) or bovine aorta endothelial cells (BAE) grown in Glasgow-MEM medium complemented with 10% inactivated fetal calf serum, tryptose–phosphate broth (2.9 g/l) and penicillin (100 IU/ml)/streptomycin (100 μg/ml) (Bezuidenhout et al., 1985) at 37 °C, 5% CO_2 with a weekly passage on fresh cells (Martinez et al., 1990). Bacterial DNA was extracted from infected endothelial cell culture at passage 18 for *Erga*, at passage 14 for *Erwe*, at passage 2 for isolate Umpala and at passage 4 for isolate Senegal.

2.2. DNA extraction and PCR reactions

Elementary bodies were purified from culture supernatant as previously described (Martinez et al., 1994), resuspended in 350 μl of PBS containing 0.36 μg/ml of DNAse to remove contaminating host cell DNA and incubated for 90 min at 37 °C prior to the addition of 25 mM EDTA (Martinez et al., 1990). Extraction of DNA from elementary bodies was done as previously described (Perez et al., 1997). Contamination with host DNA was checked by dot blot hybridization using a bovine DNA as positive control and probe. PCR was run in a final volume of 25 μl with 200 ng of total DNA, 2 μl of 2.5 mM dNTP mixture, 1.5 μl of 50 mM $MgCl_2$, 0.5 u of Taq DNA polymerase, 50 ng of the sense primer and 50 ng of the anti-sense primer. PCR conditions were set as follows: denaturation for 5 min at 94 °C, 35 cycles under conditions described in Table 1 and a final extension at 72 °C for 10 min. For each PCR, a negative control sample with bi-distilled water and PCR mix was used. The sequence of the primers used for PCR reactions are shown in Table 2. PCR products were separated on 1.5% agarose gels (Seakem) in 1× TAE buffer (40 mM Tris–HCl, 6% acetic acid, 1 mM EDTA, pH 8) for 1.5 h at 80 V.

2.3. Genome annotation

The full-length sequencing and annotation of the *Erga* and *Erwe* genomes have been previously described (Frutos et al., 2006). The sequences of the complete genomes of *Erga* and *Erwe* have been deposited to the EMBL databank under the accession numbers CR925677 and CR925678, respectively. Comparative genomic analysis of the *Erga* and *Erwe* genomes was conducted as previously described (Frutos et al., 2006) using the Genostar-Iogma package (http://www.genostar.com) (Durand et al., 2003). The type stain of *E. ruminantium* was

Table 1
PCR conditions for the different pairs of primers

Pairs of primer	Denaturation	Annealing	Extension
P-4350-A + P-4350-B	1 min at 94 °C	45 s at 37 °C	1 min at 72 °C
P-4990-A + P-4990-B	1 min at 94 °C	45 s at 37 °C	1 min at 72 °C
P-5600-A + P-5600-B	1 min at 94 °C	45 s at 49 °C	1 min at 72 °C
P-5610-A + P-5610-B	1 min at 94 °C	45 s at 47 °C	1 min at 72 °C
P-7600-A + P-7600-B	1 min at 94 °C	45 s at 47 °C	1 min at 72 °C
P-8340-A + P-8340-B	1 min at 94 °C	45 s at 37 °C	1 min at 72 °C
P-1350-A + P-1350-B	1 min at 94 °C	45 s at 47 °C	1 min at 72 °C
P-4510-A + P-4510-B	1 min at 94 °C	45 s at 44 °C	1 min at 72 °C
P-5750-A + P-5750-B	1 min at 94 °C	45 s at 44 °C	1 min at 72 °C
P-7420-A + P-7420-B	1 min at 94 °C	45 s at 47 °C	1 min at 72 °C

Table 2
PCR primers for detection of unique and truncated CDS in *Erga* and *Erwe*

Primers	Orientation	CDS status	Target gene	Primer sequence
P-4350-A	Sense	Unique	ERGA_CDS_04350	atgagtcacagttttattgag
P-4350-B	Antisense	Unique	ERGA_CDS_04350	cactcaaaatcacaagaagta
P-4990-A	Sense	Unique	ERGA_CDS_04990	atgtatttagtctatttagtagctg
P-4990-B	Antisense	Unique	ERGA_CDS_04990	ataacatctaattgaacaatatc
P-5600-A	Sense	Unique	ERGA_CDS_05600	atgaaaggatctttatctgc
P-5600-B	Antisense	Unique	ERGA_CDS_05600	ccttcttcttcttcattatg
P-5610-A	Sense	Unique	ERGA_CDS_05610	aagaattacatgatgcagc
P-5610-B	Antisense	Unique	ERGA_CDS_05610	tcttctcttgttatactctctg
P-7600-A	Sense	Unique	ERGA_CDS_07600	atggatttaaataaactaataaa
P-7600-B	Antisense	Unique	ERGA_CDS_07600	gcattttctctacctacga
P-8340-A	Sense	Unique	ERWE_CDS_08340	ccaaatatataaatgatctattc
P-8340-B	Antisense	Unique	ERWE_CDS_08340	gtacatagtatgtctttatataaaag
P-1350-A	Sense	Truncated	ERGA_CDS_01350, ERWE_CDS_01390	tccaccagagatgttatttgtaaag
P-1350-B	Antisense	Truncated	ERGA_CDS_01350, ERWE_CDS_01390	caacagaactttcagtattaaaagc
P-4510-A	Sense	Truncated	ERGA_CDS_04510, ERWE_CDS_04590, ERWE_CDS_04600	gttaagtgtgaaatgtattgtttag
P-4510-B	Antisense	Truncated	ERGA_CDS_04510, ERWE_CDS_04590, ERWE_CDS_04600	cactttctgttaattcaaaagtaga
P-5750-A	Sense	Truncated	ERGA_CDS_05750, ERWE_CDS_05840	gtaggccaaaaagtataggtaatag
P-5750-B	Antisense	Truncated	ERGA_CDS_05750, ERWE_CDS_05840	caacaaatacatcatcttcaagttg
P-7420-A	Sense	Truncated	ERWE_CDS_07420, ERGA_CDS_07340	agggttacttattgtagtcagagtg
P-7420-B	Antisense	Truncated	ERWE_CDS_07420, ERGA_CDS_07340	cctcttcgtatacaggattaccatt

described separately using a different annotation strategy (Collins et al., 2005). It is also a Welgevonden strain which is the parental strain of *Erwe* and was referred to as *Erwo* in a previously reported comparative genomic analysis as *Erwo* (Frutos et al., 2006). Comparative genomic analysis of *Erga*, *Erwe* and *Erwo* showed that the two Welgevonden strains, *Erwe* and *Erwo*, are identical with respect to the CDS considered here for strain-specific diagnosis (Frutos et al., 2006). Differences appear in the annotated genes between *Erwe* and *Erwo*, however this is only due to differing annotation strategy and parameter set-up (Frutos et al., 2006). When using the same parameters for both genomes the very same predicted CDS are obtained for the target CDS considered here. PCR amplification using the primers described in Table 2 provides the exact same products for both *Erwe* and *Erwo*. Either strain could be considered as representative of the Welgevonden genotype and are thus subsequently referred to as *Erwe/Erwo*.

3. Results

3.1. CDS differences between Erga and Erwe/Erwo

The previously reported comparative genomic analysis of *Erga*, *Erwe* and *Erwo* showed the presence of unique CDS in each isolate (Frutos et al., 2006). These unique CDS are sequences for which no predicted ortholog is found in the other genome. Twenty-two CDS are strictly specific to *Erga* whereas 27 are only found in *Erwe/Erwo* (Supplemental Table 3). Forty-three unique CDS in both genomes out of 49 do not display a predicted orthologous CDS in the other isolates because of the presence of a point mutation generating a stop codon and early gene termination. This results in CDS too short to be predicted according to the prediction parameters. The other six unique CDS correspond to major CDS alterations in the other genome. Five are found in *Erga* and one in *Erwe/Erwo* (Table 3, Supplemental Table 1). ERGA_CDS_04350 and ERGA_CDS_05610 correspond to extensively mutated CDS in *Erwe/Erwo*, i.e. numerous mutations throughout the CDS, resulting in the loss of a predicted open reading frame in the *Erwe/Erwo* genome. ERGA_CDS_05600 and ERGA_CDS_07600 correspond to fully deleted CDS in *Erwe/Erwo* whereas the *Erwe/Erwo* ortholog of ERGA_CDS_04990 is partially deleted only (Table 3, Supplemental Table 1). Similarly, ERWE_CDS_08340 (only the *Erwe* annotation is referred to here to facilitate the reading since the annotation strategy is the same as for *Erga*, however comparison with the *Erwo* annotation can be found in Frutos et al., 2006) corresponds to a partial deletion in the *Erga* orthologous CDS (Table 3, Supplemental Table 1). The comparative genomic analysis of the *Erga* and *Erwe/Erwo* genomes also showed the presence of both fragmented and partial CDS differentiating *Erga* from *Erwe/Erwo* (Frutos et al., 2006). Occurrence of a stop codon may also result in shorter but still predicted CDS depending upon the size of the remaining fragments. Truncated genes resulting in a single CDS are denominated partial CDS, whereas those resulting in two or more predicted CDS are described as fragmented CDS. Twelve such truncated CDS are specific to *Erga* whereas 17 are found only in *Erwe/Erwo* (Supplemental Table 1).

3.2. Regions of accumulated differences

Three regions in the genome of *E. ruminantium* can be identified as mutational hot spots or regions of accumulated differences (RAD) where more than three contiguous CDSs were affected over time by different kinds of mutations or genome rearrangements (Fig. 1, Supplemental Table 2). RAD 1 comprises 12 CDS spanning over 116 kb and bears 3 fragmented CDS, 1 partial CDS and a unique gene (Fig. 1). RAD 2, a smaller region of six conserved CDS distributed over 8 kb, contains two fragmented CDS, two partial CD and a

Table 3
Target CDS for strain-specific differential diagnostic of *E. ruminantium*

Target gene in *Erga*	Status	Size (size of deletion)	Function	Target gene in *Erwe*/*Erwo*	Status	Size (size of deletion)	Function
ERGA_CDS_04350	Unique gene	186 bp (NA)	Unkown	None	Extensive mutations	NA	NA
ERGA_CDS_04990	Unique gene	270 bp (N.A.)	Similarity to *rpoB*	None	Partial deletion	NA	NA
ERGA_CDS_05600	Unique gene	630 bp (NA)	Similarity to *cysC*	None	Full deletion	NA	NA
ERGA_CDS_05610	Unique gene	828 bp (NA)	Unknown	None	Extensive mutations	N.A	N.A
ERGA_CDS_07600	Unique gene	303 bp (NA)	Unknown	None	Full deletion	N.A	N.A
ERGA_CDS_01350	Full-length gene	3252 bp (NA)	Unknown	ERWE_CDS_01390/ Erum1430	Partial deletion[a]	2856 bp (396 bp)	Unknown
ERGA_CDS_04510	Full-length gene	3570 bp (NA)	Unknown	ERWE_CDS_04590/ Not predicted[b]	Partial deletion[c]	873 bp (585 bp)	Unknown
				ERWE_CDS_04600/ Erum4400	Partial deletion[c]	1470 bp	Unknown
ERGA_CDS_05750	Full-length gene	1836 bp (NA)	Unknown	ERWE_CDS_05840/ Erum5570	Partial deletion[a]	1659 bp (177 bp)	Unknown
ERGA_CDS_07340	Partial deletion[a]	3522 bp (600 bp)	Unknown	ERWE_CDS_07420/Erum	Full-length gene	4122 bp (NA)	Unknown
None	Partial deletion	NA	NA	ERWE_CDS_08340/ Not predicted[b]	Unique gene	225 bp (NA)	Unknown

N.A.: not applicable.
[a] The truncated CDS is a partial CDS.
[b] The CDS are not predicted in the *Erwo* published genome because of differing prediction parameters. Adjusting the parameters to those used for prediction of CDS in *Erwe* leads to identical results in *Erwe* and *Erwo*.
[c] The truncated CDS is a fragmented CDS.

unique gene (Fig. 1). RAD 3 is in the same size range as RAD 2 with 5.7 kb and bears more extensive genome rearrangements than the other two RAD with only one conserved orthologous gene out of six CDS (Fig. 1). RAD3 bears large unique CDS coming from deletions in the other isolate, a partial CDS and remaining pseudogenes (Fig. 1).

3.3. Target genes for differential diagnosis

Most of the unique, fragmented and partial CDS differentiating *Erga* from *Erwe*/*Erwo* are affected by a single or very few point mutations resulting in stops codons. They are not suited for differential diagnosis since most of the homologous sequence is still present in the other isolates although not predicted as a coding sequence. However, 10 genes bear large deletions or extensive mutations and thus represent targets of choice for differential diagnosis (Table 3). Two CDS are associated to the full-length deletion of the orthologous gene, six display partial deletions and two are characterized by extensive mutations (Table 3). ERGA_CDS_5610 whose ortholog is deleted in *Erwe*/*Erwo*, falls into this group of extensively mutated genes, due to remaining homology with the nucleotide sequence of *Erwe*/*Erwo*. *In silico* analysis of the specificity of PCR primers designed to amplify each of the 10 target genes showed that the 4 truncated genes and the 6 unique CDS could be differentially discriminated. PCR primers were

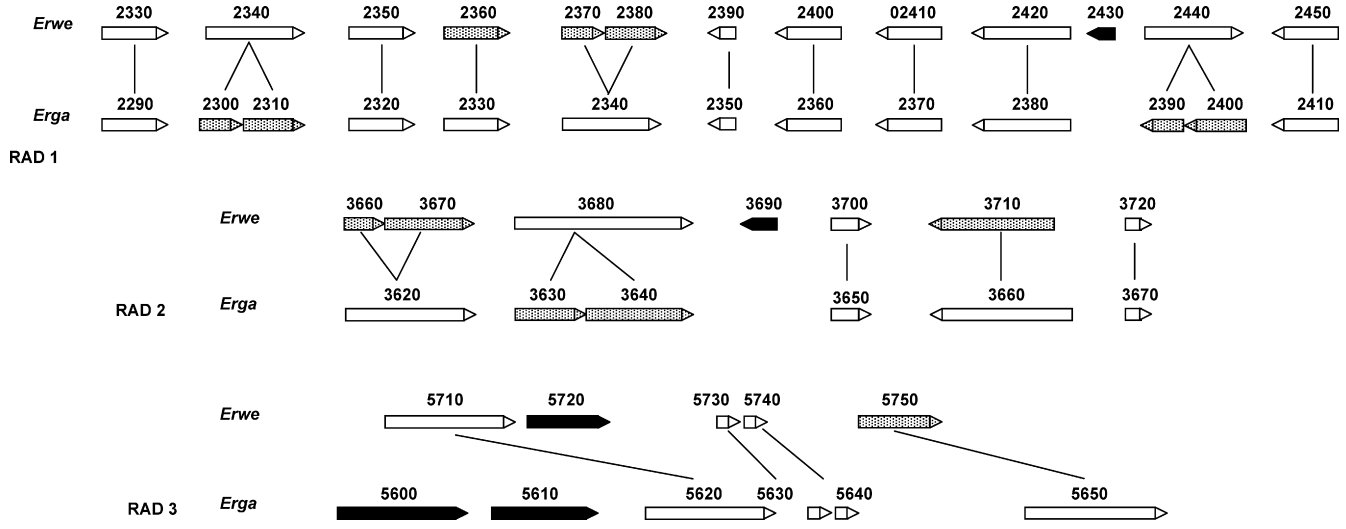


Fig. 1. Representation of the regions of accumulated differences (RAD). Mutants are shown in grey, unique CDS are shown in black and conserved orthologs are shown in white. Lines connect orthologs.

Table 4
Strain-specific differential PCR screening of *E. ruminantium*

Primer combinations	Strain									
	Erga	*Erwe*	*Erwo*[a]	Umpala	Senegal	Bankouma 421	Bekuy 255	Lamba 479	Banan 112	Banan 455
P-4350-A + P-4350-B	172	None	None	172	172	Multibands	172	172	172	Multibands
P-4990-A + P-4990-B	217	None	None	515	500 + 900	280 + 500 + 1200	500 + 900	500 + 900	500 + 900	500 + 1200
P-5600-A + P-5600-B	508	None	None	508	None	560	508 + 1900	508 + 1900	None	508
P-5610-A + P-5610-B	642	None	None	642	642	642[a]	642	642	642	642
P-7600-A + P-7600-B	238	None	None	238	None	238	238	238	238	238
P-8340-A + P-8340-	None	127	127	None	127	None	127	127	127	None
P-1350-A + P-1350-B	2791	2395	2395	2791 + 500	2395	2395	2395	2395	2395	None
P-4510-A + P-4510-B	552 + 1071 + 480	492	492	1200 + 500[b]	492	492	492	492	492	552 + 1071 + 480
P-5750-A + P-5750-B	1361	1178	1178	1361	1000	1000	1000	1000	1000	1178 + 1361
P-7420-A + P-7420-B	1095 + 300	1691	1691	820	820	820	820	820	820	1095

[a] Expected sizes were calculated *in silico* from the *Erwo* type-strain genome (Collins et al., 2005).
[b] Presence of multiple bands is observed.

thus developed to detect the targeted unique CDS based on presence–absence or size difference of the target sequences (Table 2).

3.4. PCR strain-specific differential diagnosis of *E. ruminantium*

Results of strain-specific differential PCR screening of *E. ruminantium* are summarized in Table 4. Primers developed to discriminate *Erga* and *Erwe*/*Erwo* upon differential detection of truncated CDS yielded the expected results. The primer pairs P-1350-A and -B, P-4510-A and -B, P-5750-A and -B and P-7420-A and -B yielded PCR products of the respective expected size of 2791 pb, 552 and 1071 pb, 1361 and 1095 pb on *Erga* and 2395 pb, 492 pb, 1178 pb and 1691 pb on *Erwe*/*Erwo*, respectively (Table 4, Fig. 2). An additional band of 480 bp is observed on *Erga* with the pair P-4510-A and -B. This additional band is most likely due to a single low specificity response occurring in *Erga*. Similar results were obtained using primers specific to unique CDS (Table 4, Fig. 3). All primers for unique CDS yielded the expected PCR products and differential response in both *Erga* and *Erwe*/*Erwo* (Table 4). The PCR primers were further tested firstly on two other strains obtained from other groups and conserved in the laboratory, i.e. Umpala and Senegal (Table 4, Figs. 2 and 3). Secondly, the analysis was extended to five more strains, i.e. Bankouma 421, Bekuy 255, Lamba 479, Banan 112 and Banan 455, which were field-isolated by two of the authors (F. Stachurski and H. Adakal) in

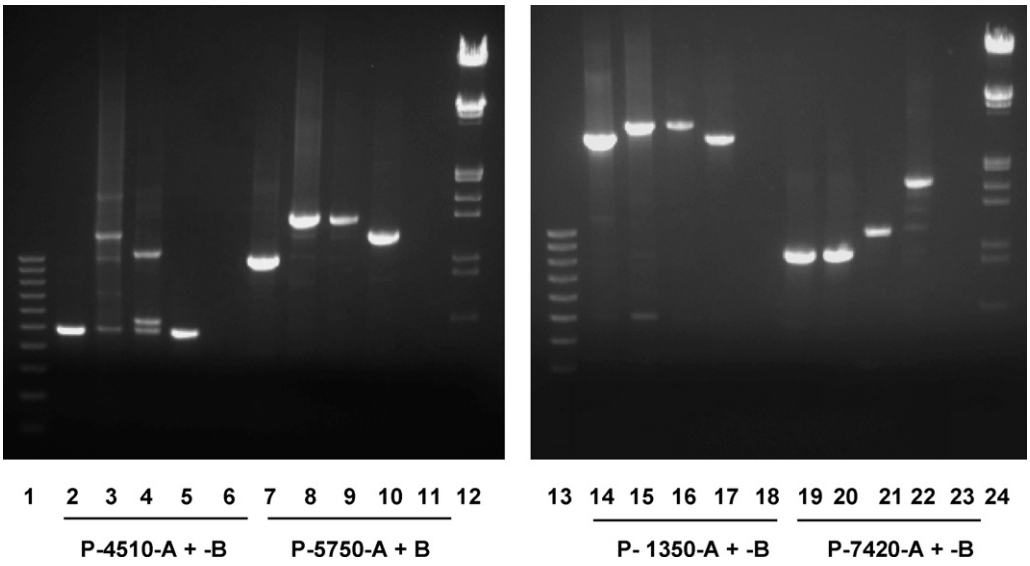


Fig. 2. Strain-specific differential diagnostic of *E. ruminantium* by PCR detection of truncated CDS. 1: Molecular weight marker (100-bp ladder); primers P-4510-A and -B on 2: senegal; 3: umpala; 4: *Erga*; 5: *Erwe*; 6: negative control sample; primers P-5750-A and -B on 7: senegal; 8: umpala; 9: *Erga*; 10: *Erwe*; 11: negative control sample; 12: molecular weight marker (λ EcoRI–HindIII); 13: molecular weight marker (100-bp ladder); primers P-1350-A and -B on 14: senegal; 15: umpala; 16: *Erga*; 17: *Erwe*/*Erwo*; 18: negative control sample; primers P-7420-A and -B on 19: senegal; 20: umpala; 21: *Erga*; 22: *Erwe*; 23: negative control sample; 24: molecular weight marker (λ EcoRI–HindIII).

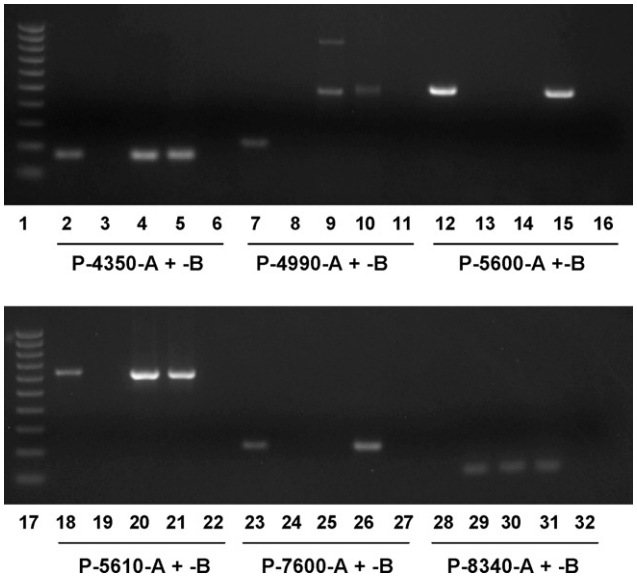


Fig. 3. Strain-specific differential diagnostic of *E. ruminantium* by PCR detection of unique CDS. 1: Molecular weight marker (100-bp ladder); primers P-5610-A + P-5610-B on 2: *Erga*; 3: *Erwe* 4: senegal; 5: umpala; 6: negative control sample; primers P-7600-A + -B on 7: *Erga*; 8: *Erwe*; 9: senegal; 10: umpala; 11: negative control sample; primers P-8340-A + -B on 12: *Erga*; 13: *Erwe*; 14: Senegal; 15: Umpala; 16: negative control sample; 17: molecular weight marker (100-bp ladder); primers P-4350-A + -B on 18: *Erga*; 19: *Erwe*; 20: senegal; 21: umpala; 22: negative control sample; primers P-4990-A + -B on 23: *Erga*; 24: *Erwe*; 25: senegal; 26: umpala; 27: negative control sample; primers P-5600-A + -B on 28: *Erga*; 29: *Erwe*; 30: senegal; 31: umpala; 32: negative control sample.

Burkina–Faso (Table 4). No genomic information was available for any of these seven additional strains. The use of primers originally designed to discriminate between *Erwe*/*Erwo* and *Erga* based on comparative genomic analysis yields strain-specific patterns allowing for discrimination between all nine isolates (Table 4).

4. Discussion

Efficient vaccines were developed under controlled virulent challenges, but their efficacy was proved limited in the field (Collins et al., 2003a,b; Jongejan, 1991; Mahan et al., 2003; Martinez et al., 1994, 1996; Zweygarth et al., 2005) owing to the simultaneous circulation of different populations of *E. ruminantium*, potentially displaying differing phenotypes in limited geographic areas while infected ticks bear composite populations of this bacterium (Martinez et al., 2005). Intracellular bacteria are considered to have reached stability with respect to genome plasticity following an initial massive gene loss as a consequence of adaptation to parasitism and intracellular life (Berg and Kurland, 2002). However, genome stability is a dynamic process in which mutations, duplications and pseudogenes development and decay are involved in the creation of genetic diversity while key genes are kept due to specific host-driven selective pressure (Berg and Kurland, 2002; Mira et al., 2002; Tamas et al., 2002). The potential and mechanisms for genome plasticity remains and might be expressed following environmental changes. The previously reported comparative genomic analysis of the Gardel and Welgevonden genotypes of *E. ruminantium* has shown the presence of such active mechanisms of genomic plasticity (Frutos et al., 2006). Plasticity is further exemplified by the differential evolution and recombination of the *map1-2* gene in two different subsets of the Gardel isolates grown under different conditions (Bekker et al., 2005) and by the development of an attenuated Welgevonden vaccine isolate through propagation in an unusual environment, i.e. a canine macrophage-monocyte cell line (Zweygarth et al., 2005). The presence of specific mechanisms of genome plasticity which could facilitate adaptation and evolution of differing populations (Frutos et al., 2006) may explain why vaccination with single isolates has until now yielded limited effect in the field. This apparent ability of evolution under host-driven selective pressure suggests that even efficient vaccines may become inefficient due to genotypic evolution of the bacteria. Furthermore, the current Map1 ELISA tests do not achieve strain-specific diagnostic (Jongejan et al., 1993). Molecular diagnostic with *map-1* nested PCR allowed for characterization of *E. ruminantium* strains, but on the *map-1* gene only. No multilocus molecular test is currently available. The sole identification of the presence of *E. ruminantium* is therefore clearly insufficient for effective control and prevention of heartwater and means to differentiate between several circulating strains are needed. Furthermore, since vaccine derivates of both the Welgevonden and Gardel strains have been developed (Vachiery et al., 2006; Zweygarth et al., 2005), diagnostic of circulating strains targeting these two main strains would be helpful for designing and monitoring of multivalent vaccines.

The comparative genomic analysis of the Gardel and Welgevonden genotypes has also shown the presence of strain-specific truncated CDS and unique CDS (Frutos et al., 2006). In accordance with the need for specific diagnosis of strains circulating in limited geographic areas (Martinez et al., 2005), the data reported here demonstrate the interest and potential of a selected set of altered genes for strain-specific diagnosis and monitoring. The ability of the PCR-based screening method targeting the selected altered genes to discriminate between *Erga* and *Erwe*/*Erwo* as predicted by comparative genomics demonstrates the potential of this approach for the identification of strain-specific targets and the development of diagnosis tools. Furthermore, the diagnosis primers described here are also capable of discriminating strains different from that of *Erga* and *Erwe*/*Erwo* for which no genomic sequence is available as shown by the detection of specific patterns in the Senegal, Umpala, Lamba 479, Bekuy 255, Bankouma 421, Banan 112 and Banan 455. This clearly reinforces their interest as diagnosis tools. Analysis of the genetic diversity of *E. ruminantium* to determine the structure of the population and assessment of the strain-specificity of the PCR primers reported here on a broader range of strains and isolates will provide a deeper understanding of their potential for strain-specific diagnosis and monitoring. This also indicates that the genetic diversity of strains observed in South Africa using a

different method (Allsopp and Allsopp, 2007) is also observed in West Africa.

The target CDS are associated to extensive mutations or deletions leading either in the affected strain to a complete loss or to a truncated CDS. These mutations correspond to the description of ''major genomic accidents'' which are considered to affect preferentially non-selected genes (Berg and Kurland, 2002). According to this hypothesis, these target CDS might be non-functional and dispensable remnants left over following adaptation to intracellular parasitism and genome shrinkage. As an obligate intracellular bacterium with a very low probability of transversal intraspecific genetic exchange, *E. ruminantium* corresponds to model of a population fragmented into numerous patches developing independently. In such a model, the introduction of novel sequence rarely occurs (Berg and Kurland, 2002) and major genomic accidents are likely to affect CDS no longer under selective pressure. If the target CDS considered here are non-functional remnants, the expected evolution of isolated populations suggests that they might be stably maintained in the strain (i.e. patch) as non-selected and non-detrimental remnants evolving through time by regular gene decay (Berg and Kurland, 2002). According to this theory, they are expected to reflect the particular genomic history of a given strain and thus to act as strain-specific markers. The results reported here would tend to support this hypothesis. Interestingly, this is also apparently in contradiction with the high genomic plasticity described in *E. ruminantium* both in terms of genomic size plasticity and mutations (Frutos et al., 2006), creation of pseudogenes (Collins et al., 2005) and tandem repeats (Collins et al., 2005; Frutos et al., 2006). This discrepancy is however most likely only apparent since these various studies did not address the same aspects. A high rate of substitution was observed between *Erga* and *Erwe*/*Erwo* but it was associated to a high level of conservation at the protein level, indicative of selection pressures for conservation of the proteins structure and function (Frutos et al., 2006, in press). Similarly, active movement of large tandem repeats was also observed, but was associated to genome size plasticity in non-coding regions (Frutos et al., 2006, in press). Collins et al. (2005) also reported the creation of pseudogenes through duplication and fusion, however this occurred in regions orthologs in all the sequenced genomes. In contrast, the CDS targeted here for diagnosis were specifically selected because they were associated to large deletions and extensive mutations.

This leads to another specific trait worth investigating further which is the presence of regions of accumulated differences or RAD in *E. ruminantium*. Phenotypic strain differences like pathogenesis or host range were often related to large insertion–deletion events or to large regions accumulating mutations (Brosh et al., 2001; Carroll et al., 2004; Fitzgerald and Musser, 2001; Fitzgerald et al., 2001). However, the regions identified in the genome of *E. ruminantium* are instead mutational hot spots where contiguous CDSs were affected by different kinds of mutations or genome rearrangements over time. The target CDS located in these regions (ERGA_CDS-5600 and ERGA-CDS-5610) are unique CDS and accordingly only a variation in presence–absence was observed depending upon the strain. Beyond their potential for strain-specific diagnosis, further investigations should be conducted to determine if a differential in gene expression exists in these regions. Combining the strain-specific unique features reported here with diagnostic methods providing multiplex analyses such as DNA arrays or PCR, more accurate and reliable means of discrimination could thus be developed not only to implement relevant survey, risk assessment and epidemiological analysis but also to ensure monitoring and use of appropriate vaccines.

Acknowledgements

The Welgevonden isolate was kindly provided by Dr. Durr Bezuidenhout (Onderstepoort Veterinary Institute, South Africa), in 1988, as an exchange for the Gardel isolate from Cirad. The Umpala and Senegal isolates were kindly provided by Dr. Frans Jongejan (Faculty of Veterinary Medicine, University of Utrecht, The Netherlands). This work was supported by Cirad-Cnrs grant #751745/00 and by the EC grant Inco-Dev #FP6-003713: ''Epigenevac''.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.meegid.2007.06.001.

References

Allsopp, M.T.E.P., Allsopp, B.A., 2007. Extensive genetic recombination occurs in the field between different genotypes of *Ehrlichia ruminantium*. Vet. Microbiol. 124, 58–65.

Allsopp, M.T., Visser, E.S., Du Plessis, J.L., Vogel, S.W., Allsopp, B.A., 1997. Different organisms associated with heartwater as shown by analysis of 16S ribosomal RNA gene sequences. Vet. Parasitol. 71, 283–300.

Barré, N., Uilenberg, G., Morel, P.C., Camus, E., 1987. Danger of introducing heartwater onto the American mainland: potential role of indigenous and exotic *Amblyomma* ticks. Onderstepoort J. Vet. Res. 54, 406–416.

Bekker, C.P.J., De Vos, S., Taoufik, A., Sparagano, O.A.E., Jongejan, F., 2002. Simultaneous detection of *Anaplasma* and *Ehrlichia* species in ruminants and detection of *Ehrlichia ruminantium* in *Amblyomma variegatum* ticks by reverse line blot hybridization. Vet. Microbiol. 89, 223–238.

Bekker, C.P.J., Postigo, M., Taoufik, A., Bell-Sakyi, L., Ferraz, C., Martinez, D., Jongejan, F., 2005. Transcription analysis of the major antigenic protein 1 multigene family of three in vitro cultured *Ehrlichia ruminantium* isolates. J. Bacteriol. 187, 4782–4791.

Berg, O.G., Kurland, C.G., 2002. Evolution of microbial genomes: sequence acquisition and loss. Mol. Biol. Evol. 19, 2265–2276.

Bezuidenhout, J.D., 1989. Cowdria vaccines. In: Wright, I.G. (Ed.), Veterinary Protozoan and Hemoparasite Vaccines. CRC Press Inc., Boca Raton, pp. 31–42.

Bezuidenhout, J.D., Paterson, C.L., Barnard, B.J., 1985. In vitro cultivation of *Cowdria ruminantium*. Onderstepoort J. Vet. Res. 52, 113–120.

Brosh, R., Pym, A.S., Gordon, S.V., Cole, S.T., 2001. The evolution of mycobacterial pathogenicity: clues from comparative genomics. Trends Microbiol. 9, 452–458.

Burridge, M.J.J., Simmons, L.A., Peter, T.F., Mahan, S.M., 2002. Increasing risks of introduction of heartwater onto the American mainland associated with animal movements. Ann. N. Y. Acad. Sci. 969, 269–274.

Carroll, I.M., Khan, A.A., Ahmed, N., 2004. Revisiting the pestilence of *Helicobacter pylori*: insights into geographical genomics and pathogen evolution. Infect. Genet. Evol. 4, 81–90.

Collins, N.E., Pretorius, A., Van Kleef, M., Brayton, K.A., Allsopp, B.A., Zweygarth, E., Allsopp, B.A., 2003a. Development of improved attenuated and nucleic acid vaccines for heartwater. Dev. Biol. 114, 121–136.

Collins, N.E., Pretorius, A., Van Kleef, M., Brayton, K.A., Zweygarth, E., Allsopp, B.A., 2003b. Development of improved vaccines for heartwater. Ann. N. Y. Acad. Sci. 990, 474–484.

Collins, N.E., Liebenberg, J., De Villiers, E.P., Brayton, K.A., Louw, E., Pretorius, A., Faber, F.E., Van Heerden, H., Josemans, A., Van Kleef, M., Steyn, H.C., van Strijp, M.F., Zweygarth, E., Jongejan, J., Maillard, J.C., Berthier, D., Botha, M., Joubert, F., Corton, C.H., Thomson, N.R., Allsopp, T.M., Allsopp, B.A., 2005. The genome of the heartwater agent *Ehrlichia ruminantium* contains multiple tandem repeats of actively variable number. Proc. Natl. Acad. Sci. U.S.A. 102, 838–843.

Du Plessis, J.L., 1985. A method for determining the *Cowdria ruminantium* infection rate of *Amblyomma hebraeum*: effects in mice injected with tick homogenates. Onderstepoort J. Vet. Res. 52, 55–61.

Durand, P., Medigue, C., Morgat, A., Vandenbrouck, Y., Viari, A., Rechenmann, F., 2003. Integration of data and methods for genome analysis. Curr. Opin. Drug Discov. Dev. 6, 346–352.

Fitzgerald, J.R., Musser, J.M., 2001. Evolutionary genomics of pathogenic bacteria. Trends Microbiol. 9, 547–553.

Fitzgerald, R.J., Sturdevant, D.E., Mackie, S.M., Gill, S.R., Musser, J.M., 2001. Evolutionary genomics of *Staphylococcus aureus*: insight into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. Proc. Natl. Acad. Sci. U.S.A. 98, 8821–8826.

Frutos, R., Viari, A., Ferraz, C., Morgat, A., Eychenié, S., Kandassamy, Y., Chantal, I., Bensaid, A., Coissac, E., Vachiery, N., Demaille, J., Martinez, D., 2006. Comparative genomic analysis of three strains of *Ehrlichia ruminantium* reveals an active process of genome size plasticity. J. Bacteriol. 188, 2533–2542.

Frutos, R., Viari, A., Vachiery, N., Boyer, F., Martinez, D., in press. *Ehrlichia ruminantium*: *genomic and evolutionary features*. Trends Parasitol., in press.

Jongejan, F., 1991. Protective immunity to heartwater (*Cowdria ruminantium* infection) is acquired after vaccination with in vitro-attenuated *Rickettsiae*. Infec. Immun. 59, 729–731.

Jongejan, F., De Vries, N., Nieuwenhuijs, J., Wassink, L.A., Van Vliet, A.H.M., 1993. The immunodominant 32-kilodalton protein of *Cowdria ruminantium* is conserved within the genus *Ehrlichia*. Rev. Elev. Med. Vet. Trop. 46, 145–152.

Jongejan, F., Uilenberg, G., Franssen, F.F.J., Gueye, A., Nieuwenhuijs, J., 1988. Antigenic differences between stocks of *Cowdria ruminantium*. Res. Vet. Sci. 44, 186–189.

Katz, J.B., DeWald, R., Dawson, J.E., Camus, E., Martinez, D., Mondry, R., 1997. Development and evaluation of a recombinant antigen, monoclonal antibody-based competitive ELISA for heartwater serodiagnosis. J. Vet. Diagn. Invest. 9, 130–135.

Mahan, S.M., Barbet, A.F., Burridge, M.J., 2003. Development of improved vaccines for heartwater. Dev. Biol. 114, 137–145.

Martinez, D., Maillard, J.C., Coisne, S., Sheikboudou, C., Bensaid, A., 1994. Protection of goats against heartwater acquired by immunisation with inactivated elementary bodies of *Cowdria ruminantium*. Vet. Immunol. Immunopathol. 41, 153–163.

Martinez, D., Perez, J.M., Sheikboudou, C., Debus, A., Bensaid, A., 1996. Comparative efficacy of Freund's and Montanide ISA50 adjuvants for the immunisation of goats against heartwater with inactivated *Cowdria ruminantium*. Vet. Parasitol. 67, 175–184.

Martinez, D., Swinkels, J., Camus, E., Jongejan, J., 1990. Comparison between 3 antigens for the serodiagnosis of heartwater disease by indirect immunofluorescence. Rev. Elev. Med. Vet. Pays Trop. 43, 159–166.

Martinez, D., Vachiery, N., Stachurski, F., Kandassamy, Y., Raliniaina, M., Aprelon, R., Gueye, A., 2005. Nested PCR for the detection and genotyping of *Ehrlichia ruminantium*. Use in genetic diversity analysis. Ann. N. Y. Acad. Sci. 1026, 106–113.

Mira, A., Klasson, L., Andersson, S.G.E., 2002. Microbial genome evolution: sources of variability. Curr. Opin. Microbiol. 5, 506–512.

Mondry, R., Martinez, D., Camus, E., Liebisch, A., Katz, J.B., Dewald, R., Van Vliet, A.H., Jongejan, J., 1998. Validation and comparison of three enzyme-linked immunosorbent assays for the detection of antibodies to *Cowdria ruminantium* infection. Ann. N. Y. Acad. Sci. 849, 262–272.

Mukhebi, A.W., Chamboko, T., O'Callaghan, C.J., Peter, T.F., Kruska, R.L., Medley, G.F., Mahan, S.M., Perry, B.D., 1999. An assessment of the economic impact of heartwater (*Cowdria ruminantium* infection) and its control in Zimbabwe. Prev. Vet. Med. 39, 173–189.

Perez, J.M., Martinez, D., Debus, A., Sheikboudou, C., Bensaid, A., 1997. Detection of genomic polymorphisms among isolates of the intracellular bacterium *Cowdria ruminantium* by random amplified polymorphic DNA and Southern blotting. FEMS Microbiol. Lett. 154, 73–79.

Peter, T.F., Deem, S.L., Barbet, A.F., Norval, R.A., Simbi, B.H., Kelly, P.J., Mahan, S.M., 1995. Development and evaluation of PCR assay for detection of low levels of *Cowdria ruminantium* infection in *Amblyomma* ticks not detected by DNA probe. J. Clin. Microbiol. 33, 166–172.

Provost, A., Bezuidenhout, J.D., 1987. The historical background and global importance of heartwater. Onderstepoort J. Vet. Res. 54, 165–169.

Tamas, I., Klasson, L., Canback, B., Naslund, A.K., Eriksson, A.S., Wernegreen, J.J., Sandstrom, J.P., Moran, N.A., Andersson, S.G.E., 2002. 50 million years of genomic stasis in endosymbiotic bacteria. Science 296, 2376–2379.

Uilenberg, G., Camus, E., Barré, N., 1985. Some observations on a stock of *Cowdria ruminantium* from Guadeloupe (French West Indies). Rev. Elev. Med. Vet. Pays Trop. 34, 34–42.

Vachiery, N., Lefrancois, T., Esteves, I., Molia, S., Sheikboudou, C., Kandassamy, Y., Martinez, D., 2006. Optimisation of the inactivated vaccine dose against heartwater and in vitro quantification of *Ehrlichia ruminantium* challenge material. Vaccine 24, 4747–4756.

Van Vliet, A.H., Van der Zeijst, B.A., Camus, E., Mahan, S.M., Martinez, D., Jongejan, F., 1995. Use of a specific immunogenic region on the *Cowdria ruminantium* MAP1 protein in a serological assay. J. Clin. Microbiol. 33, 2405–2410.

Zweygarth, E., Josemans, A.I., Van Strijp, M.F., Lopez-Rebollar, L., Van Kleef, M., Allsopp, B.A., 2005. An attenuated *Ehrlichia ruminantium* (Welgevonden stock) vaccine protects small ruminants against virulent heartwater challenge. Vaccine 23, 1695–1702.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 467–473
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Promiscuous T cell epitope prediction of *Candida albicans* secretory aspartyl protienase family of proteins

Songsak Tongchusak [a], Vladimir Brusic [b,c], Sansanee C. Chaiyaroj [a,d,*]

[a] *Department of Microbiology, Faculty of Science, Mahidol University, Bangkok, Thailand*
[b] *Cancer Vaccine Center, Dana-Farber Cancer Institute, Boston, MA, USA*
[c] *School of Land, Crop and Food Sciences, University of Queensland, Brisbane, QLD, Australia*
[d] *Laboratory of Pharmacology, Chulabhorn Research Institute, Bangkok, Thailand*

Received 16 April 2007; received in revised form 18 September 2007; accepted 19 September 2007
Available online 25 September 2007

Abstract

Candida albicans is one of the most important opportunistic dimorphic fungi responsible for hospital acquired fungal infection in humans. *Candida* infection rarely occurs in healthy individuals but it is frequently associated with patients who suffer from acquired immunodeficiency syndromes. To date, there is no effective vaccine against this fungal infection. Herein we demonstrated the use of immunomics to characterize promiscuous T cell epitope of *C. albicans* virulence factors by utilizing CandiVF, a *C. albicans* database previously constructed to be equipped with protein sequence analysis tool, three dimensional structure visualization software, sequence variable analysis program and Hotspot Hunter epitope prediction tool. Secretory aspartyl proteinase (Sap) family was chosen as a model to validate the Hotspot Hunter prediction. Analysis of Saps1–10 protein entries from CandiVF database revealed that a consensus T cell epitope was located at the C-terminal region of Saps1–10. The result of the *in silico* prediction was subsequently validated by conventional immunological methods. By using overlapping peptides span the predicted consensus T cell epitopes of Saps1–10 as stimulators, it was demonstrated that peptides S6 and S7 could stimulate PBMC proliferation in 9 of 12 blood donors. Interestingly, S2, the predicted T cell epitope of Sap2, was able to induce proliferation of all donors' PBMC. ELISpot assay for the detection of gamma-interferon producing clones confirmed that the peptide S2 actually stimulated T cell proliferation. The results suggest that S2 might be a potential candidate for vaccine development against *C. albicans* infection or to be utilized as an adjuvant to stimulate the pre-existing $CD4^+$ T cell in other vaccine development.

Keywords: *Candida albicans*; Immunomics; Secretory aspartyl proteinase; T cell epitope; Antigenic prediction; Interferon ELISpot assay; PBMC proliferation

1. Introduction

Candidiasis is a disease caused by *Candida* spp. It is an infection of skin, oral cavity, gastrointestinal tract, vagina and vascular system. Generally, it rarely occurs in healthy individual but it is frequently associated with patients, who receive immunosuppressive drug or long term catheterization, and patients who suffer from AIDS (Senet, 1997). High incidence of oral and gastrointestinal candidiasis has been reported among HIV infected patients (Johnson, 2000). Various molecules of *C. albicans* have been reported to be involved in candidiasis. Upon colonization in host tissue, these fungal molecules are induced and expressed in response to host environmental stimuli such as pH, nutrient, serum factor and immune response. Such molecules assist the fungus on their survival and host immune evasion. We earlier developed a molecular database of *C. albicans* virulence factors for immunological validation of *Candida* antigen called CandiVF (Tongchusak et al., 2005). The database contained 153 virulence proteins (until January 2005). These includes secretory proteins, cell wall associated proteins, membrane proteins and allergic proteins (Fig. 1). Of these, only six proteins were characterized as allergens (Table 1). The database can be accessible via WWW at URL http://antigen.i2r.a-star.edu.sg/Templar/DB/CandiVF/.

Among these virulence molecules, secretory aspartyl proteinases (Saps) are the most important virulence factor of

* Corresponding author at: Department of Microbiology, Faculty of Science, Mahidol University, Bangkok 10400, Thailand. Tel.: +66 2201 5689; fax: +66 2644 5411.
E-mail address: scscy@mahidol.ac.th (S.C. Chaiyaroj).

1567-1348/$ – see front matter
doi:10.1016/j.meegid.2007.09.006

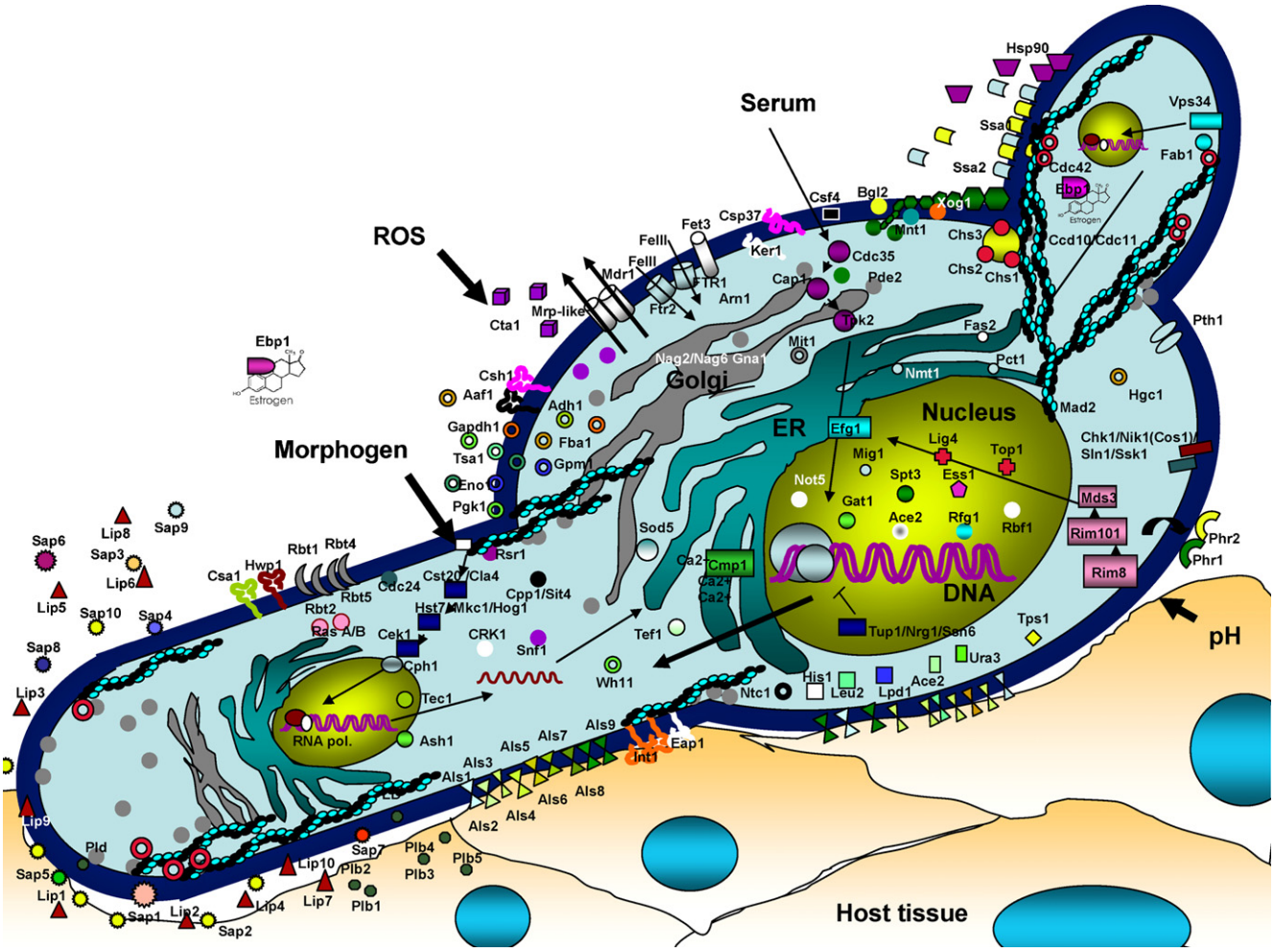


Fig. 1. Schematic representation of *C. albicans* virulence proteins expressed during host infection. Full names of virulence factors can be retrieved from CandiVF database (URL http://antigen.i2r.a-star.edu.sg/Templar/DB/CandiVF/).

C. albicans infections. *C. albicans* produces ten members of secretory aspartyl proteinases (Saps) which are associated with fungal virulence. These proteinases are encoded by *SAP1* to *SAP10* genes. The genes encode pre-proenzymes that are cleaved by Kex2 enzyme in Golgi complex before subjected to *N*-glycosylation. Upon catalytic cleavage, the enzymes become functionally mature enzymes (Calderone and Ronald, 2002). Saps have two conserved aspartic acid residues that serve as catalytic sites for its biological function. Additionally, several Saps possess conserved cysteine residues involving in disulfide bond formation (Naglik et al., 2004).

Table 1
List of *Candida albicans* allergic proteins

Allergens	References
Alcohol dehydrogenase (Adh1)	Horner et al. (1995)
Enolase (Eno1)	Horner et al. (1995)
Aldolase (Fba1)	Horner et al. (1995)
Glyceraldehyde-3-phosphate dehydrogenase (Gapdh)	Navarro-Garcia et al. (2001)
Phosphoglycerate kinase (Pgk1)	Horner et al. (1995)
Secretory aspartyl proteinase 2 (Sap2)	Akiyama et al. (1996)

Amino acid sequence analysis reveals that Sap1–Sap3 are closely related in term of sequence identity. These three enzymes have shown to be important in fungal adherence whereas other three closely related enzymes, namely, Sap4–Sap6, are necessary for deep-seated candidiasis (Naglik et al., 2004). Experiments showed that not all Saps were secreted to the outside of the cell; for example, Sap9 and Sap10 were localized at the fungal cell membrane via GPI anchor (Naglik et al., 2004). Expression and secretion of Saps are pH dependent and tissue specific. Saps1–3 are produced at pH 2–5 during superficial infection whereas Saps4–6 are secreted at pH 3–7 during the course of invasive candidiasis (Kretschmar et al., 1999; Schaller et al., 1999; Naglik et al., 2003). Expression of Sap7 and Sap8 remains unproven (Kretschmar et al., 1999; Schaller et al., 1999).

Sap2 is a major protein that induces antibody response during *Candida* infection in human. It has been experimentally reported that Sap2 targets on host extracellular matrix such as mucin, keratin, collagen, laminin and fibronectin (Naglik et al., 2004). Moreover, this enzyme effectively catalyzes the immune components such as saliva lactoferrin, macrophage enzyme, immunoglobulin and a proteinase inhibitor, α-macroglobulin, allowing the fungus to escape from the first line of host defense.

Sap2 also converts blood clotting factor X to Xa, inactivates cystatin A, and cleaves the vasoconstrictive peptide endothelin 1, leading to vascular malfunction (Tsushima et al., 1994; Tsushima and Mine, 1995; Hube and Naglik, 2001).

In this study we demonstrated the use of immunoinformatics approach for the identification of a consensus T cell epitope of secretory aspartyl protinases that promiscuously stimulated T cell proliferation via multiple HLA-DRB1 supertypes. Conventional immunological methods were subsequently performed to assess the ability of peptide to stimulate T cell proliferation and cytokine production. Identification of the novel immunogenic peptides of *C. albicans* Sap2 may assist future design and development of vaccine or cancer vaccine adjuvant.

2. Materials and methods

2.1. Prediction of promiscuous class II HLA-DRB1 binding peptides

Amino acid sequences of *C. albicans* virulence factor were obtained from *Candida albicans* virulence factor database (Tongchusak et al., 2005). Promiscuous $CD4^+$ restricted T cell epitopes were predicted by Hotspot Hunter Program through the CandiVF website (http://antigen.i2r.a-star.edu.sg/Templar/DB-/CandiVF/). The program identified HLA-DRB1 binding peptides by MULTIPRED computation system that can compute and select peptides that have high affinity to eight common HLA-DRB1 alleles (Zhang et al., 2005). The alleles selected for this analysis were: DRB1*0101; DRB1*0301; DRB1*0401; DRB1*0701; DRB1*0801; DRB1*1101; DRB1*1301; and DRB1*1501. In this study, 452 primary sequences including those of *SAP1–SAP10* were submitted to the Hotspot Hunter Program for prediction. The threshold of binding score was set at 75. Peptides with binding score of $\geq$75 would be predicted as binders whereas those with binding scores of lower than 75 would be predicted as non-binders.

2.2. Peptides synthesis

Short peptides of 20-mer in length, with an overlapping sequence of 10 amino acids, were prepared corresponding to the epitope selections at the C-terminal of Sap family of proteins (Fig. 2). Synthesis of overlapping peptides was performed through a custom made facility of Chiron Mimotope Company (Melbourne, Australia). The synthesis relied on solid phase peptide synthesis 9-fluorenylmethoxycarbonyl (fmoc) strategy. The peptides were delivered 90% pure as ascertained by HPLC, mass spectrometry and UV scan. Each synthetic peptide was reconstituted in concentrated dimethyl sulfoxide (DMSO) as a stock solution at 1 mg/ml (Sigma–Aldrich Chemical Company, St. Louis, MO, USA) and stored at −20 °C.

Table 2
Information of donors and their HLA-DRB1 allelic types

Donor	Gender	Age	HLA-Allele 1	HLA-Allele 2
D1	Male	28	DRB1*12	DRB1*13
D2	Female	26	DRB1*04	DRB1*15
D3	Male	27	DRB1*03	DRB1*08
D4	Female	29	DRB1*04	DRB1*15
D5	Female	35	DRB1*04	DRB1*0701
D6	Female	29	DRB1*03	DRB1*15
D7	Male	47	DRB1*03	DRB1*12
D8	Male	39	DRB1*12	DRB1*12
D9	Female	28	DRB1*04	DRB1*14
D10	Female	33	DRB1*04	DRB1*04
D11	Male	24	DRB1*04	DRB1*0701
D12	Male	42	DRB1*03	DRB1*15

2.3. Study subjects and blood collection

Fifty milliliters of peripheral blood were drawn from 12 healthy volunteers after obtaining informed consent. Table 2 provides detailed information related to the donors. The study was under the documentary approval of the Ethical Clearance Committee on Human Rights Related to Researches Involving Human Subjects of Faculty of Medicine, Ramathibodi Hospital, Mahidol University (Protocol Number ID 05-48-29).

2.4. Genomic DNA extraction of peripheral blood leukocytes and HLA typing

Extraction of genomic DNA from peripheral blood leukocytes was performed as described (Sambrook et al., 1989). Genomic DNA was precipitated and resuspended in distilled water and stored at −20 °C. DNA concentration was determined by UV spectrophotometry (OD_{260}/OD_{280}) and run through agarose gel for quality checking. HLA-DRB1 typing was performed by polymerase chain reaction with primer mixes utilizing allele sequence specific primers (PCR-SSP) (Olerup and Zetterquist, 1992). The typing was kindly undertaken by Dr. Chanvit Leelayuwat, Khon Kaen University, Thailand.

```
SAP1  299 KCQLLLGISDANILGDNFLRSAYLVYDLDDDKISLAQVKYTSASNIAALT 391
SAP2  306 KCQLLFDVNDANILGDNFLRSAYIVYDLDDNEISLAQVKYTSASSISALT 398
SAP3  306 QCQLLFGTSDYNILGDNFLRSAYIVYDLDDNEISLAQVKYTTASNIAALT 398
SAP4  325 KCEIRVRESEDNILGDNFMRSAYIVYDLDDRKISMAQVKYTSQSNIVAIN 417
SAP5  326 KCEVRIRESEDNILGDNFLRSAYVVYNLDDKKISMAPVKYTSESDIVAIN 418
SAP6  326 KCEIRVRESEDNILGDNFMRSAYIVYDLDDKKISMAQVKYTSESNIVAIN 418
SAP7  496 LCVFGILPGTHSILGDNFMRSVYAVFDLEDHVISIAQAAYNDNHAVVPIE 588
SAP8  321 SCYFGVSRDSATILGDNFLRRAYAVYDLDGNTISLAQVKYTTSSSISTL- 405
SAP9  402 GVMQQSSSSSYMLFGDNILRSAYIVYDLDDYEVSLAQVSYTNKESIEVIG 493
SAP10 297 IMSNSVIGGGGILFGDDILRQIYLVYDLQDMTISVAPVVYTEDEDIEEIL 386
```

Fig. 2. Amino acid sequence alignment of secretory aspartyl proteinases, Sap1–Sap10, showing consensus T cell epitope at the C-termini. Underlined sequences represent the predicted T cell epitopes.

2.5. Peripheral blood mononuclear cells (PBMC) isolation

Human PBMC were isolated by density gradient centrifugation. Thirty milliliters of peripheral blood were mixed with equal volume of 1× PBS. A total of 30 ml diluted blood was then carefully layered on the top of 15 ml Ficoll-hypaque (Isoprep™, Nycomed Pharma, Norway) in a non-pyrogenic conical tube (CORNING®, Corning, NY). After separation, PBMC layer was transferred into a new pyrogen free conical tube and washed three times with 1× PBS by centrifugation. Contaminated red blood cells were lysed osmotically by the addition of sterile distilled water. After washing, PBMC pellet was resuspended in complete RPMI-1640 (GIBCO/BRL, Grand Island, NY, USA). Cell numbers were determined using Bright-Line® Improved Newbauer Haemacytometer and the concentration of the cells was adjusted to 2×10^6 cells/ml.

2.6. PBMC proliferation assay

PBMC at 2×10^5 cells were seeded in an individual well of a 96-well plate. Sap synthetic peptides at the final concentration of 10 μg/ml were used to stimulate PBMC. PBMC stimulated with 5 μg/ml phytohemagglutinin (PHA) (Sigma–Aldrich) served as positive control while PBMC cultured in complete RPMI-1640 alone served as negative control. Cultures were incubated at 37 °C in 5% CO_2 incubator for 24 h. At the end of the incubation period, proliferation of PBMC was determined by enumeration of proliferative colonies under an inverted microscope (Olympus, CK2). Proliferative colonies were determined for five microscopic fields under 20× magnifying objective lens. Equations below were employed to determine proliferative index.

$$N = \left(\sum_{i=1}^{5} \mathrm{Xi}\right) Y$$

N is the number of proliferative colonies; Xi is the summation of proliferative colonies from five microscopic fields under 20× magnifying objective lens (colonies/mm^2); $Y = 32.2$ mm^2 (surface area of at the bottom of well).

Proliferative response was expressed as proliferative index (PI) by the following equation below;

$$\text{Proliferative Index (PI)} = \frac{\mathrm{Ns} - \mathrm{Nr}}{\mathrm{Nr}}$$

Ns is the number of proliferative colonies in the presence of stimuli and Nr is the number of proliferative colonies in RPMI medium alone.

2.7. Human PBMC ELISpot assay

The frequency of Sap epitope-specific T lymphocytes was determined with an IFN-specific ELISpot assay according to the manufacturer's recommendation (R&D systems, Minneapolis, MN). In brief, polyvinylidene difluoride membrane (PVDF) in each well was pre-wet with RPMI-1640 and 100 μl of 2×10^6 PBMC were added to each well. Then, RPMI 1640, phytohemaaglutinin (PHA) or synthetic peptide (final concentration at 10 μg/ml) was added into each assigned well. The culture was incubated in a humidified 37 °C 5% CO_2 incubator for 18 h. At the end of the incubation period, culture medium in each well was aspirated and the wells were washed by wash buffer using semi-autowasher (Nunclonâ, Nalge Nunc International, Rochester, NY, USA). Then, 100 μl of diluted detection antibody were added to each well and incubated at 4 °C overnight. The plate was subsequently washed and 100 μl of diluted alkaline phosphatase-conjugated streptavidin were added to each well. The plate was incubated further for 2 h at room temperature. After washing, 100 μl of 5-bromo-4-chloro-3′-indoylphosphate-*p*-toluidine salt/nitro blue tetrazolium chloride (BCIP/NBT) chromogen were added into each well and the reaction was allowed to continue for 1 h at room temperature in a dark chamber. Finally, chromogenic solution was removed from the plate and the plate was rinsed with deionized water. The plate was inverted, tapped on a paper towel to remove excess water and dried completely at room temperature. The number of spots was assessed by automated KS ELISpot reader (Model: ZEISS Stemi 2000-C) (Carl Zeiss, Germany).

2.8. Statistical analysis

SPSS® statistical software package Version 9.0 (SPSS Inc., 1998) was used for data analysis of peptides stimulation. Paired *T*-test was used to determine the differences between groups. *P* values of ≤ 0.05 were considered to be statistically significant.

3. Results

Localization of 153 *C. albicans* virulence factors, taken from 80 scientific literatures and listed in CandiVF database (Tongchusak et al., 2005), is shown in Fig. 1. When 452 protein entries from CandiVF database were submitted to the Hotspot Hunter Program (Zhang et al., 2005) to predict promiscuous HLA-DRB1 binding peptides, the results showed that there were 2639 peptide sequences predicted as T cell epitopes. Among these entries, no T cell hot spot was observed in 27 primary sequences. The sequences either contained antibody binding peptides, truncated sequences or tandem repeats of Als7 adhesin. Most T cell epitopes were predicted from four protein families including secretory aspartyl proteinase (Sap1–Sap10) family, lipase (Lip1–Lip10) family, phospholipase (Pla-Pld) family and agglutinin-like sequence (Als1–Als9) family.

By using Hotspot Hunter Program to predict consensus T cell epitope of Sap family of proteins, we found that there was a T cell epitope located at the carboxyl-(C-)terminal region (Fig. 2). T cell epitopes of Saps1–3 demonstrated 100% homology and shared approximately 94% homology to those of Saps4–6. The predicted T cell epitope of Sap8 is 4-mer longer than those of Saps1–3 but shared greater than 90% identity with the 15-mer predicted sequence of Saps1–3. Lesser to insignificant degrees of homology with T cell epitopes of Sap5, Sap7, Sap9 and Sap10 were found as compared to those of the former group (Fig. 2).

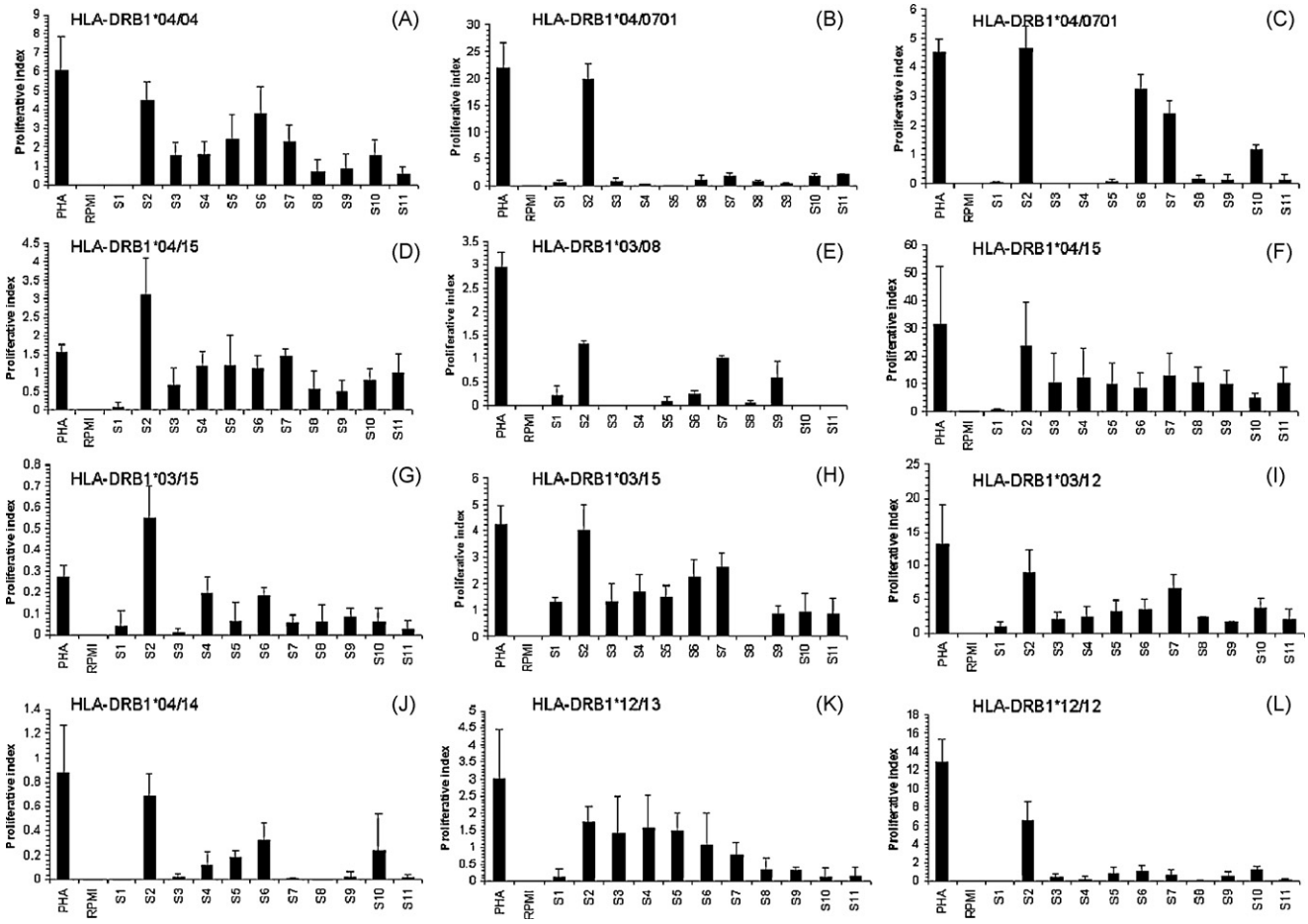


Fig. 3. Degrees of PBMC proliferation of 12 donors upon stimulation with overlapping peptides of Sap. Values are means ± S.D. of three determinations.

Among the proteins of Sap family, Sap2 is the common antigen expressed in all cases of candidiasis. Thus, 11 overlapping peptides were constructed based on the Sap2 amino acid sequence. Peptides S6–S10 span the predicted consensus T cell epitope of Sap sequence. These peptides were used to stimulate PBMC from 12 HLA-DRB1 blood donors (including HLA-DRB1*04/04, *04/0701, *04/0701, *04/15, *04/15, *03/15, *03/08, *03/15, *03/12, *04/14, *12/13 and *12/12) (Table 2). Fig. 3 showed different degrees of proliferative response to Sap peptides of all 12 donors. Peptides S6 (QLLFDVNDANILGDNFLRSA) and S7 (QLLFDVNKANILGDNFLRSA), which lie in the predicted area, were able to significantly induce PBMC activation of 9 of 12 donors. Notably, these two peptides contain almost identical amino acid sequences except for one amino acid at position 8 (D for S6, K for S7) from the amino-(N-)terminal. There were two donors (HLA-DRB1*03/08 and *04/14) that responded to only one of the two peptides. The other two donors did not show any response to peptides S1 and S3–S10. The latter two non-responders to the consensus T cell epitope (S6–S10) possess HLA-DRB1*12/12 and 04/0701.

Interestingly, peptide S2 (DSNLSGDVVFNFSKNAKISV), that lays upstream of the conserved T cell epitope of Sap family, was able to stimulate PBMC of all donors (Fig. 3). The ability of S2 peptide to stimulate T cell was confirmed by ELISpot assay based on interferon secretion by PBMC. S2 enabled to activate gamma-interferon release from PBMC of all donors (Fig. 4). The response to this peptide was significant as the number of interferon producing clones exceeded 3500 spots/10^6 PBMC (Fig. 4).

Three-dimensional structure data of Sap2 were gathered from PDB and homology modeling assignment (Fig. 5). The resolution Sap2 three-dimensional structure was 2.10 Å. From the structural analysis, peptide S2 displays two β-sheets with –DVVFNFS-(β261–β267) and –AKIS-(β270–β273) sequences. The constituted hydrophobic amino acids possess side chains that point down towards the inner part of Sap2 molecule (Fig. 5).

4. Discussion

Our strategy of peptide-based vaccines is to identify promiscuous, as well as consensus, T cell epitope of Sap family of proteins. The epitope should be a relevant representative of large proportion of the human population. The majority of publicly available programs usually predicted the peptides that specifically bound to an individual HLA molecule and has not been properly assessed for predictive accuracy. MULTIPRED was developed based on the fact that a peptide could bind to multiple MHC allelic variants and closely

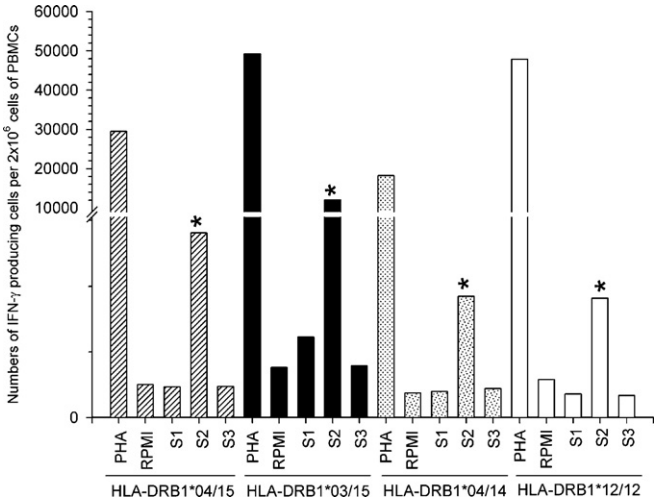


Fig. 4. The number of gamma-interferon producing cells in response to the S2 synthetic peptide. T cell proliferation of four representative donors HLA-DRB1 alleles *04/15, *03/15, *04/14, and *12/12 were assessed by ELISpot assay. *Significantly different from controls ($P < 0.05$).

related MHC alleles sometime contain common position-specific amino acids that interact with the peptide. MULTIPRED therefore probably replaced the earlier requirements for individual prediction model for each HLA. It was first trained to be both sensitive and highly specific for HLA-A2 and HLA-A3 (Bian et al., 2003; Zhang et al., 2005).

In this study, *in silico* experiment (dry lab) was performed to simulate Class II HLA antigen presentation of *C. albicans* virulent antigens of the Sap family. Unlike virus, *C. albicans* is an extracellular pathogen. Therefore, after fungal antigens are processed, the digested peptides will be presented in the context of Class II HLA molecule to $CD4^+$ T cell (Claudia et al., 2002). With the prediction of peptide binding to Class II MHC, there have been only two computation systems available, namely, MULTIPRED and TEPITOPE (Zhang et al., 2005). Hotspot Hunter Program, which is a suitable prediction tool for Class II HLA-DRB1 based on MULTIPRED, was chosen as a computational method attached to CandiVF. The program used artificial neural network (ANN) method and Hidden Markov Model (HMM) as predictive engines for identifying antigenic cluster of peptides that had ability to fit in the groove of Class II HLA molecules. ANN had been used for the prediction of peptides that bound to both Class I and Class II HLA molecules with the sensitivity and specificity closed to 80% (Brusic et al., 2004). HMM was a novel predictive engine for T cell epitope prediction (Noguchi et al., 2003). The use of HMM in the prediction of peptide antigenicity had been demonstrated in Class I HLA-A2 but not Class II HLA and the model gave a high accuracy of prediction (Noguchi et al., 2003). Notably, the difficulty of Class II HLA epitope prediction was the length of peptides (approximately 11–30 residues) presented in context of HLA. Therefore, the prediction was selected as a cluster of peptides. Noguchi et al. were the first to combine HMM with successive state splitting (SSS) algorithm for optimization of the HMM structure and used it for the prediction of peptides that bind to Class II HLA-DRB1*0101 (Noguchi et al., 2003). They demonstrated that S-HMM prediction accuracy was comparable to fully connected HMM and ANN methods.

In this study, Hotspot Hunter Program identified immunological hot spots of about 14 amino acids in length at the C-termini of Sap1–Sap10 proteins. The sequence was consensus among Saps. In order to verify the quality of prediction, ''wet'' laboratory testing was conducted. Overlapping peptides of Sap2 sequence were synthesized and used to stimulate PBMC isolated from Thai blood donors with different HLA-DRB1 background. S6–S10 contained the predicted consensus sequence. PBMC of most responsive donors, which were stimulated with S6 and S7, consistently showed the significant proliferative indices. Notably, S6 and S7 have only one amino acid different from each other and as a result, most donors' PBMC responded well to both peptides. Only PBMC of the donors with HLA-DRB1*03/08 or *01/14 were able to proliferate upon stimulation by either S7 or S6, respectively. There were only two non-responsive donors whose PBMC could not be activated by any tested peptides within the predicted area. Of these two non-responders, one has HLA-DRB1 allele (HLA-DRB1*12/12) that does not match with the alleles selected by Hotspot Hunter analysis.

It is interesting to note that S2, whose sequence is not conserved among the Saps, could significantly stimulate PBMC

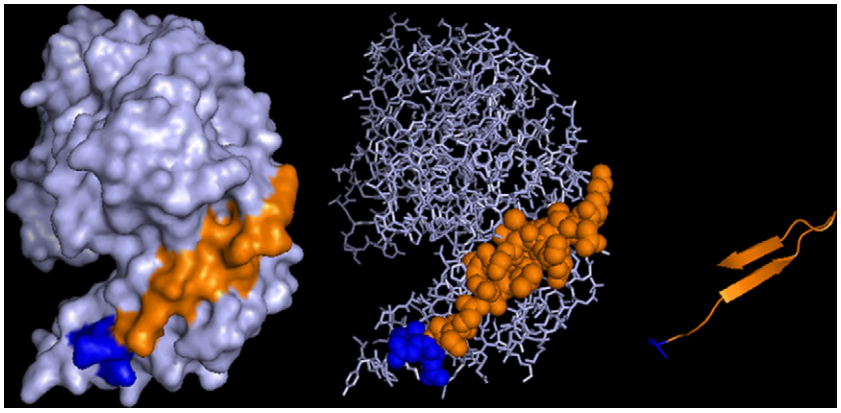


Fig. 5. Three-dimensional structure of the predicted S2 epitope located on the Sap2 molecule generated by Pymol.

and T cells of every donor tested. More importantly, S2 was identified by the Hotspot Hunter Program as T cell epitope of Sap2, one of the most immunogenic proteins during *C. albicans* infection. S2 could activate T cells of all donors to release gamma-interferon regardless of their HLA-DRB1 types. The three-dimensional structure analysis of Sap2 showed that S2 lies on the surface of Sap2 molecule. The two β-pleated sheets of S2 are enriched with hydrophobic amino acids Y, L, V, F and I, which were previously reported to be common anchor residues at peptide position 1 for HLA-DRB1 binding (Jemmerson and Paterson, 1986). The prediction results were also corroborated with the general perception that T cell binding peptides are those with short segment of continuous structure, which is limited to primary or secondary structure of the protein.

In summary, Hotspot Hunter Program could thus predict not only high binding individual peptides but also regions of immunological hot spots in an antigen, which could have potential therapeutic significance as peptide vaccines. This immunomics approach to vaccine design benefits over the conventional approach as it increases the efficiency of T cell epitope screening without the need of antigen purification. The technology itself will be further enhanced by additional experimental data enrichment of training datasets.

Acknowledgements

This work was supported by BIOTEC, the National Science and Technology Development Agency (NSTDA), Bangkok, Thailand. Songsak Tongchusak was supported by the Thailand Research Fund through the Royal Golden Jubilee PhD. Program (Grant No. PHD/021/46). We are grateful to Dr. Chanvit Leelayuwat for his kind assistance in HLA-typing and Dr. Pokrath Hansasuta for his technical assistance in ELISpot assay. V. Brusic acknowledges support in part from the ImmunoGrid project under the EC contract FP6-2004-IST-4, No. 028069.

References

Akiyama, K., Shida, T., Yasueda, H., Mita, H., Yannagihara, Y., Hasegawa, M., et al., 1996. Allergenicity of acid protease secreted by *Candida albicans*. Allergy 51, 887–892.

Bian, H., Reidhaar-Olson, J.F., Hammer, J., 2003. The use of bioinformatics for identifying class II-restricted T-cell epitopes. Methods 29, 299–309.

Brusic, V., Bajic, V.B., Petrovsky, N., 2004. Computational methods for prediction of T-cell epitopes—a framework for modeling, testing, and applications. Methods 34, 436–443.

Calderone, R.A., Ronald, L.C., 2002. Host recognition. In: Calderone, R.A., Ronald, L.C. (Eds.), Fungal Pathogenesis: Principles and Clinical Applications. Marcel Dekker, New York, pp. 1–24.

Claudia, M., Bacci, A., Silvia, B., Gaziano, R., Spreca, A., Romani, L., 2002. The interaction of fungi with dendritic cells: implications for Th immunity and vaccination. Curr. Mol. Med. 2, 507–524.

Horner, W.E., Helbling, A., Salvaggio, J.E., Lehrer, S.B., 1995. Fungal allergens. Clin. Microbiol. Rev. 2, 161–179.

Hube, B., Naglik, J., 2001. *Candida albicans* proteinases: resolving the mystery of a gene family. Microbiology 147, 1997–2005.

Jemmerson, R., Paterson, Y., 1986. Mapping epitopes on a protein antigen by the proteolysis of antigen–antibody complexes. Science 232, 1001–1004.

Johnson, R.A., 2000. HIV disease: mucocutaneous fungal infections in HIV disease. Clin. Dermatol. 18, 411–422.

Kretschmar, M., Hube, B., Bertsch, T., Sanglard, D., Merker, R., Schroder, M., Hof, H., Nichterlein, T., 1999. Germ tubes and proteinase activity contribute to virulence of *Candida albicans* in murine peritonitis. Infect. Immunol. 67, 6637–6642.

Naglik, J.R., Challacombe, S.J., Hube, B., 2003. *Candida albicans* secreted aspartyl proteinases in virulence and pathogenesis. Microbiol. Mol. Biol. Rev. 67, 400–428.

Naglik, J., Albrecht, A., Bader, O., Hube, B., 2004. *Candida albicans* proteinases and host/pathogen interactions. Cell. Microbiol. 6, 915–926.

Navarro-Garcia, F., Sanchez, M., Nombela, C., Pla, J., 2001. Virulence genes in the pathogenic yeast *Candida albicans*. FEMS Microbiol. Rev. 25, 245–268.

Noguchi, H., Kato, R., Hanai, T., Matsubara, Y., Honda, H., Brusic, V., Kobayashi, T., 2003. Hidden Markov model-based prediction of antigenic peptides that interact with MHC class II molecules. J. Biosci. Bioeng. 94, 264–270.

Olerup, O., Zetterquist, H., 1992. HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 h: an alternative to serological DR typing in clinical practice including donor–recipient matching in cadaveric transplantation. Tissue Antigens 39, 225–235.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, second ed. Cold Spring Harbor Laboratory Press, New York.

Schaller, M., Hube, B., Ollert, M.W., Schafer, W., Borg-von Zepelin, M., Thoma-Greber, E., Korting, H.C., 1999. *In vivo* expression and localization of *Candida albicans* secreted aspartyl proteinases during oral candidiasis in HIV-infected patients. J. Invest. Dermatol. 112, 383–386.

Senet, J.M., 1997. Risk factors and physiopathology of candidiasis. Rev. Iberoam. Micol. 14, 6–13.

Tongchusak, S., Chaiyaroj, S.C., Veeramani, A., Koh, J.L.Y., Brusic, V., 2005. CandiVF—*Candida albicans* virulence factor database. Int. J. Pep. Res. Ther. 11, 271–277.

Tsushima, H., Mine, H., 1995. Cleavage of human big endothelin-1 by *Candida albicans* aspartic proteinase. FEMS Immunol. Med. Microbiol. 11, 69–72.

Tsushima, H., Mine, H., Kawakami, Y., Hyodoh, F., Ueki, A., 1994. *Candida albicans* aspartic proteinase cleaves and inactivates human epidermal cysteine proteinase inhibitor, cystatin A. Microbiology 140, 167–171.

Zhang, G.L., Khan, A.M., Srinivasan, K.N., August, J.T., Brusic, V., 2005. MULTIPRED: a computational system for prediction of promiscuous HLA peptides. Nucleic Acids Res. 33, 172–179.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 474–483
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Genomic interrogation of ancestral *Mycobacterium tuberculosis* from south India

Sujatha Narayanan [a,b,*], Sebastien Gagneux [b,1], Lalitha Hari [a], Anthony G. Tsolaki [b,2], Suganthi Rajasekhar [a], P.R. Narayanan [a], Peter M. Small [b,1,3], Susan Holmes [b], Kathryn DeRiemer [b,4]

[a] *Tuberculosis Research Centre (TRC), Mayor VR Ramanathan Road, Chetput, Chennai 600031, India*
[b] *Division of Infectious Diseases and Geographic Medicine, Stanford University Medical Center, 300 Pasteur Drive, Stanford, CA 94305-5107, USA*

Received 27 March 2007; received in revised form 19 September 2007; accepted 24 September 2007
Available online 29 September 2007

Abstract

Mycobacterium tuberculosis is a very important global pathogen. One quarter of the world's TB cases occur in India. The tuberculosis strains isolated from south Indian patients exhibit certain phenotypic characteristics like low virulence in guinea-pigs, resistance to isoniazid, thiophene-2-carboxylic acid hydrazide (TCH) and *para*-amino salicylic acid (PAS), and enhanced susceptibility to H_2O_2. Besides this, a large percentage of the isolates harbor only a single copy of IS 6110 which makes these strains distinct. Hence, we have studied the genotypic characteristics of these strains by using advanced techniques like Deletion Micro array, deletion PCR, allelic discrimination RT-PCR using several lineage specific markers and KatG G1388T (non-synonymous) polymorphism along with spoligotyping. The analysis of 1215 tuberculosis patient isolates from south India revealed that 85.2% belonged to the ancestral lineage of *M. tuberculosis*. Comparative whole-genome hybridization identified six new genomic regions within this lineage that were variably deleted.

Keywords: Tuberculosis; Spoligotyping; Genomic deletion; Large sequence polymorphism; Lineage; Vaccine

1. Introduction

Mycobacterium tuberculosis is a very important global pathogen. Without HIV, the tuberculosis (TB) epidemic would now be in decline everywhere. India has the highest number of incident TB cases in the world. With effective DOTS implementation tuberculosis is expected to decline but as HIV continues to spread, the decline of TB is far from reality.

Although one fourth of the global tuberculosis burden stems from India there are a few studies which define the tuberculosis genogroup (Bhanu et al., 2002; Das et al., 2005; Mistry et al., 2002; Narayanan et al., 1997; Kulkarni et al., 2005; Singh et al., 2004; Radhakrishnan et al., 2001). There has been a population based study from south India on transmission dynamics and risk factors associated with transmission (Narayanan et al., 2002). The information obtained in this study suffers from a lack of portability which hinders comparison between global studies. In addition, the south Indian isolates included by others in the previous studies have been limited in number (Guitierrez et al., 2006; Singh et al., 2007). To avoid the problem in portability of data and limitation in number we have used spoligotyping to genotype the isolates from a population based study.

Spoligotyping is a PCR based genotyping method that allows to assess the *M. tuberculosis* genetic biodiversity and provides enough information about the epidemiologically

* Corresponding author at: Tuberculosis Research Centre (TRC), Mayor VR Ramanathan Road, Chetput, Chennai 600031, India.
Tel.: +91 44 2836 9627/9600; fax: +91 44 2836 2528/2529.
E-mail address: sujatha.sujatha36@gmail.com (S. Narayanan).
[1] Present address: Institute for Systems Biology, 1441 North 34th Street, Seattle, WA 98103, USA.
[2] Present address: Center for Molecular Biology and Infection (CMMI), Division of Investigative Sciences, Flowers Building, South Kensington Campus, Imperial College, London SW7 2AZ, United Kingdom.
[3] Present address: Bill and Melinda Gates Foundation, P.O. Box 23350, Seattle, WA 981012, USA.
[4] Present address: School of Medicine, University of California, One Shields Avenue, Davis, CA 95616, USA.

1567-1348/$ – see front matter
doi:10.1016/j.meegid.2007.09.007

important clones in various settings (Cowan et al., 2004; Streicher et al., 2004). The establishment of an international database is an advantage to this method because it is very informative regarding the endemicity or the ubiquitous strains (Filliol et al., 2002, 2003).

The tuberculosis strains isolated from south Indian patients exhibit certain phenotypic characteristics like low virulences in guinea pigs, resistance to isoniazid (Mitchison et al., 1960; Middlebrook and Cohn, 1953), thiophene-2-carboxylic acid hydrazide (TC4) and *para*-amino salicylic acid (PAS) (Joseph et al., 1964) enhanced susceptibility to H_2O_2 and a majority of them belonging to intermediate phage type (Bhatia, 1961; Grange et al., 1978). Besides this a large percentage of the isolates harbor only a single copy of IS6110 (Das et al., 1995) which make these strains distinct. Hence, we wanted to study the genotypic characteristics of these *M. tuberculosis* strains by using advanced techniques like Deletion Microarray, deletion PCR, allelic discrimination RT-PCR using several lineage specific markers and KatG G1388T (non-synonymous) polymorphism (Sreevatsan et al., 1997) along with spoligotyping and IS6110 RFLP which is considered the gold standard for genotyping techniques.

The deletion micro array is an approach complementary to comparative genomics. This involves the interrogation of unsequenced genomes by DNA-microarray to identify sequences present in a fully sequenced isolate but absent from interrogated isolate (Kato-Maeda et al., 2001; Tsolaki et al., 2004). Nucleotide sequences provide robust, portable and comparable data for studying population variation. The mutational processes that generate these variations are understood and sequence data have been successfully used in the study of evolution, bacterial epidemiology and population biology (Baker et al., 2004).

2. Materials and methods

2.1. IS6110 RFLP & spoligotyping

We used a standardized international protocol for IS*6110* RFLP genotyping (van Embden et al., 1993). Spoligotyping was performed as previously described (Kamerbeek et al., 1997). All of the spoligotype patterns were coded using the octal code system (Dale et al., 2001) with *M. bovis* P_3 and H37Rv as positive controls and autoclaved milli Q water as negative control. The DR region was amplified using the DRa (5′biotinylated) and DRb primers. The amplified product was hybridized to a set of 43 immobilized oligonucleotides, each corresponding to a unique spacer sequence within the DR locus. Hybridization was detected by enhanced chemiluminescence (ECL Enhanced Chemo-Luminescence Detection Kit; Amersham, Hongkong) followed by exposure to X-ray film according to the, manufacturer's recommendations. The results were documented in the form of binary code (*x* and 0) according to the hybridization (respectively positive or negative) for each spacer nucleotide probe and entered into excel format.

We then referred to a standardized international database of spoligotype patterns, SpolDB4 (Brudey et al., 2006) to determine whether each pattern had been previously reported. The updated SpolDB4 version contained 39,295 patterns distributed into 1939 shared types (patterns reported at least twice that grouped 36,925 clinical isolates) and 337 orphan patterns from more than 120 countries. The database provides detailed information about the country of origin and the geographic distribution within eight regions of the world.

2.2. Identification of large sequence polymorphism

We selected 25 isolates for comparative whole-genome hybridization using an Affymetrix DNA chip (Santa Clara, California, USA) following procedures previously described. We identified putative deletions in the experimental strains relative to the sequenced reference strain H37Rv using DelScan software (AbaSci, San Pablo, California, USA), and confirmed the putative deletions by PCR and direct sequencing. We used phylogenetically informative large sequence polymorphisms (LSPs) or genomic deletions, to screen one isolate per spoligotype patterns (102 isolates) by PCR or multiplex real-time PCR. The screening results from the clustered isolates were extrapolated to the remaining isolates of their respective clusters.

We used previously published studies of genomic deletions (Tsolaki et al., 2004) to identify phylogenetically important genomic deletions and screen for them using PCR. For the detection of LSPs by multiplex real-time PCR, we designed a series of assays based on different TaqMan primer/probe combinations. The presence or absence of RD105 and RD239 was evaluated using following primers and probes: RD105-F, 5′-AACGAACTGCGCACTGAACTC-3′; RD105-R, 5′-TCCCGCACCGGTTGAG-3′; RD105-probe, 5′-FAM-AGAGTGGACAGTTTCG-MGBNFQ-3′; RD239-F′ 5′-CGAGCTCAATCCGAACGAAA-3′; RD239-R, 5′-CCGGGCTTGGCTTTAACTG-3′; RD239-probe, 5′-VIC®-CCAGGTGCTTGCCATGMGBNFQ-3′. Samples with no increase in fluorescence of either FAM or VIC® were considered deleted for the corresponding region of interest. Appropriate positive and negative controls were run on each plate. Isolates with an increase in both FAM and VIC® signal were considered not deleted for either RD and analyzed further. Similarly, the presence or absence of TbD1 and RD9 was evaluated with the following primers and probes: TbD1-F, 5′-CCGATTGACCACAGCTCGAT-3′; TBD1-R, 5′-CTGGCCGACGCTTTGC-3′; TbD1-probe, 5′-FAM-CCGTTTCAGATCAGC-3′; RD9-F, 5′-TGGTGGCGGTAGGTTTCAC 3′; RD9-R, 5′-ATGACCCGCGCGATGT-3′; RD9-probe, 5′-VIC®-TTCGACCCCAA GAC.

2.3. Real-time PCR

The Euro-American lineage was defined based on a characteristic seven base pair deletion in pks15/1 (Constant et al., 2002) or the ctg to cgg substitution at codon 463 of *katG*[i]. These two markers are known to be equivalent based on previous studies (Baker et al., 2004; Gutacker et al., 2002; Marmiesse et al., 2004). The *katG*463 SNP was analyzed by allelic discrimination using a TaqMan multiplex real-time PCR assay (Applied Biosystems, Foster City, CA, USA). The

TaqMan-probes were labeled with 6-carboxyfluorescein (FAM) or VIC® at the 5′-end, and conjugated with a nonfluorescent quencher (NFQ) and a minor groove binding (MGB) group at the 3′-end. The following primers and probes were used: *katG*463-L: 5′-CCGAGATTGCCAGCCTTAAG-3′, *katG*463-R: 5′-GAA-ACTAGCTGTGAGACAGTC AATCC-3′, *katG*463-cgg probe: 5′-FAM-CAGATCCGGGCATC-MGBNFQ 3′, *katG*463-ctg probe: 5′-VIC®-CCAGATCCTGGCATC-MGBNFQ 3′. Samples were run in 25 μl reactions in 96-well plates on an ABI 7000 sequence detection system (Applied Biosystems, Foster City, California, USA). The reaction conditions were as recommended by the manufacturer.

3. Results

3.1. Strain diversity based on spoligotyping

The study was approved by the ethical review committee of the Tuberculosis Research Centre (TRC, 1999) and Stanford University. There were 3036 persons diagnosed with pulmonary tuberculosis during the study period, June 1999–June 2002, in the Tiruvallur District, Tamil Nadu state, south India. Of these, 1456 (48.0%) were culture positive for *M. tuberculosis*. Spoligotyping was performed on 83.4% (1215/1456) of patient isolates. Based on the spoligotype patterns, there were 102 different genotypic clusters, each with at least two different patients (median, three patients; range, 2–338 patients) and a total of 1038 tuberculosis patients (85.4%) in the clusters (Supplementary table). The remaining 177 (14.6%) tuberculosis patients had an isolate with a unique spoligotype pattern.

The spoligopattern of all the isolates were compared to SpolDB4 (Brudey et al., 2006) to determine whether each pattern has been previously reported. We found 42 clusters with spoligotypes matching the SpolDB4 (Fig. 1). There were 23 unique isolates with spoligopatterns which matched SpolDB4 database. 56 clusters had spoligotypes which did not match the SpolDB4 database (Fig. 2). These were considered as clusters of orphan isolates. There were 152 orphan strains which were not involved in clusters. The pattern of the 42 major clusters which matched the SpolDB4 data base has been shown (Fig. 1). Three clustered clinical isolates had spoligotype pattern lacking spacers 20, 21 and 33–36 a profile generally associated with *M. tuberculosis* H37Rv and 1 isolate had a pattern of *M. microti*. When we analyzed our data using spotclust (online software) the tuberculosis isolates were classified into nine major clades with sub clades (Tables 1A and 1B). The Spotclust assigned the orphans to be variants of major clades and thus reduced the number of orphans to 30 (2.4%) (Tables 1A and 1B).

Table 1A and Supplementary table shows the different spoligotype clades and spoligotypes present in the Thiruvallur region, respectively. EAI 3 and EAI 5 are represented predominantly upto 41% each whereas 5% of the isolates belong to EAI 1 and 4 and 1.6% belong to EAI 2. Compared to northern regions like Delhi and Lucknow, the CAS isolates are represented by low numbers in Thiruvallur (4.0%) (Singh et al., 2007). The Beijing, LAM, T1 and Family 33 genotypes accounted for 2.0%, 1.0%, 3.0% and 1.0%, respectively.

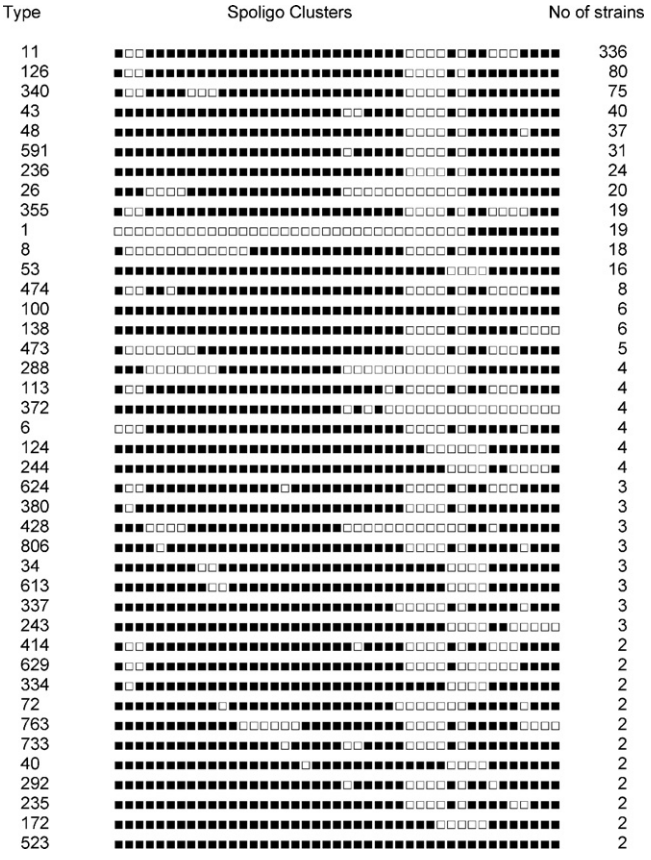


Fig. 1. Clustered spoligo types matching Spol DB 4.0.

Table 1B shows the phylogenetic tree. Phylogenetic analysis was carried out with spoligotyping data of all the isolates included in the study. Nei's original measures of genetic identity and genetic distance analysis was carried out using POPgene software version 1.32 (http://www.ualberta.ca/fyeh/

Table 1A
Spoligotyping of *M. tuberculosis* strains from Tiruvallur, Tamil Nadu

Spoligotype	Isolates	Clusters	Unique
EAI 1	5	4	1
EAI 2	20 (1.6%)	8	12
EAI 3	498 (40.9%)	448	50
EAI 4	5	4	1
EAI 5	498 (40.9%)	406	92
Beijing	24 (1.9%)	23	1
CAS	48 (3.9%)	34	14
Harleem 1	12 (0.9%)	7	5
LAM 7	1	0	1
LAM 8	7	2	5
LAM 9	6	2	4
T1	36 (2.9%)	22	14
T2	2	2	0
X2	1	0	1
X3	1	0	1
Family 33	14 (1.1%)	10	4
Family 36	3	2	1
H37RV	3	2	1
Microti	1	0	1
Orphan	30 (2.4%)	16	14
Total	1215	992 (81.6%)	223 (14.4%)

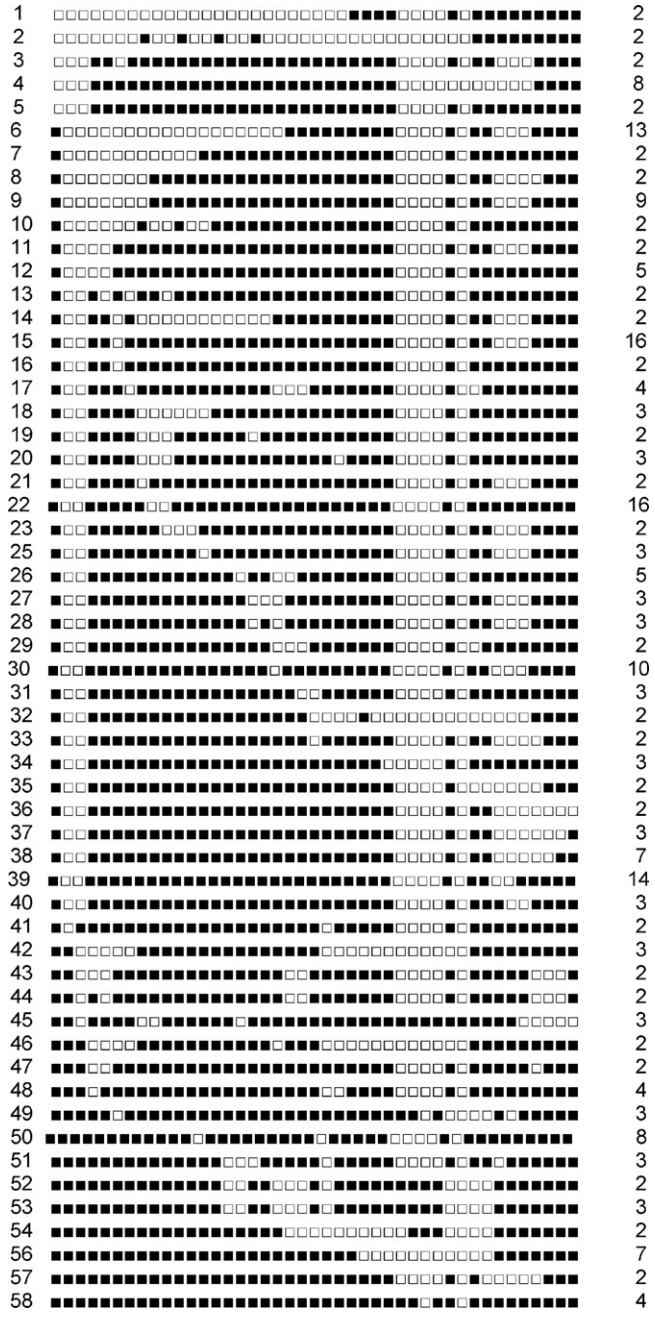
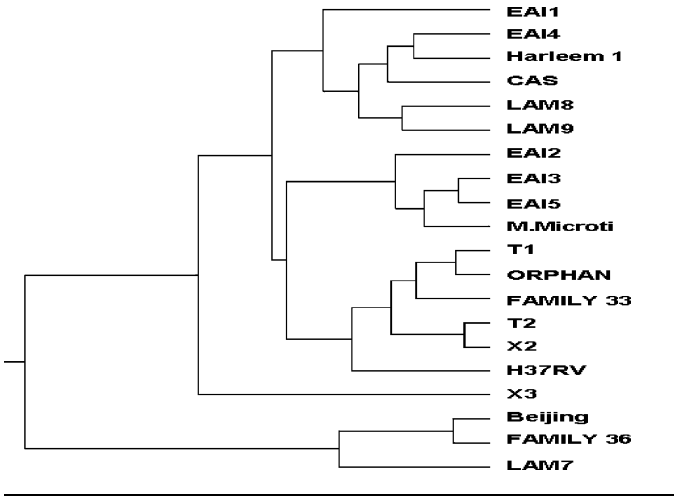


Fig. 2. Orphan strains in cluster.

Table 1B
Spoligotyping of *M. tuberculosis* strains from Tiruvallur, Tamil Nadu



download). The phylogenetic tree has grouped the subclades EAI 2, 3 and 5 together which is farther from EAI 1 and 4 subgroups, represented poorly in Thiruvallur. EAI 3 and 5 are present as predominant sub clades in Thiruvallur. More number of EAI 1 and 4 subclades are present in Denmark, The Netherlands, USA, Bangladesh and Vietnam (SpolDB4) EAI 1 and 4 are grouped closer to Harlem, CAS and LAM 8 and 9.

3.2. Spoligotyping and IS6110 RFLP

Among the 1192 isolates which had been typed by IS6110, 20 isolates had no copy. Four hundred and ninety three isolates had single copy, 249 isolates had 2–5 copies. Four hundred and thirty isolates were IS6110 high copy isolates. Among the 20 no copy IS6110 strains, 14 had unique spoligotype patterns and remaining six isolates were in three clusters of two each. Thus, among the no copy IS6110 strains, spoligotyping has been highly discriminatory (17 types among 20 strains). We tried to find out whether the major spoligotype clusters could be differentiated by IS6110 RFLP. Tables 2 and 3 show the typed clusters (>2 strains in a cluster) and orphan clusters (<than 3 isolates in a cluster) and the degree of differentiation by IS6110 RFLP. Among the total 733 isolates involved in clusters with types matching the database 332 strains (45%) had a single copy, 195 (26.6%) showed low copy strains and 8 (1.0%) showed no copy isolates. Totally 63% of strains could not be discriminated by both spoligotype and IS6110. Similarly among the isolates involved in orphan clusters, all the spoligotype clusters (except 2) had isolates with low copy IS6110 strains. The East African variants or the ancient TB lineage forms the predominant group among the south Indian isolates and 63% of the EAI isolates are low copy isolates. The TbD1 ancestral isolates belonging to EAI family were less discriminatory by IS6110 RFLP typing. Seventy-six percent of the isolates belonging to the typed and orphan clusters by spoligotyping are low copy IS6110 isolates (Tables 2 and 3).

3.3. Genomic deletions detected by comparative whole-genome hybridization

To identify genomic differences putatively associated with the altered virulence, we performed whole-genome hybridization on a subset of 25 strains belonging to the dominant ancestral lineage using the Affymetrix *M. tuberculosis* GeneChip Trade Mark (Santa Clara, CA, USA) and published methods (Gingeras et al., 1998; Kato-Maeda et al., 2001; Tsolaki et al., 2004, 2005). We confirmed putative deletions by PCR and sequencing as previously reported (Chen, 2006; Gagneux et al., 2006; Glynn et al., 2002; Gutacker et al., 2006).

Table 2
Orphan clusters and IS6110 patterns

S. no.	Orphan clusters	No. of isolates	IS single copy	IS low copy	IS high copy	No copy
1	467 777 777 413 071	15	14	1	0	0
2	400 000 377 413 071	14	11	1	1	1
3	477 777 777 413 171	14	3	7	4	0
4	476 377 777 413 771	13	7	5	1	0
5	477 777 377 413 071	9	7	2	0	0
6	077 777 777 400 071	8	0	2	6	0
7	777 737 757 413 771	8	0	1	7	0
8	777 777 774 000 771	7	0	0	7	0
9	477 777 777 413 011	8	7	0	0	1
10	417 777 777 413 771	5	4	1	0	0
11	477 773 177 413 771	5	5	0	0	0
12	401 177 777 413 071	4	4	0	0	0
13	473 777 077 410 771	4	1	3	0	0
14	737 777 747 413 771	4	1	3	0	0
15	777 777 777 733 771	4	0	1	3	0

In the 25 strains, we identified a total of 150 putative deletions, including 38 that involved highly repetitive PE-/PPE-PGRS sequences and which were excluded from further analysis because of the difficulties in precisely determining deletion boundaries (Tsolaki et al., 2004). Analysis of the remaining 112 putative deletions determined that 31 (27.6%) were false positives and 81 (72.3%) were true deleted sequences. The locations of the true deletions were mapped to the base pair and their sequences were named regions of difference (RD) with respect to H37Rv (Brosch et al., 2002; Cole et al., 1998).

All 25 strains had a genomic deletion in RD239 (as expected), RD147C and RD198a. These three large sequence polymorphisms (LSPs) have been previously reported (Gagneux et al., 2006; Tsolaki et al., 2005). In addition, we found six new LSPs (Table 4). To confirm that the RD239-deleted strains from south India belonged to the ancestral lineage of *M. tuberculosis*, we used multiplex real-time PCR to screen for TbD1 (Brosch et al., 2002) in the 25 strains following published procedures (Gagneux et al., 2006). As expected, we found that the TbD1 region was intact in all 25 strains.

To estimate the frequency of the six newly identified LSPs (RD768–RD773), 100 isolates belonging to the ancestral lineage were chosen from the largest spoligotype clusters and screened by PCR. We found that RD769 and RD771 were each detected in only 2 of the 100 isolates, and RD768, RD770, RD772 and RD773 were not deleted in any of the 100 isolates. Earlier studies have reported that 40% of the strains in south India are low virulent in guinea pigs (Prabhakar et al., 1987). Because of the low frequency of these LSPs, these genomic alterations are unlikely to account for the low virulence of south Indian strains in animal models.

3.4. Ancestral strain lineage predominates in south India

To define the population structure of *M. tuberculosis* in the study area, we selected one isolate per spoligotype pattern (102

Table 3
Typed soligo clusters and IS6110 RFLP pattern

S. n.	Spoligotype	No. of iolates	IS single copy	IS low (2– copies)	IS high copy	No copy
1	11	333	144	129	58	3
2	126	80	59	69	3	2
3	340	75	48	19	7	1
4	43	40	5	1	34	0
5	48	37	26	5	6	0
6	591	31	2	2	27	0
7	355	19	14	3	2	0
8	236	24	9	2	11	2
9	26	20	0	2	18	0
10	53	16	2	1	13	0
11	8	18	13	4	1	0
12	473	5	0	5	0	0
13	474	8	6	2	0	0
14	288	4	0	0	4	0
15	25	4	0	0	4	0
16	138	5	2	1	2	0
17	100	6	2	3	0	1
18	372	4	0	0	4	0
19	124	4	0	0	4	0

Table 4
Description of six new large sequence polymorphisms (LSPs) or genomic deletions observed in the clinical isolates from south India

Isolate	Name*	Coordinates		Primer sequence	Size† (bp)	Putative function of deleted genes and base pairs
		Start	Stop			
M394	RD768	70289	72023	GGG GCG GCT GTT GGA CCC GCA TAT CCT	1735	Rv0064 Trans membrane protein Rv0065
				ATC AAG ACC GGT AAC GAG CGG GCC AAA CTC		Conserved hypothetical protein
M357	RD769	1957030	1959686	GAC AGC AAC CGC GAC GCC CGG AAT C	2657	Rv1730c penicillin binding protein
				CCC GCC CTC GTC GTC ACC TTC ATC TGT AA		Rv1731 (*gabD2*) succinate-semialdehyde dehydrogenase (NADP+) dependent (SSDH) *gabD2*
						Rv1732c conserved hypothetical protein
M395	RD770	2196902	2200337	CCG GTG ACC GTC GTG GTG AGC ACC A	3436	Rv1946 (*lppG*) possible lipoprotein
				CCA GGA CGG AGG TCA CAG TTG CGG GGT		Rv1947 hypothetical protein
						Rv1948 hypothetical protein
M118	RD771	89500	90450	CCG GGC GCG CGA ACA TGG ACT GC	951	Rv0081 probable transcriptional regulatory protein
				GGC TCG GCG CCT CCG GGT GG		Rv0082 probable oxidoreductase
						Rv0083 probable oxidoreductase
M165	RD772	30669	34074	GCC ATC GCG GAG GCG GAA GCA GCT CT	3406	Rv0027 conserved hypothetical protein
						Rv0028 conserved hypothetical protein
				TTT GCC CGG CCT AGC GGT TGC CCA TC		Rv0029 conserved hypothetical protein
						Rv0030 conserved hypothetical protein
						Rv0031 possible remnant of a transposase
M461	RD773	3434523	3441337	CGG CCC TGA CGG TGG CAA TCT GGA TC	6815	Rv3071-Rv3076 conserved hypothetical
						Rv3077 possible hydrolase
						Rv3078 (*hab*) probable hydroxylaminobenzene mutase *Hab*
				GAG CAG GGT CGC CAG CCA GTT GCC		Rv3079c conserved hypothetical protein
						Rv3080(*pknK*) probable serine/threonine-protein kinase
						Transcriptional regulatory protein *pknK* (protein kinase k)

bp = base pairs.
* Name assigned to the specific genomic deletion.
† Size of genomic deletion (bp) = (stop coordinates − start coordinates).

isolates) and screened by multiplex real-time PCR for the main strain lineages using the lineage-specific DNA polymorphisms (Gagneux et al., 2006). We first tested for the presence or absence of the genomic region of difference (RD) 105 and RD239 using the primers and probes (Gagneux et al., 2006). RD105 and RD239 are markers for the East Asian (''W-Beijing'') and ancestral strain lineages, respectively (Gagneux et al., 2006; Tsolaki et al., 2005). The screening results from the clustered isolates were extrapolated to the remaining isolates of their respective clusters. Based on this first round of screening,

Table 5
Results of real-time pCR screening of one isolate per strain, as defined by spoligotyping

Genomic deletion RD105	Genomic deletion RD239	Number of isolates	Percentage of isolates
Deleted	Present	28	13.0
Present	Present	143	11.8
Deleted	deleted	4	0.3
Present	Deleted	1035	85.2
Indeterminate	Indeterminate	5	0.4
Total		1215	100.0

we found that 85.2% (1035/1215) of the isolates had a deletion in RD239, 2.3% (28/1215) of isolates had a deletion in RD105, 11.8% (143/1215) of isolates had no deletion, and the genomic deletions in 0.7% (9/1215) of isolates remained undetermined (Table 5). We performed a second round of screening on the isolates without a confirmed deletion in RD105 or RD239 using a TaqMan allelic discrimination assay for the *katG*463 ctg to cgg single nucleotide polymorphism, which defines the Euro-American lineage (Principal Genetic Groups 2 and 3) (Gagneux et al., 2006). We found 6.8% (83/1215) of isolates belonged to this lineage. Finally, close inspection of the spoligotypes of the remaining isolates (Supplementary table) revealed that 4.4% (54/1215) belonged to the CAS lineage or the Delhi genogroup (Baker et al., 2004; Brudey et al., 2006).

4. Discussion

For decades there has been a belief that the outcome of a person's exposure to and infection with *M. tuberculosis* depends only on individual host characteristics such as age, gender and genetic background. However, recent studies increasingly suggest that bacterial factors also contribute to the differences in outcomes seen in the human host population. Hence, it becomes imperative to characterize the clinical isolates of *M. tuberculosis* from a large population based study in order to design vaccine and new drugs for the population. One quarter of the world's TB cases occur in India and the major lineage which has been identified in this study is likely to be responsible for a high percentage of the TB cases.

This study describes the spoligotyping of 1215 isolates of *M. tuberculosis* from Tiruvallur area in south India where the largest BCG trial was conducted. The total population of the study area has been 580,000. The incidence of smear positive TB is 76 per 100,000 in the population. We have compared the spoligotyping with the widely accepted IS6110 RFLP typing from a 3-year population based study. In addition to these two methods Deletion Micro array based large sequence polymorphism (LSP), deletion PCR and RT-PCR were used as large scale screening methods to delineate the phylogeny and evolutionary characteristics of the south Indian isolates of *M. tuberculosis*. To our knowledge this is the first study from India which combines deletion micro array along with the widely used genotyping tools in a population based study.

The genetic diversity observed in our study has been 22.5% which is corroborating with the fact that endemic area might have relatively few circulating strains (Hermans et al., 1995). Ours is a population based study and restricted to southern most region of the country. Hence, this data contradicts the study of (Singh et al., 2007) which included seven different regions of India and showed a large number of circulating strains. This highlights the rich diversity among different regions of India.

Spoligotyping has been a very useful technique to differentiate the *M. tuberculosis* isolates into different geographical clades. It is not a highly discriminatory marker. Studies from parts of the globe have highlighted the predominant spoligotypes responsible for most of the pulmonary TB cases (Brudey et al., 2003; Dale et al., 1999; Puusteinen et al., 2003). These studies reveal that different spoligotypes occur at variable frequencies in different countries and continents.

The CAS or Delhi type corresponding to ST26 which predominates in north India (Bhanu et al., 2002; Singh et al., 2004; Kulkarni et al., 2005; Guitierrez et al., 2006) represents 1.02 % of isolates SpolDB4 database. This type has been reported from 34 countries. None of the orphan strains identified by Kulkarni et al. (2005) or Singh et al. (2007) were found in our study. Spoligotyping has categorically identified the major clade present in south India to East African lineage with 28,32 and 34 spacers missing. This clearly shows the difference in origin and evolution of *M. tuberculosis* in north and south of India. In our study, the EAI family corresponding to ubiquitous spoligotype ST11, is present as a large group among the clusters matching the SpolDB4. 336 isolates among 1215 isolates belong to this group. This family is present in other countries like Bangladesh, Malaysia, Indonesia, Pakistan, UK, Denmark, Netherlands, France, New Zealand and USA.

The family W-Beijing which has been reported from several countries has a distinct genotype has been associated with rapid transmission across large parts of different continents, accounted to only 1.9% in Thiruvallur. The other clades have been represented by a fewer percentage in Thiruvallur (Tables 1A and 1B).

Sixteen strains of type 53 were seen in the present study from Tiruvallur. This type is linked to be a major clade of European descent in a multicentric study (Kremer et al., 1999). The presence or absence of spacer31 discriminates between the two highly prevalent spoligotypes worldwide, ST 50 and ST 53 with its absence being linked to IS6110 insertional events. While Mumbai study (Kulkarni et al., 2005) showed the presence of one strain of ST 50, we have seen ST 53 in the south Indian region.

In spite of the endemicity of the disease a large proportion of the patients were infected with the tuberculosis isolates having spoligotype that had not been described in the global database, of more than 39,000 tuberculosis isolates. But there has been a major cluster of isolates constituting 85% of the total isolates from this endemic area. Twelve major clusters ST 1, 8, 11, 126, 340, 43, 48, 591, 236, 26, 355, 53 comprised 86%. Except 1 and 53 all the remaining clusters are variants of EAI with the characteristic spoligotype pattern with absence of spacers 28–32 and 34 which may be considered as a major type in the Tiruvallur district.

Overall the spoligotyping analysis showed diversity in clinical isolates especially among the strains with no copy of IS6110. The 20 no copy isolates were differentiated to 17 different spoligotypes. Thirty percent of the IS6110 low copy strains were differentiated by spoligotyping. Spoligotyping alone is not sufficient as a secondary marker for IS6110 low copy strains, instead this genotyping method has been an excellent tool to define the geno clade prevalent in the Tiruvallur region of south India in comparison with the other regions of the globe. This region has been predominated by IS6110 low copy strains. 63% the EAI variants are comprised of single and low copy IS6110 isolates. As Table 5 shows that

major spoligo clusters could not be differentiated by IS6110 typing because of the predominance of low copy isolates. This warrants the use of other secondary markers like VNTR, MIRU besides spoligotyping, and other PCR based methods to study the transmission dynamics of *M. tuberculosis* in the region. Even though the discriminating capacity of MIRU is not greater than IS6110, it has been shown to be efficient to discriminate the low IS6110 strains (Sun et al., 2004).

The excellent congruence observed between all the independent sets of genetic markers used here lends strong support to the assignment of different prevalent lineages. Our LSP followed by large scale screening by RT-PCR and PCR has confirmed the prevalent genotypic clade in this region. The presence of TbD1 in a large percentage of isolates indicates that south India could have been the initial focus of the ancient lineage. Another characteristic feature of these strains is the deletion of RD239 region which is also a characteristic feature of the Manila strains (Tsolaki et al., 2004). Aggregation of 236–239 deletion is confined to the so-called ''Manila'' clade (Tsolaki et al., 2004), and is likely to result from a genetic event that occurred in the ancestor of these isolates. The spoligotyping of 41 of the 48 *M. tuberculosis* isolates from metropolitan Manila yielded characteristic identical patterns lacking hybridization to eight spacers (Bhanu et al., 2002) the remaining 35 spacers showed a positive hybridization signal.

A similar systematic association has been observed in strains from Singapore (Sun et al., 2004) and from Bangladesh (Sola et al., 2001). It remains to be tested whether the deletion of TbD1 from the genome of ancient strains conferred or not some selective advantage to the descendent strains enhancing their epidemic potential. Due to trade links the probable route of migration to and fro from south India to East Africa could have been through East Asian countries like Philippines.

This study is population based and has been spread over 3 years. There have also been studies from this region describing the prevalence of low copy IS6110 strains in the past (Das et al., 1995) In spite of the presence of all the nine clades in this region the only lineage which has been predominating in this region is the EAI clade. Recently, there have been reports on the notion that *M. tuberculosis* has adapted to human population. In this region the ancient genotype or the EAI strain has adapted to the host during several centuries and thus has a capacity for higher transmission. The prevalence of such a high number in this part of the continent and also its prevalence for a long time raise the question whether these have a unique transmission dynamics and are they suppressing the spread of other genotypes. In the largest BCG trial conducted in the same region earlier, BCG did not afford any protection. It is known that genetic variation of the strains is a method of vaccine escape and has been demonstrated in several bacterial species (Van Loo and Mooi, 2002; Mbelle et al., 1999). Further investigation in this line would be helpful to ascertain the role of these highly prevalent strains on the host immunity.

Starting 50 years ago, studies using the guinea pig and other animal models have repeatedly reported that strains of *M. tuberculosis* from south India were less virulent than other strains (Mitchison et al., 1960; Singh and Sharma, 1964; Prabhakar et al., 1987). Furthermore, BCG vaccine trails in south India showed no significant efficacy (TRC, 1999). Considering the global phylogeography of *M. tuberculosis* and the dominance of the ancestral lineage in south India, BCG failure in this part of the world could partially be due to strain-specific effects (Brewer and Colditz, 1995). Based on our approach, which was designed to detect genomic regions that are absent relative to the laboratory reference strain, H37Rv, deletion array was performed.

Deletion micro array followed by large scale screening by PCR had identified six new deletions in the genome of only 1–2% of the south Indian isolates. The phenotypic characteristics lost or gained by these deletions have to be explored. However, considering that this strain lineage is ancestral to all *M. tuberculosis* strains that have been sequenced thus far (Gagneux et al., 2006), this lineage could harbor unique sets of genes that are absent (i.e. deleted) in other *M. tuberculosis* and thus not detectable using H37Rv as a reference.

Acknowledgments

We thank the patients and staff of the Tiruvallur District and the Tuberculosis Research Centre, Chennai, India. This work was funded in part by a grant from the United States Agency of International development provided through the World Health Organisation, partly by the NIH through NIAID/TRC ICER programme and NIH Grant 5R03AI063576-02.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.meegid.2007.09.007.

References

Baker, L., Brown, T., Maiden, M.C., Drobniewski, F., 2004. Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. Emerg. Infect. Dis. 10, 1568–1577.

Bhanu, N.V., van Soolingen, D., van Embden, J.D., Dar, L., Pandey, R.M., Seth, P., 2002. Predominace of a novel *Mycobacterium tuberculosis* genotype in the Delhi region of India. Tuberculosis 82, 105–112.

Bhatia, A.L., 1961. Characteristics of Indian tubercle bacilli. Ind. J. Chest Dis. 3, 147–153.

Brewer, T.F., Colditz, G.A., 1995. Relationship between bacille Calmette-Guerin (BCG) strains and the efficacy of BCG vaccine in the prevention of tuberculosis. Clin. Infect. Dis. 20, 126–135.

Brosch, R., Gordon, S.V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., Pym, A.S., Samper, S., van Soolingen, D., Cole, S.T., 2002. A new evolutionary scenario for the *Mycobaterium tuberculosis* complex. Proc. Natl. Acad. Sci. U.S.A. 99, 3684–3689.

Brudey, K., Filliol, I., Sola, C., Bebear, C., Elia-Pasquet, S., Texier-Maugein, J., Rastogi, N., 2003. Molecular characterization and biodiversity of *Mycobacterium tuberculosis* in the Antilles-Guiana region and comparative analysis in a metropolitan region, Aquitaine. Pathol. Biol. 51, 282–289.

Brudey, K., Driscoll, J.R., Rigouts, L., Prodinger, W.M., Gori, A., Al-Hajoj, S.A., Allix, C., Aristimuno, L., Arora, J., Baumanis, V., Binder, L., Cafrune, P., Cataldi, A., Cheong, S., Diel, R., Ellermeier, C., Evans, J.T., Fauville-Dufaux, M., Ferdinand, S., Garcia de Viedma, D., Garzelli, C., Gazzola, L., Gomes, H.M., Guttierez, M.C., Hawkey, P.M., van Helden, P.D., Kadival,

G.V., Kreiswirth, B.N., Kremer, K., Kubin, M., Kulkarni, S.P., Liens, B., Lillebaek, T., Ho, M.L., Martin, C., Martin, C., Mokrousov, I., Narvskaia, O., Ngeow, Y.F., Naumann, L., Niemann, S., Parwati, I., Rahim, Z., Rasolofo-Razanamparany, V., Rasolonavalona, T., Rossetti, M.L., Rusch-Gerdes, S., Sajduda, A., Samper, S., Shemyakin, I., Singh, U., Somoskovi, A., Skuce, R., van Soolingen, D., Streicher, E., Suffys, P.N., Tortoli, E., Tracevska, T., Vincent, V., Victor, T.C., Warren, R., Yap, S.F., Zaman, K., Portaels, F., Rastogi, N., Sola, C., 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC Microbiol. 6, 6–23.

Chen, I., 2006. Thinking big about global health. Cell 124, 661–663.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E., Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.-A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S.C., Barrell, B.G., 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393, 537–544.

Constant, P., Perez, E., Malaga, W., Laneelle, M.A., Saurel, O., Daffe, M., Guilhot, C., 2002. Role of the pks15/1 gene in the biosynthesis of phenolglycolipids in the *Mycobacterium tuberculosis* complex. Evidence that all strains synthesize glycosylated p-hydroxybenzoic methly esters and that strains devoid of phenolglycolipids harbor a frameshift mutation in the pks15/1 gene. J. Biol. Chem. 277, 38148–38158.

Cowan, L.S., Diem, L., Brake, M.C., Crawford, J.T., 2004. Transfer of a *Mycobacterium tuberculosis* genotyping method, Spoligotyping, from a reverse line-blot hybridization, membrane-based assay to the Luminex multianalyte profiling system. J. Clin. Microbiol. 42, 474–477.

Dale, J.W., Nor, R.M., Ramayah, S., Tang, T.H., Zainuddin, Z.F., 1999. Molecular epidemiology of tuberculosis in Malaysia. J. Clin. Microbiol. 37, 1265–1268.

Dale, J.W., Brittain, D., Cataldi, A.A., Cousins, D., Crawford, J.T., Driscoll, J., Heersma, H., Lillebaek, T., Quitugua, T., Rastogi, N., Skuce, R.A., Sola, C., van Soolingen, D., Vincent, V., 2001. Spacer oligonucleotide typing of bacteria of the *Mycobacterium tuberculosis* complex: recommendations for standardised nomenclature. Int. J. Tuberc. Lung Dis. 5, 216–219.

Das, S., Paramasivan, C.N., Lowrie, D.B., Prabhakar, R., Narayanan, P.R., 1995. IS6110 restriction fragment length polymorphism typing of clinical isolates of *Mycobacterium tuberculosis* from patients with pulmonary tuberculosis in Madras, South India. Tuberc. Lung Dis. 76, 550–554.

Das, S.D., Narayanan, S., Hari, L., Hoti, S.L., Thangathurai, R.K., Charles, N., Jaggarajamma, K., Narayanan, P.R., 2005. Differentiation of highly prevalent IS6110 single-copy strains of *Mycobacterium tuberculosis* from a rural community in South India with an ongoing DOTS programme. Infect. Genet. Evol. 5, 67–77.

Filliol, I., Driscoll, J.R., Van Soolingen, D., Kreiswirth, B.N., Kremer, K., Valetudie, G., Anh, D.D., Barlow, R., Banerjee, D., Bifani, P.J., Brudey, K., Cataldi, A., Cooksey, R.C., Cousins, D.V., Dale, J.W., Dellagostin, O.A., Drobniewski, F., Engelmann, G., Ferdinand, S., Gascoyne-Binzi, D., Gordon, M., Gutierrez, M.C., Haas, W.H., Heersma, H., Kallenius, G., Kassa-Kelembho, E., Koivula, T., Ly, H.M., Makristathis, A., Mammina, C., Martin, G., Mostrom, P., Mokrousov, I., Narbonne, V., Narvskaya, O., Nastasi, A., Niobe-Eyangoh, S.N., Pape, J.W., Rasolofo-Razanamparany, V., Ridell, M., Rossetti, M.L., Stauffer, F., Suffys, P.N., Takiff, H., Texier-Maugein, J., Vincent, V., De Waard, J.H., Sola, C., Rastogi, N., 2002. Global distribution of *Mycobacterium tuberculosis* spoligotypes. Emerg. Infect. Dis. 8, 1347–1349.

Filliol, I., Driscoll, J.R., van Soolingen, D., Kreiswirth, B.N., Kremer, K., Valetudie, G., Dang, D.A., Barlow, R., Banerjee, D., Bifani, P.J., Brudey, K., Cataldi, A., Cooksey, R.C., Cousins, D.V., Dale, J.W., Dellagostin, O.A., Drobniewski, F., Engelmann, G., Ferdinand, S., Gascoyne-Binzi, D., Gordon, M., Gutierrez, M.C., Haas, W.H., Heersma, H., Kassa-Kelembho, E., Ho, M.L., Makristathis, A., Mammina, C., Martin, G., Mostrom, P., Mokrousov, I., Narbonne, V., Narvskaya, O., Nastasi, A., Niobe-Eyangoh, S.N., Pape, J.W., Rasolofo-Razanamparany, V., Ridell, M., Rossetti, M.L., Stauffer, F., Suffys, P.N., Takiff, H., Texier-Maugein, J., Vincent, V., de Waard, J.H., Sola, C., Rastogi, N., 2003. Snapshot of moving and expanding clones of *Mycobacterium tuberculosis* and their global distribution assessed by spoligotyping in an international study. J. Clin. Microbiol. 41, 1963–1970.

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B.C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., Hilty, M., Hopewell, P.C., Small, P.M., 2006. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. U.S.A. 103, 2869–2873.

Gingeras, T.R., Ghandour, G., Wang, E., Berno, A., Small, P.M., Drobniewski, F., Alland, D., Desmond, E., Holodniy, M., Drenkow, J., 1998. Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic Mycobacterium DNA arrays. Genome Res. 8, 435–448.

Glynn, J.R., Whiteley, J., Bifani, P.J., Kremer, K., van Soolingen, D., 2002. Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. Emerg. Infect. Dis. 8, 843–849.

Grange, J.M., Aber, V.R., Allen, B.W., Mitchison, D.A., Goren, M.B., 1978. The correlation of bacteriophage types of *Mycobacterium tuberculosis* with guinea pig virulence and *in vitro* indicators of virulence. Tubercle 45, 354–359.

Guitierrez, M.C., Ahmed, N., Willery, E., Narayanan, S., Hasnain, S.E., Chauhan, D.S., Katoch, V.M., Vincent, V., Locht, C., Supply, P., 2006. Predominance of ancestral lineages of *Mycobacteirum tuberculosis* in India. Emerg. Infect. Dis. 12, 1367–1374.

Gutacker, M.M., Smoot, J.C., Migliaccio, C.A., Ricklefs, S.M., Hua, S., Cousins, D.V., Graviss, E.A., Shashkina, E., Kreiswirth, B.N., Musser, J.M., 2002. Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: resolution of genetic relationships among closely related microbial strains. Genetics 162, 1533–1543.

Gutacker, M.M., Mathema, B., Soini, H., Shashkina, E., Kreiswirth, B.N., Graviss, E.A., Musser, J.M., 2006. Single-nucleotide polymorphism-based population genetic analysis of *Mycobacterium tuberculosis* strains from 4 geographic sites. J Infect Dis. 193, 121–128.

Hermans, P.W., Messadi, F., Guebrexabher, H., van Soolingen, D., de Haas, P.E., Heersma, H., de Neeling, H., Ayoub, A., Portaels, F., Frommel, D., 1995. Analysis of the population structure of *Mycobacterium tuberculosis* in Ethiopia, Tunisia, and The Netherlands: usefulness of DNA typing for global tuberculosis epidemiology. J. Infect. Dis. 171, 1504–1513.

Joseph, S., Mitchison, D.A., Ramachandran, K., Selkon, J.B., Subbaiah, T.V., 1964. Virulence in the guinea-pig and sensitivity to pas and thiacetazone of tubercle bacilli from south Indian patients with pulmonary tuberculosis. Tubercle 45, 354–359.

Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R., Goyal, M., van Embden, J., 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J. Clin. Microbiol. 35, 907–914.

Kato-Maeda, M., Rhee, J.T., Gingeras, T.R., Salamon, H., Drenkow, J., Smittipat, N., Small, P., 2001. Comparing genomes within the species *Mycobacterium tuberculosis*. Genome Res. 11, 1796–1799.

Kremer, K., van Soolingen, D., Frothingham, R., Haas, W.H., Hermans, P.W., Martin, C., Palittapongarnpim, P., Plikaytis, B.B., Riley, L.W., Yakrus, M.A., Musser, J.M., van Embden, J.D., 1999. Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. J. Clin. Microbiol. 37, 2607–2618.

Kulkarni, S., Sola, C., Filliol, I., Rastogi, N., Kadival, G., 2005. Spoligotyping of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in Mumbai, India. Res. Microbiol. 156, 588–596.

Marmiesse, M., Brodin, P., Buchrieser, C., Gutierrez, C., Simoes, N., Vincent, V., Glaser, P., Cole, S.T., Brosch, R., 2004. Macro-array and bioinformatic analyses reveal mycobacterial 'core' genes, variation in the ESAT-6 gene family and new phylogenetic markers for the *Mycobacterium tuberculosis* complex. Microbiology 150, 483–496.

Mbelle, N., Huebner, R.E., Wasas, A.D., Kimura, A., Chang, I., Klugman, K.P., 1999. Immunogenicity and impact on nasopharyngeal carriage of a non-avalent pneumococcal conjugate vaccine. J. Infect. Dis. 180, 1171–1176.

Middlebrook, G., Cohn, M.L., 1953. Some observations on the pathogenicity of isoniazid-resistant variants of tubercle bacilli. Science 118, 297–299.

Mistry, N.F., Iyer, A.M., D'souza, D.T., Taylor, G.M., Young, D.B., Antia, N.H., 2002. Spoligotyping of *Mycobacterium tuberculosis* isolates from multiple-drug-resistant tuberculosis patients from Bombay, India. J. Clin. Microbiol. 40, 2677–2680.

Mitchison, D.A., Wallace, J.G., Bhatia, A.L., Selkon, J.B., Subbaiah, T.V., Lancaster, M.C., 1960. A comparison of the virulence in guinea-pigs of South Indian and British tubercle bacilli. Tubercle 41, 1–22.

Narayanan, S., Sahadevan, R., Narayanan, P.R., Krishnamurthy, P.V., Paramasivan, C.N., Prabhakar, R., 1997. Restriction fragment length polymorphism of *Mycobacterium tuberculosis* strains from various regions of India, using direct repeat probe. Indian J. Med. Res. 106, 447–454.

Narayanan, S., Das, S., Garg, R., Hari, L., Rao, V.B., Frieden, T.R., Narayanan, P.R., 2002. Molecular epidemiology of tuberculosis in a rural area of high prevalence in South India: implications for disease control and prevention. J. Clin. Microbiol. 40, 4785–4788.

Prabhakar, R., Venkataraman, P., Vallishayee, R.S., Reeser, P., Musa, S., Hashim, R., Kim, Y., Dimmer, C., Wiegeshaus, E., Edwards, M.L., 1987. Virulence for guinea pigs of tubercle bacilli isolated from the sputum of participants in the BCG trial, Chingleput District, South India. Tubercle 68, 3–17.

Puusteinen, K., Marjamaki, M., Rastogi, N., Sola, C., Filliol, I., Ruutu, P., Holmstrom, P., Viljanen, M.K., Soini, J., 2003. Characterization of Finnish *Mycobaterium tuberculosis* isolates by spoligotyping. J. Clin. Microbiol. 41, 1525–1528.

Radhakrishnan, I., Manju, Y.K., Kumar, R.A., Mundayoor, S., Harris, 2001. Implications of low frequency of IS6110 in fingerprinting field isolates of *Mycobacterium tuberculosis* from Kerala, India. J. Clin. Microbiol. 39, 1683.

Singh, B., Sharma, M.D., 1964. Niacin production and virulence in the guinea-pig of Indian tubercle bacilli. Tubercle 45, 239–240.

Singh, U.B., Suresh, N., Bhanu, N.V., Arora, J., Pant, H., Sinha, S., Aggarwal, R.C., Singh, S., Pande, J.N., Sola, C., Rastogi, N., Seth, P., 2004. Predominant tuberculosis spoligotypes, Delhi India. Emerg. Infect. Dis. 10, 1138–1142.

Singh, U.B., Arora, J., Suresh, N., Pant, H., Rana, T., Sola, C., Rastogi, N., Pande, J.N., 2007. Genetic biodiversity of *Mycobacterium tuberculosis* isolates from patients with pulmonary tuberculosis in India. Infect. Genet. Evol. 7, 441–448.

Sola, C., Filliol, E., Legrand, E., Mokrousov, I., Rastogi, N., 2001. *Mycobacterium tuberculosis* phylogeny reconstruction based on combined numerical analysis with IS1081, IS6110, VNTR and DR-based spoligotyping suggests the existence of two new phylogeographical clades. J. Mol. Evol. 53, 680–689.

Sreevatsan, S., Pan, X., Stockbauer, K.E., Connell, N.D., Kreiswirth, B.N., Whittam, T.S., Musser, J.M., 1997. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. Proc. Natl. Acad. Sci. U.S.A. 94, 9869–9874.

Streicher, E.M., Warren, R.M., Kewley, C., Simpson, J., Rastogi, N., Sola, C., van der Spuy, G.D., van Helden, P.D., Victor, T.C., 2004. Genotypic and phenotypic characterization of drug-resistant *Mycobacterium tuberculosis* isolates from rural districts of the Western Cape Province of South Africa. J. Clin. Microbiol. 42, 891–894.

Sun, Y.J., Lee, A.S., Ng, S.T., Ravindran, S., Kremer, K., Bellamy, R., Wong, S.Y., van Soolingen, D., Supply, P., Paton, N.I., 2004. Characterization of ancestral *Mycobacterium tuberculosis* by multiple genetic markers and proposal of genotyping strategy. J. Clin. Microbiol. 42, 5058–5064.

Tuberculosis Research Centre (ICMR) Chennai, 1999. Fifteen-year follow up of trial of BCG vaccines in south India for tuberculosis prevention. Indian J. Med. Res. 110, 56–69.

Tsolaki, A.G., Hirsh, A.E., DeRiemer, K., Enciso, J.A., Wong, M.Z., Hannan, M., Goguet de la Salmoniere, Y.-O.L., Aman, K., Kato-Maeda, M., Small, P.M., 2004. Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains. Proc. Natl. Acad. Sci. U.S.A. 101, 4865–4870.

Tsolaki, A.G., Gagneux, S., Pym, A.S., Goguet de la Salmoniere, Y.-O.L., Kreiswirth, B.N., van Soolingen, D., Small, P.M., 2005. Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of *Mycobacterium tuberculosis*. J. Clin. Microbiol. 43, 3185–3191.

van Embden, J.D., Cave, M.D., Crawford, J.T., Dale, J.W., Eisenach, K.D., Gicquel, B., Hermans, P., Martin, C., McAdam, R., Shinnick, T.M., 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J. Clin. Microbiol. 31, 406–409.

Van Loo, I.H., Mooi, F.R., 2002. Changes in the Dutch Bordetella pertussis population in the first 20 years after the introduction of whole-cell vaccines. Microbiol. 148, 2011–2018.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 484–488
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Genetic diversity of *Trypanosoma evansi* in beef cattle based on internal transcribed spacer region☆

Supatra Areekit, Pirom Singhaphan, Pornpimon Kanjanavas, Sintawee Khuchareontaworn, Thayat Sriyapai, Arda Pakpitcharoen, Kosum Chansiri *

Department of Biochemistry, Faculty of Medicine, Srinakharinwirot University, Sukhumvit 23, Bangkok 10110, Thailand

Received 17 October 2007; accepted 17 October 2007
Available online 1 November 2007

Abstract

This study was focused on genetic diversity of *Trypanosoma evansi* which is a widely distributed haemoflagellate of veterinary importance that infects a variety of larger mammals including horses, mules, camels, buffalo, cattle and deer. The genetic diversity of *T. evansi* of beef cattle LAM19 was accomplished by using phylogenetic analysis based on internal transcribed spacer region (ITS). Blood sample was collected from a naturally infected beef cattle LAM 19 and parasitemia was raised by mouse inoculation. The parasites were collected and isolated by using DE 52 DEAE cellulose anion exchange column prior to DNA extraction. Upon PCR amplification of ITS region, the product of 1300 bp in size was obtained. The ITS nucleotide sequences were analyzed and revealed that it could demonstrate the genetic diversity of *T. evansi* of beef cattle LAM19. Based on the ITS tree, beef cattle LAM 19 *T. evansi* were categorized into two main groups where the genetic diversity occurred within Group 1. The data could be applicable for the survey of parasite dynamics, epidemiological studies as well as prevention and control of the disease.

Keywords: Internal transcribed spacer; ITS; Cattle

1. Introduction

Trypanosomes are protozoan parasites that cause major disease in human and other animals. Trypanosomes (genus *Trypanosoma*) are widespread blood parasites of vertebrates, usually transmitted by arthropod or leech vectors. Several *Trypanosome* species are agents of disease in humans and/or livestock particularly in the tropics. In Thailand, *T. evansi* are widespread where they cause considerable loss in productivity in domestic animals, such as horses (Boonyawong et al., 1975), beef cattle (Chaichanapunpol et al., 1985; Tuntasuvan et al., 1997), dairy cattle (Trisanarom et al., 1987).

The finding that trypanosomes exhibit considerable intra-species diversity has led to the proposal that these parasites have a clonal population structure (Tibayrenc et al., 1990; Tibayrenc and Ayala, 1991; Tibayrenc et al., 1991a,b). The clonal theory has important implications for the diagnosis and treatment of trypanosomiasis and for the development of anti-trypanosome drugs and vaccine. Genetically diverse clonets are likely to differ in their pathological and other biological properties. In fact, there is evidence for these in *T. cruzi* (Morel et al., 1980; Engman et al., 1987; Aymerich and Goldenberg, 1989), *T. brucei* (Stevens and Godfrey, 1992) and *T. congolense* (Majiwa and Webster, 1987). It is clear that much work remains to be done to understand the molecular basis for these differences. It will be particularly important to correlate molecular variation at the gene level with changes in pathophysiological properties of the parasites.

The internal transcribed spacer (ITS region) has been used for phylogenetic analysis, evaluation of the evolutionary process, as well as for determination of taxonomic identities. It contains two regions (ITS1 and ITS2) that are located between the repeating array of nuclear 18S, 5.8S and 28S ribosomal RNA genes (Lodish et al., 2001), a locus that has 100–200 copies per genome. Beltrame-Botelho and co-worker described the sequences of both ITS spacers flanking the *T. rangeli* and 5.8S rDNA. These sequences could be used to identify both intra- and inter-specific variability. Most of the data indicated that the ITS-2 spacer is

☆ GenBank accession numbers of ITS regions of *T. evansi* isolated from beef cattle LAM19 were EF545993-EF546014. The accession numbers of ITS regions of other organisms used in this study were indicated in the text.

* Corresponding author. Tel.: +66 2260 2122x4605; fax: +66 2260 0125.
E-mail address: kosum@swu.ac.th (K. Chansiri).

1567-1348/$ – see front matter
doi:10.1016/j.meegid.2007.10.008

more variable than the ITS-1 spacer and the 5.8S rDNA when performing inter-specific comparisons among the studied trypanosomatids (Beltrame-Botelho et al., 2005). In 1999, the phylogenetic analysis of ITS sequences was used for studying the polymorphism of *T. rangeli* strains isolated from different hosts and geographic areas (southern Brazil, Central America, and northern South America) (Grisard et al., 1999). In 2004, Sarataphan and colleagues have reported the 4 group-specific genotypes within *T. evansi* isolates from six different hosts (elephant, horse, buffalo, cattle, pig and deer) based on PCR-RFLP of the open reading frame 2 (ORF2) of glucose transporter genes located on chromosome X of *T. brucei* (Sarataphan et al., 2004). However, the data were not sufficient for demonstration of the genetic diversity and relationship of *T. evansi* within each single host. In 2007, Khuchareontaworn and colleagues demonstrated the genetic diversity of *T. evansi* isolated from a buffalo based on ITS 1 and ITS2 regions. They reported that ITS 2 tree was able to classify *T. evansi* into 4 groups whereas ITS 1 tree showed lower efficacy in discrimination (Khuchareontaworn et al., 2007).

Here in, ITS region of *T. evansi* from beef cattle was analyzed and the genetic diversity was examined. The phylogenetic trees were constructed to support the evidence of genetic diversity of the parasite.

2. Materials and methods

2.1. Parasite collection

A naturally *T. evansi* infected blood sample from beef cattle namely LAM19 was collected during 1999 from a farm in Lampraya Klang district, Saraburi province, Thailand. The infection was examined by using thin blood smear technique. The 3.0 ml of infected blood were collected from cattle and preserved in the phosphate saline glucose buffer (PSG), pH 8.0 (50 mM $Na_2HPO_4{\cdot}H_2O$, 2 mM $NaH_2PO_4{\cdot}2H_2O$, 36 mM NaCl and 1.5% glucose) and stored in the liquid nitrogen tank (−196 °C) until use.

2.2. Mouse inoculation and parasite isolation

Infected blood in the liquid nitrogen tank was thawed at room temperature and 0.1 ml was injected to a mouse using 1 ml Tuberculin syringe and needle gauge No 27. Parasitemia was daily checked by wet blood smear, starting from the third day after injection. At the highest parasitemia (10^8 cells/ml) which was usually in the 3rd to 5th day of injection, blood was withdrawn from euthanized mouse by cardiac punctured. An anion exchange column (DE 52 DEAE cellulose) was used to purify parasite from the horde blood cells according to the method of described by Chao (Choa et al., 1984). The eluent with infect protozoa was collected and proceeding to parasite lysis and DNA extraction.

2.3. DNA extraction and purification

Pellet of parasites was resuspended in PSG buffer in the presence of SDS (final concentration was 2%) and proteinase K (final concentration was 1 mg/ml). The solution was incubated at 42 °C for 14 h. Parasite DNA was extracted by conventional phenol/chloroform which responded for deproteinization of the aqueous solution containing the desired nucleic acid. The purified DNAs were precipitated by the addition of 2 volumes of cold absolute ethanol. The pellet was dried, dissolved in sterile distilled water and kept at 4 °C until use.

2.4. PCR amplification

PCR amplifications of purified DNA from blood samples were performed by a pair of primers used for amplification of ITS regions including 5.8S rDNA were ITS/F, forward primer, 5′ GGT GAT CGG ACC GTC GCT CGT CT 3′ and reverse primer ITS/R, 5′ CCT CTT CGC TCG CCG CTG ACT G 3′. The primers were designed based on the nucleotide sequence comparison of *T. evansi* ITS1 and ITS2 (Accession number D89527) in GenBank database. All reactions were manipulated in 25 μl volume containing 50 ng of genomic DNA samples. The PCR amplification of ITS region will contain 1X PCR buffer, 2 mM of each primer, 400 mM of dNTP, 6 mM $MgCl_2$ and 1.5 units of proof reading *Taq* DNA polymerase (Gibco BRL®). Sterile distilled water was used to make volume to 25 μl. A PCR cycle was performed using Peltier Thermal Cycler (MJ Research, PTC-200) for 30 cycles. Each cycle consisted of denaturation at 94 °C for 1 min, annealing at 60 °C for 1 min and extension at 72 °C for 1 min. PCR products were analyzed by using electrophoresis in 1.2% agarose gel at 110 V approximately 45 min prior to staining by ethidium bromide and the gel was observed under ultraviolet light.

2.5. Cloning and sequencing of PCR fragment

The PCR fragment was eluted from the gel and purified by QIAGEN DNA purification kit (QIAGEN®). The purified 200 ng of PCR fragment were further ligated to the pGEM® – T Easy vector by using protocol of pGEM® – T Easy Vector Systems (Invitrogen, Promega). The reaction was consisted of 4 μl of PCR product, 1 μl of 2X Rapid Ligation Buffer (60 mM Tris–HCl, pH 7.8, 20 mM $MgCl_2$, 20 mM ATP 10% PEG), 1 μl of the pGEM® – T Easy vector and 3 units of T4 DNA ligase. The reaction was mixed, and incubated for 24 h at 4 °C prior to transformation by electroporation using BioRad MicroPulser (BIO-RAD MicroPulser™). The transformants were then spread on the LB-ampicillin agar plates containing X-gal and IPTG (40 μl of 20 mg/ml X-gal and 20 μl of 200 mg/ml IPTG) and incubated for 16 h at 37 °C. White colonies were randomly selected and DNAs were purified by using QIAGEN plasmid purification kit (QIAGEN®). DNA sequencing was performed in both forward and reverse direction based on the Big Dye Terminator Cycle Sequencing procedure (Macrogen). The nucleotide sequence data were analyzed using software of ABI PRISM Model 3100 version 3.7 and BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) in comparison to ITS regions of the other *T. evansi* that have been previously reported in GenBank.

Table 1
GenBank accession numbers of ITS regions of *Trypanosoma* spp.

Parasites names	Specific host/lab host	Origin	GenBank accession numbers
T. evansi clone No.1.1,3.1	Beef cattle	Saraburi, Thailand	EF546003—EF546004
T. evansi clone No.3	Beef cattle	Saraburi, Thailand	EF545993
T. evansi clone No.4	Beef cattle	Saraburi, Thailand	EF545996
T. evansi clone No.6	Beef cattle	Saraburi, Thailand	EF546005
T. evansi clone No.8	Beef cattle	Saraburi, Thailand	EF545994
T. evansi clone No.8.1,11,12.1	Beef cattle	Saraburi, Thailand	EF546006—EF546008
T. evansi clone No.12	Beef cattle	Saraburi, Thailand	EF545997
T. evansi clone No.13	Beef cattle	Saraburi, Thailand	EF546009
T. evansi clone No.16	Beef cattle	Saraburi, Thailand	EF545998
T. evansi clone No.16.1,17,24,27	Beef cattle	Saraburi, Thailand	EF546010—EF546013
T. evansi clone No.28.1	Beef cattle	Saraburi, Thailand	EF546014
T. evansi clone No.29,38,42	Beef cattle	Saraburi, Thailand	EF546000—EF546002
T. brucei Da1972	*Homo sapiens*	Daloa, Ivory Coast	AF306774
T. brucei TH2	*Homo sapiens*	Koudougou, Ivory Coast	AF306777
T. brucei KP2	*Glossina palpalis*	Kouassi-Perita, Ivory Coast	AF306773
T. brucei NW2	*Homo sapiens*	Uganda	AF306776
T. brucei STIB215	Lion	Serengeti, Tanzania	AF306771
T. brucei B8/18	Pig	Nsukka, Nigeria	AF306772
T. brucei H3	Lion	Luangwa Valley, Zambia	AF306770
T. brucei Strain 427	Unknown	–	X05682
T. evansi Strain Tansui-Taiwan	Unknown	Taiwan	D89527
T. congolense Strain riverine/forest-type	*Rattus norvegicus*	–	U22319

2.6. Phylogenetic analysis

The 21 white colonies containing PCR products of ITS region (1300 bp) amplified from beef cattle LAM 19 were selected, sequenced and analyzed. Three sets of nucleotide sequence data, ITS1 (341 bp), ITS2 (587 bp) and ITS (ITS1+ITS2+5.8S rDNA), were separately aligned against those of *T. brucei* and other reported *T. evansi* by using Clustal X software Version 1.83 (multiple sequence alignment) program (Thompson et al., 1997) and refined using the manual method (Table 1). The phylogenetic trees were constructed according to PAUP program version 4.0 (Swofford, 2002) using *T. congolense* as the outgroup. The cladograms were created by means of maximum likelihood (complete ITS region) and maximum parsimony (ITS1 and ITS2) methods. Bootstrap values were replicated 1000 times and computed with PAUP program version 4.0.

3. Results

Upon PCR amplification of *T. evansi* genomic DNA of beef cattle LAM 19, the product of 1300 bp in size was obtained. After ligation and transformation processes of this PCR product, 21 positive white colonies were randomly selected. BLAST data of this ITS nucleotides sequences revealed that they were corresponded to those of *T. evansi* Taiwan stock (D89527.1) with the homology of 91%.

Initially, the individual two data sets of ITS 1 (341 bp) and ITS 2 (587 bp) regions were employed for phylogenetic tree construction in comparison to those of *T. evansi* and *Trypanosoma* spp. from GenBank. However, both trees showed low bootstrap values among clades that were significantly unable to discriminate the genetic diversity of ITS regions in *T. evansi* of beef cattle LAM 19 (data not shown). On contrary, the phylogenetic tree inferred from the complete ITS nucleotide sequences (1300 bp) clearly showed the genetic diversity of the parasites. According to the tree, two main groups (Group 1 and Group 2) of ITS regions in *T. evansi* of beef cattle LAM 19 were clearly separate with bootstrap value 89% (Fig. 1). Group 1 was further divided into 6 subgroups, subgroup 1a–1f. Subgroup 1a contained ITS regions of *T. brucei* (X05682) and *T. brucei* of lions (AF306770 and AF306771) with bootstrap value 61%. Subgroup 1b, 1c, 1e and 1f consisted of 5, 5, 2 and 3 ITS clones of *T. evansi* of beef cattle LAM 19 with bootstrap value 84, 70, 91 and 89%, respectively. Subgroup 1d included ITS regions of *T. brucei* of *Homo sapiens* (AF306774, AF306776 and AF306777), pig (AF306772), *Glossina palpalis* (AF306773), and one ITS clone of *T. evansi* of beef cattle LAM 19 with bootstrap value 81%. Finally, Group 2 was categorized in the same clade as ITS of *T. congolense* encompassing two ITS clones of *T. evansi* of beef cattle LAM 19 with bootstrap value 59%.

4. Discussion

Previously, Khuchareontaworn and colleques (Khuchareontaworn et al., 2007) suggested that ITS 2 region was informative for genetic diversity exploitation in water buffalo B18 (four groups were detected). However, this study revealed that ITS 2 tree inferring from nucleotide sequences of 341 bp failed to elucidate the genetic diversity of *T. evansi* from beef cattle LAM19 since the tree formed the low bootstrap value (~50%) indicating the high feasibility of tree collapse. Similarly, the ITS 1 nucleotide sequences (587 bp) could not support the tree and also showed the low bootstrap values at 50–55%. Hence, the tree construction based on merely either ITS 1 or ITS2 data could be affected by their limited nucleotide sequence length as

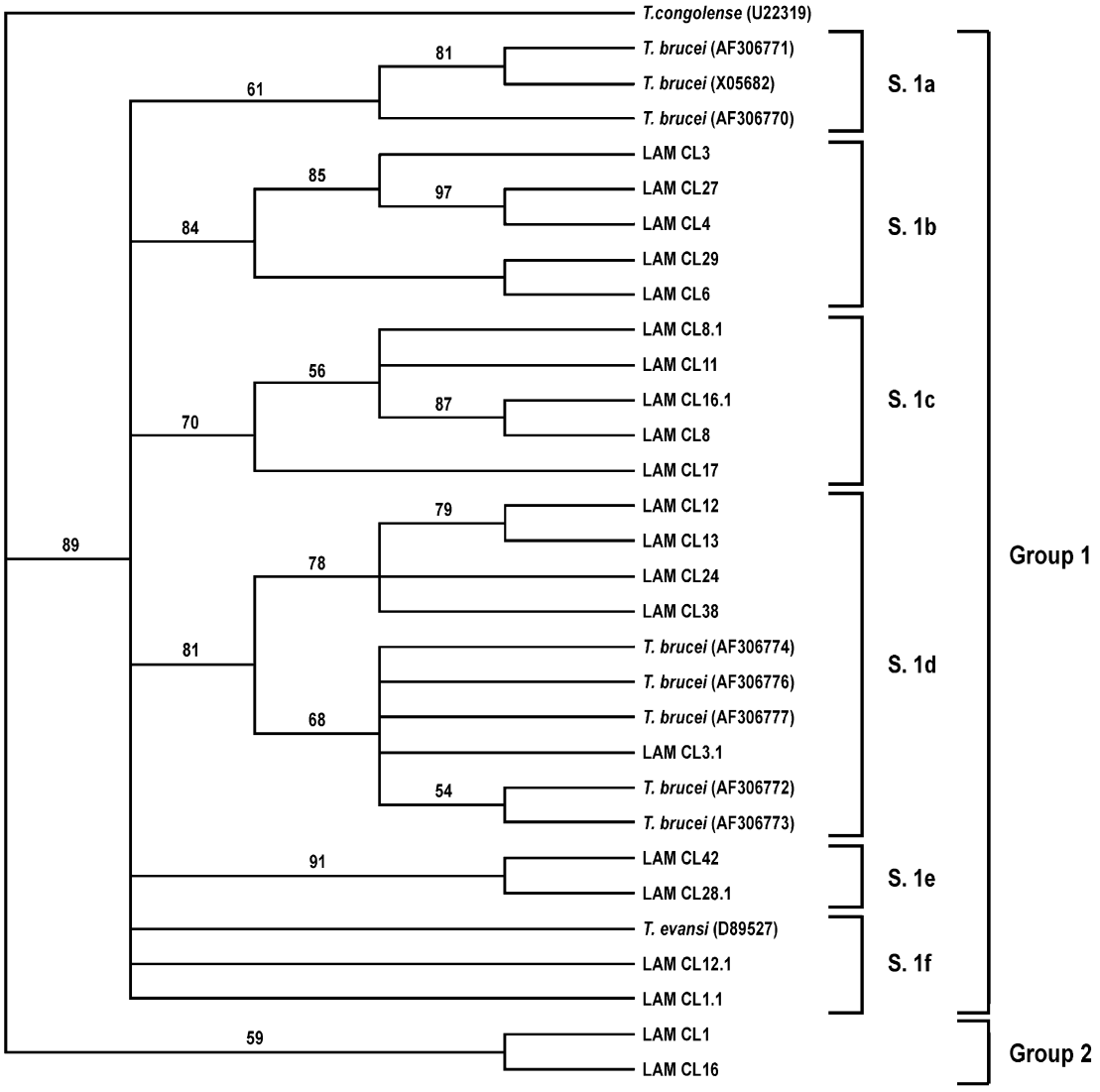


Fig. 1. A maximum parsimony tree inferring from ITS regions of *Trypanosoma* spp. Twenty-one ITS clones of *T. evansi* of beef cattle LAM 19 were analyzed against those of *T. brucei* and *T. congolense*. The tree was produced by the heuristic search of PAUP version 4.0 using a bootstrap resembling of the data set. The numbers at the nodes represent the percentage of times the group occurred out of 1000 trees. Branches with less than 50% bootstrap support were collapsed. No significance is placed on the lengths of the branches connecting the species.

well as the high similarity among isolates. Nevertheless, the tree was improved when both ITS 1 and ITS 2 including 5.8 S rDNA data were complied.

According to the failure of ITS 1 and ITS 2 trees in genetic discrimination of *Trypanosoma* spp. including *T. evansi*, our alternative rational have also been focused on the different exposure of water buffalo and cattle to vectors. Referring to our sample collection, beef cattle LAM19 was obtained from a certain hygiene Livestock whereas water buffalo B18 was taken from domestic area. Hence, it is possible that beef cattle LAM 19 may have less exposure to vectors.

Based on the ITS tree, *Trypanosome* spp. were categorized into two main groups. In fact, the genetic diversity was occurred within Group 1 where seven subgroups were generated including one subgroup (S.1a) that contained only *T. brucei* from GenBank. This result supported the genetic diversity within *T. evansi* that have been previously reported by Sarataphan et al. (2004) and Khuchareontaworn et al. (2007). Since ITS of *T. evansi* has approximately 100–200 copies spreading out through the genome, it is more likely to be genetic diversity rather than mix infection.

In conclusion, genetic diversity of *T. evansi* of beef cattle LAM 19 has been demonstrated. The results could be employed as the preliminary data for studying genetic diversity, molecular epidemiology, and the dynamic of *T. evansi*. In addition, it could be applicable for the prevention and control of the disease.

Acknowledgement

Blood sample of naturally *T. evansi* infected beef cattle LAM19 was gifted from Dr. Nopporn Sarataphan, National Institute of Animal Health, Bangkok, Thailand.

References

Aymerich, S., Goldenberg, S., 1989. The karyotype of *Trypanosoma cruzi* Dm 28c: comparison with other *T. cruzi* strains and trypanosomatids. Exp. Parasitol. 69 (2), 107–115.

Beltrame-Botelho, I.T., Gaspar-Silva, D., Steindel, M., Davila, A.M., Grisard, E.C., 2005. Internal transcribed spacers (ITS) of *Trypanosoma rangeli* ribosomal DNA (rDNA): a useful marker for inter-specific differentiation. Infect. Genet. Evol. 5 (1), 17–28.

Boonyawong, T., Chantaraprateep, P., Muangyai, M., 1975. Surra in horse. Thai J. Vet. Med. 5, 665–672.

Chaichanapunpol, I., Sisrrangkamanont, T., Trongwongsa, L., Suvanwasi, P., 1985. Observations on trypanosomiasis in a native bull. Kasetsart Vet. 6, 1–9.

Chao, D., Dusanic, D.G., 1984. Comparative studies of the isolation of metacyclic trypomastigotes of Trypanosoma cruzi by DEAE ion exchange chromatography. Zhonghua Min Guo Wei Sheng Wu Ji Mian Yi Xue Za Zhi 17, 146–152.

Engman, D.M., Reddy, L.V., Donelson, J.E., Kirchhoff, L.V., 1987. *Trypanosoma cruzi* exhibits inter- and intra-strain heterogeneity in molecular karyotype and chromosomal gene location. Mol. Biochem. Parasitol. 22 (2–3), 115–123.

Grisard, E.C., Campbell, D.A., Romanha, A.J., 1999. Mini-exon gene sequence polymorphism among *Trypanosoma rangeli* strains isolated from distinct geographical regions. Parasitology 118, 375–382.

Khuchareontaworn, S., Singhaphan, P., Viseshakul, N., Chansiri, K., 2007. Genetic diversity of *Trypanosoma evansi* in buffalo based on internal transcribed spacer regions. J. Vet. Med. Sci. 69 (5), 487–493.

Lodish, H., Berk, A., Zipursky, S.L., Mastsudaira, P., Balimore, D., Darnell, J., 2001. Molecular cell biology, fourth ed. W.H.Freeman and Company, New York.

Majiwa, P.A., Webster, P., 1987. A repetitive deoxyribonucleic acid sequence distinguishes *Trypanosoma simiae* from *T. congolense*. Parasitology 95 (3), 543–558.

Morel, C., Chiari, E., Camargo, E.P., Mattei, D.M., Romanha, A.J., Simpson, L., 1980. Strains and clones of *Trypanosoma cruzi* can be characterized by pattern of restriction endonuclease products of kinetoplast DNA minicircles. Proc. Natl. Acad. Sci. USA 77 (11), 6810–6814.

Sarataphan, N., Boonchit, S., Sirivan, C., Indrakamhang, P., 2004. Genetic diversity of Trypanosoma evansi in Thailand based on a repeated DNA coding sequence marker.In: Livestock Development Proceeding. The Goden Jubilee Museum of Agriculture, Pathumthani.

Stevens, J.R., Godfrey, D.G., 1992. Numerical taxonomy of Trypanozoon based on polymorphisms in a reduced range of enzymes. Parasitology 104 (1), 75–86.

Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), 4.0b10 edn. Sinauer Associates, Sunderland, Massachusetts.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Hoiggins, D.G., 1997. The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25.

Tibayrenc, M., Kjellberg, F., Ayala, F.J., 1990. A clonal theory of parasitic protozoa: the population structures of Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas, and Trypanosoma and their medical and taxonomical consequences. Proc. Natl. Acad. Sci. USA 87 (7), 2414–2418.

Tibayrenc, M., Ayala, F.J., 1991. Towards a population genetics of microorganisms: The clonal theory of parasitic protozoa. Parasitol. Today 7 (9), 228–232.

Tibayrenc, M., Kjellberg, F., Arnaud, J., Oury, B., Breniere, S.F., Darde, M.L., et al., 1991a. Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. Proc. Natl. Acad. Sci. USA 88 (12), 5129–5133.

Tibayrenc, M., Kjellberg, F., Ayala, F.J., 1991b. The clonal theory of parasitic protozoa. BioScience 41, 767–774.

Trisanarom, A., Markmee, S., Ped-Usorn, C., 1987. Trypanosoma evansi infection in Chaing Mai dairy cattle. In: Proceeding of the 6th Annual Veterinary, 18–19 May, Conference of Thai Veterinary Medicine Association, Thailan, pp. 1–12.

Tuntasuvan, D., Sarataphan, N., Nishikawa, H., 1997. Cerebral trypanosomiasis in native cattle. Vet. Parasitol. 73 (3–4), 357–363.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 489–503
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Bionomics, taxonomy, and distribution of the major malaria vector taxa of *Anopheles* subgenus *Cellia* in Southeast Asia: An updated review

S. Manguin [a,*], C. Garros [a,1], I. Dusfour [a,2], R.E. Harbach [b], M. Coosemans [c,d]

[a] *Institut de Recherche pour le Développement (IRD), Centre de Biologie et de Gestion des Populations, Montpellier, France*
[b] *Department of Entomology, The Natural History Museum, Cromwell Road, London SW7 5BD, UK*
[c] *Institute of Tropical Medicine (ITM), Antwerp, Belgium*
[d] *Faculty of Biomedical Sciences, University of Antwerp, Belgium*

Received 27 April 2007; received in revised form 22 November 2007; accepted 23 November 2007
Available online 21 February 2008

Abstract

There is high diversity of *Anopheles* mosquitoes in Southeast Asia and the main vectors of malaria belong to complexes or groups of species that are difficult or impossible to distinguish due to overlapping morphological characteristics. Recent advances in molecular systematics have provided simple and reliable methods for unambiguous species identification. This review summarizes the latest information on the seven taxonomic groups that include principal malaria vectors in Southeast Asia, i.e. the Minimus, Fluviatilis, Culicifacies, Dirus, Leucosphyrus, and Sundaicus Complexes, and the Maculatus Group. Main issues still to be resolved are highlighted. The growing knowledge on malaria vectors in Southeast Asia has implications for vector control programs, the success of which is highly dependant on precise information about the biology and behavior of the vector species. Acquisition of this information, and consequently the application of appropriate, sustainable control measures, depends on our ability to accurately identify the specific vectors.

Keywords: Malaria; Vectors; Minimus Complex; Fluviatilis Complex; Culicifacies Complex; Dirus Complex; Leucosphyrus Complex; Sundaicus Complex; Maculatus Group; Asia

1. Introduction

Southeast Asia, including the Mekong area, encompasses 15 countries[3] that experience a high burden of vector-borne diseases, among which malaria remains the most important. Some 2.5 million cases of malaria are reported annually, but it is estimated that as many as 100 million cases may actually occur in the region each year (WHO, 2007a). This region accounts for 30% of the global malaria morbidity and about 8% of the global mortality, with approximately 26,000 deaths per year (WHO, 2007b).

Major progress in malaria control was achieved during the last decade, especially in Bhutan, Cambodia, Laos, Sri Lanka, Thailand, and Vietnam. However, a high rate of malaria still occurs in hilly forested areas and some coastal foci where it is a fatal disease that is endemic in poor rural areas (Trung et al., 2004).

The epidemiology of vector-borne diseases is strongly linked to the biodiversity of known or potential insect vectors such as *Anopheles* mosquitoes that may transmit malarial pathogens. Nowadays one must consider the whole anopheline community present in an area, instead of focusing on just one vector species. Ecological, demographic, and climatic changes influence the composition of anopheline communities and consequently have an impact on malaria transmission. This is quite true in Asia where the biodiversity and specific richness of *Anopheles* species is high compared to the other regions

* Corresponding author at: Institut de Recherche pour le Développement (IRD), UMR Centre de Biologie et de Gestion des Populations (CBGP), Campus International de Baillarguet, CS 30016, 34988 Montferrier sur Lez Cedex, France. Tel.: +33 499 62 33 27; fax: +33 499 62 33 45.
E-mail address: manguin@mpl.ird.fr (S. Manguin).
[1] Present address: University Catholic of Louvain, Ecology and Biogeography Unit, Place Croix du Sud 4-5, 1348 Louvain-la-Neuve, Belgium.
[2] Present address: Uniformed Services University of the Health Sciences (USUHS), Department of Preventive Medicine and Biometrics, 4301 Jones Bridge Road, Bethesda, MD 20814, USA.
[3] Bangladesh, Bhutan, Cambodia, southern China (Yunnan), DPR Korea, India, Indonesia, Laos, Maldives, Myanmar, Nepal, Sri Lanka, Thailand, Timor-Leste, Vietnam.

1567-1348/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2007.11.004

(Foley et al., 2007), and where important environmental changes, such as deforestation (Walsh et al., 1993) and irrigation system (Klinkenberg et al., 2004), occur at a quick pace.

The main malaria vectors in Southeast Asia belong to species complexes and groups (Harbach, 2004), which include closely related species that are difficult to distinguish morphologically yet often differ in their bionomics. The sympatric occurrence of the vector species complicates our understanding of malaria transmission and epidemiology. A species may play a primary role in one area and a secondary role elsewhere, and the vector status of an individual species may vary in relation to environmental or seasonal changes. Vector control in the region is therefore potentially hampered by the number and complexity of the primary and secondary vector species. This situation dictates that scientists must come together to study anopheline communities as a whole and to integrate the diverse information about vector systems to define appropriate vector and effective control programs.

The purpose of this review is to synthesize the most recent information on the principal malaria vector taxa of genus *Anopheles* subgenus *Cellia* in Southeast Asia and to highlight the main issues still to be resolved.

2. The Minimus Complex (Funestus Group, Myzomyia Series)

The Minimus Complex comprises two formally named species, *An. minimus* (species A) and *An. harrisoni* (species C), and the informally designated *An. minimus* E. Several putative forms of *An. minimus* are mentioned in published literature that are either morphological or chromosomal variants of the genetic species (see the review of Chen et al. (2002)).

Today, the taxonomy of the Minimus Complex is nearly complete. Harbach et al. (2006) designated a neotype to fix the identity of *An. minimus* s.s., *An. harrisoni* was recently described and named by Harbach et al. (2007), and the description and naming of *An. minimus* E is underway (Harbach, personnel communication). Despite with the formal taxonomy, the three species cannot be distinguished based on morphology (Jaichapor et al., 2005; Sungvornyothin et al., 2006a) and their separation from closely related sympatric species is problematic due to overlapping characters. The situation is complicated by the morphological variability of *An. minimus* (Jaichapor et al., 2005).

The application of molecular techniques has made it possible to reliably identify species in entomological surveys (Table 1). A number of molecular identification assays are now

Table 1
Type and references on the PCR assays developed for each complex and associated species

Method[a] (reference)	*An. culicifacies* A, D/B, C/E	*An. fluviatilis* S, T, U	*An. dirus*, *An. cracens*, *An. scanloni*, *An. baimaii*	*An. minimus*, *An. harrisoni*	*An. aconitus*, *An. pampanai*, *An. varuna*[b]	*An. jeyporiensis*[1]	*An. maculatus* group[c,d]	*An. sundaicus*, species E, *An. epiroticus*
AS (Singh et al., 2004b)	A, D/B, C, E							
RFLP (Goswani et al., 2005)	A, D/B, C/E							
AS (Manonmani et al., 2001; Singh et al., 2004b)		X						
AS (Walton et al., 1999a) AS-SCAR (Manguin et al., 2002)			X					
SSCP (Sharpe et al., 2000)				X	X			
RFLP (Garros et al., 2004b; Van Bortel et al., 1999)				X	X	X		
AS-SCAR (Kengne et al., 2001)				X	X			
AS (Garros et al., 2004a; Phuc et al., 2003)				X	X			
RFLP (Torres et al., 2000)							X[c]	
AS (Ma et al., 2006; Walton et al., 2007)							X[d]	
AS (Dusfour et al., 2007b)								X

[a] AS: Allele-specific; RFLP: restriction fragment length polymorphism; SSCP: single-strand conformation polymorphism; SCAR: sequence characterized amplified region.

[b] These four species are closely related to the Minimus Complex and often sympatric with members of this complex.

[c] Two species identified by this RFLP, *An. dispar* and *An. greeni* (Torres et al., 2000).

[d] Five species identified by these two AS-PCR, *An. maculatus*, *An. dravidicus*, *An. pseudowillmori*, *An. sawadwongporni* and either *An. willmori* for Ma et al. (2006) or chromosomal form K for Walton et al. (2007).

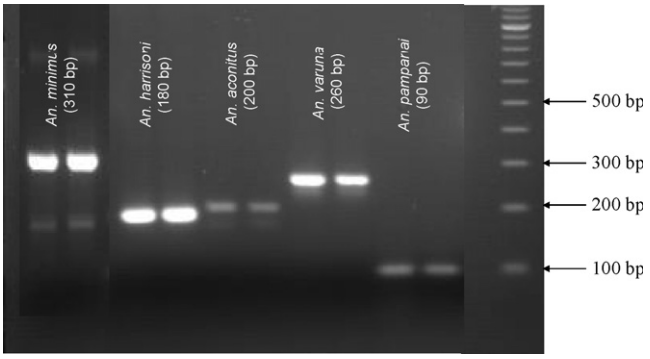


Fig. 1. AS-PCR gel of *An. minimus* and related species (Garros et al., 2004a).

available to distinguish the two sympatric sibling species, *An. minimus* and *An. harrisoni*, as well as related, sympatric species. The RFLP-PCR[4] assay (Garros et al., 2004b; Van Bortel et al., 2000) is very useful for large-scale screening of anopheline fauna, but is technically more time-consuming (two-step PCR assay) and expensive than other assays. The AS-PCR[5] assay is more frequently used to distinguish *An. minimus* and *An. harrisoni*, and also related species such as *An. aconitus*, *An. pampanai*, and *An. varuna* (Table 1, Fig. 1), as it is a quick, reliable, and easy one-step PCR application (Garros et al., 2004a; Phuc et al., 2003).

Despite the availability of molecular assays, important data on the bionomics and distribution of *An. harrisoni* are still unavailable. Studies have shown that *An. minimus* and *An. harrisoni* are considered as main malaria vectors in hilly regions in the Oriental Region. They are commonly found at elevations ranging from 200 to 900 m; they also occur at higher elevation but become quite rare at altitudes above 1,500 m (Duc and Huu, 1973; Harrison, 1980; Oo et al., 2004). *Anopheles minimus* species E is restricted to Ishigaki Island in the Ryukyu Archipelago of Japan, a malaria-free region (Fig. 2) (Green et al., 1990; Harbach et al., 2006; Somboon et al., 2001).

Anopheles minimus extends from northern India eastwards through Vietnam and northward across southern China (up to 24.5°N latitude), including Taiwan (Figs. 2 and 3) (Chen et al., 2002; Garros et al., 2005b; Jambulingam et al., 2005; Phuc et al., 2003; Van Bortel et al., 2000). *Anopheles harrisoni* has been collected in Vietnam, Laos, Thailand, Myanmar, and southern China (up to 32.5°N latitude) (Fig. 2) (Chen et al., 2002; Garros et al., 2005b; Kengne et al., 2001; Phuc et al., 2003; Sharpe et al., 2000; Singh et al., 2006; Trung et al., 2004). *Anopheles minimus* and *An. harrisoni* have been found in sympatry over a large area that includes northern and central Vietnam, southern China, northern Laos, and western Thailand (Fig. 2) (Garros et al., 2006). Whether the two smaller areas in central Vietnam and western Thailand are contiguous with the large areas of sympatry is unknown. Data from Cambodia are scarce and so far no specimens of *An. harrisoni* have been found there (Coosemans et al., 2006). Recently, Singh et al. (2006) recorded the presence of *An. harrisoni* (as *An. minimus* species C) from central Myanmar (Mandalay).

Specific trophic behavior of *An. minimus* and *An. harrisoni* has been studied in four countries, Cambodia, Laos, Vietnam (Garros et al., 2006; Trung et al., 2005; Van Bortel et al., 1999), and Thailand (Sungvornyothin et al., 2006b), but no information is available for species E from Japan. These studies showed that adult females of both species are opportunist feeders as they show a high degree of behavioral plasticity (Trung et al., 2005).

Anopheles minimus is one of the main malaria vectors throughout Southeast Asia (Trung et al., 2004). Nowadays, effective control programs make it difficult to estimate the potential role of *An. harrisoni* as a vector, but its higher exophagic and zoophilic behavior compared to *An. minimus* suggests a lower vectorial capacity in some areas of northern Vietnam (Van Bortel et al., 1999). However, the presence of *An. harrisoni*, without *An. minimus*, in central China where malaria is prevalent suggests that this species, along with three species of the Hyrcanus Group, plays an important role in malaria transmission (Chen et al., 2003, 2006, 2002).

Members of the Minimus Complex occur in the forested foothills of India, Southeast Asia, and southern China where the larvae mainly inhabit clear-water canals and streams with grassy margins and slow moving current (Harrison, 1980). However, larvae of *An. minimus* are also found in water tanks in the suburbs of Hanoi (Van Bortel et al., 1999, 2003). Unpublished data from field observations in northern Vietnam showed that *An. harrisoni* occurs in hilly open areas associated with deforested agro-ecosystems such as maize cultivation, whereas *An. minimus* occurs in more undisturbed closed environments with little anthropogenic change (Garros et al., unpublished data). In western Thailand, *An. harrisoni* was found in fewer types of habitats than *An. minimus*, which occurs in a variety of habitats ranging from agricultural fields to forests with a closed canopy (Rongnoparut et al., 2005). This difference between the two regions could be due to the high behavioral plasticity of both species. Future studies need to focus on the landscape associations of each species for the development of malaria risk maps.

3. The Fluviatilis Complex (Funestus Group, Myzomyia Series)

Cytotaxonomic studies of fixed inversions in polytene chromosomes have identified three chromosomal forms within the Fluviatilis Complex, *An. fluviatilis* S, T, and U, informally recognized as sibling species (Subbarao et al., 1994).

The taxonomic status of the Fluviatilis Complex is unresolved and complicated by the recent publication of molecular variants (mostly based on differences in ITS2 sequence), including species X in Orissa State, India (Manonmani et al., 2003; Naddaf et al., 2003, 2002) and form V recorded only in Iran (Hormozgan Province) (Djadid et al., unpublished data). So far no taxonomic study of the complex has been published and no morphological characters are known to differentiate the different forms. Crossing experiments are required to unequivocally determine whether these chromoso-

[4] RFLP-PCR: Restriction fragment length polymorphism-polymerase chain reaction.
[5] AS-PCR: Allele specific-polymerase chain reaction.

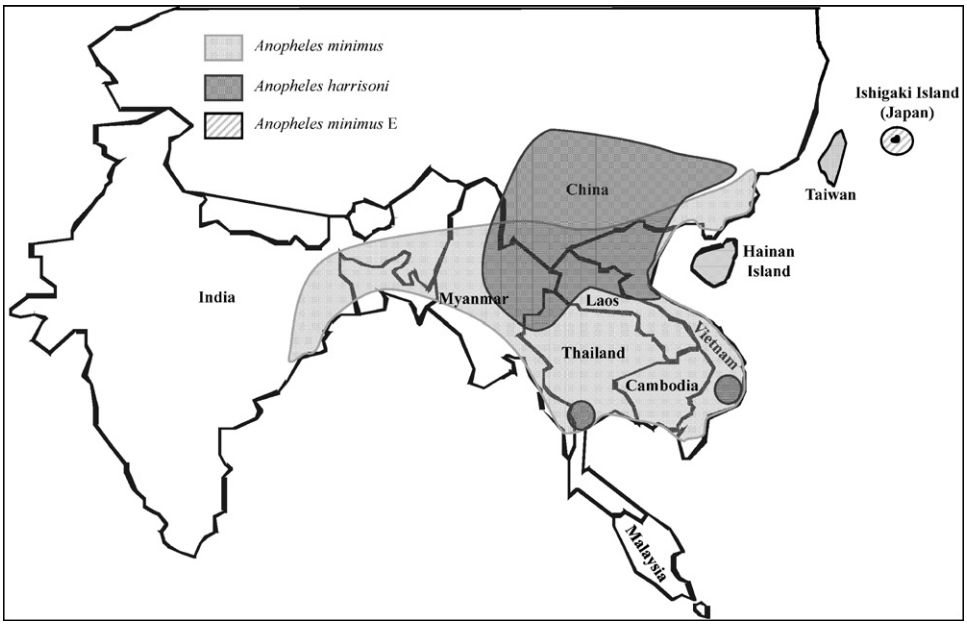


Fig. 2. Distribution of the Minimus Complex based on molecular identifications.

mal forms are definitely distinct species. A complete AS-PCR assay based on sequence differences of the Domain 3 (D3) of 28S rDNA is available for distinguishing members of the complex (Singh et al., 2004b) (Table 1). This assay needs to be applied on a broad scale to determine the precise distribution of each species of the complex.

The Minimus and Fluviatilis Complexes are phylogenetically closely related (Garros et al., 2005a). Based on morphological data and a recent comparison of the D3 sequences of *An. fluviatilis* S and *An. harrisoni*, these two species were deemed to be conspecific (Chen et al., 2003, 2006; Garros et al., 2005a). However, this conclusion was refuted by Singh et al. (2006) who found pair-wise distances of 3.6% and 0.7% for the ITS2 and 28S-D2/D3 loci, respectively, between the species. Chen et al. (2006), in a thorough review, concluded that the Fluviatilis Complex consists of two species, T (with intraspecific variations, including the putative species Y) and U, and two forms, X (different from species S) and V. Singh et al. (2006) removed *An. harrisoni* (as *An. minimus* C) from synonymy with *An. fluviatilis* S, and reported that *An. fluviatilis* X is synonymous with the latter species. Therefore, for Singh et al. (2006) the Fluviatilis Complex includes *An. fluviatilis* species S, T, U, and form V. As currently interpreted, *An. fluviatilis* S is distinct from *An. harrisoni*, which does not occur in India. Further research, however, is needed to clarify the situation.

Little information is available on the bionomics, ecology, and distribution of the taxa outside of India. The Fluviatilis Complex is widely distributed in hilly forested regions of southwestern Asia (Fig. 3) (Bhatt and Kohli, 1996; Malakar et al., 1995; Nanda et al., 2000; Nandi et al., 2000; Subbarao

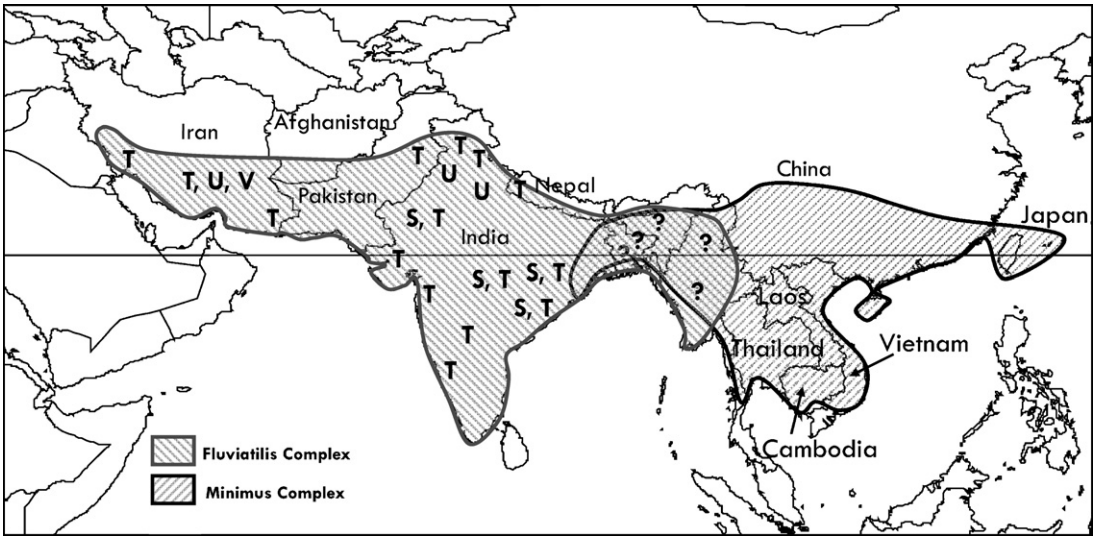


Fig. 3. Distribution of the Fluviatilis and Minimus Complexes (see Fig. 2 for distribution details of the Minimus Complex).

et al., 1994; Vatandoost et al., 2004). However, the distribution of the chromosomal forms has only been well studied in India and Iran. Species T has the largest distribution, which includes India, Nepal, Pakistan, and Iran (Chen et al., 2006). Species U has been recorded in northern India and Iran (Chen et al., 2006; Raeisi et al., 2005). The Fluviatilis and Minimus Complexes overlap in India to Myanmar, but the limits of the overlap are not precisely known (Fig. 3).

Anopheles fluviatilis S is mainly anthropophilic (90%) and endophilic (Nanda et al., 2000), and is known to be a highly competent malaria vector (Nanda et al., 2005), ranking second to *An. culicifacies* species A for total malaria cases transmitted in India (Singh et al., 2004a). *Anopheles fluviatilis* T and U are primarily zoophilic (99%), exophagic, and exophilic, and are regarded as poor or non-vectors, even though species T is known to play a role in the maintenance of malaria in mountainous and hilly regions of India, Pakistan, Iran, and Nepal (Naddaf et al., 2003; Rao, 1984). Members of the Fluviatilis Complex are restricted to forest, especially in mountainous, hilly, and foothill regions of southwestern Asia (Iran, Pakistan, Afghanistan, India, Nepal, Bangladesh, and Myanmar). A study comparing forested and deforested areas of Orissa State in India showed that *An. fluviatilis* S is predominant in forested areas (98% for S; 2% for T), whereas members of the complex are nearly absent in deforested areas where only one specimen of species T was collected (Nanda et al., 2000).

4. The Culicifacies Complex (Funestus Group, Myzomyia Series)

As for the Fluviatilis Complex, the taxonomy of the Culicifacies Complex is unresolved. The complex includes five chromosomal forms, denoted as species A, B, C, D, and E. The members of the complex are cytogenetically defined by fixed paracentric inversions of polytene chromosomes, except for species B and E which are homosequential. However, it is possible to distinguish species B and C based on mitotic chromosomes in semi-gravid females: the Y-chromosome is acrocentric in species B and submetacentric in species E (Kar et al., 1999). Thus, so far no comparative morphological study of the complex has been undertaken.

As cytogenetic analyses can only be done on semi-gravid females, routine field identification is limited during disease control programs. Isozyme analyses based on *Ldh* (Lactate dehydrogenase) distinguish species A and D from species B and C with 95% confidence (Adak et al., 1994), but the main vector, species E, cannot be distinguished from B and C (Kar et al., 1999). Recently developed molecular assays include an AS-PCR based on the D3 domain (Singh et al., 2004c) and a RFLP-PCR based on cytochrome oxidase II (COII) and ITS2 using two restriction enzymes (Goswani et al., 2005). However, these two applications only distinguish A and D from B, C and E (Table 1). In the latter group, an additional RFLP-PCR can distinguish species E from B and C (Goswani et al., 2005), which is useful where species A and B (India) or species B and E (India and Sri Lanka) are sympatric. No currently available application can directly identify all five species, which raises some doubt about their validity.

The Culicifacies Complex is widely distributed from southern China, Vietnam, Laos, Cambodia, Thailand, and Myanmar to India, Pakistan, and Iran, with a western extension into the Arabian Peninsula and Ethiopia (Fig. 4) (Amerasinghe et al., 1999; Kobayashi et al., 1997; Rowland et al., 2002; Van Bortel et al., 2002; Vatandoost et al., 2004; Zhang and Yang, 1996). In parallel with the Fluviatilis Complex, the bionomics and ecology of the species have been largely studied in India, and data are missing for other regions.

Four species of the complex, i.e. species A, C, D, and E, are malaria vectors in India; however, species E is the most efficient vector. Species B is a poor or non-vector (Subbarao et al., 1988). Its distribution is the widest of all the species—it occurs from Iran to China and is the only species of the complex in eastern areas,

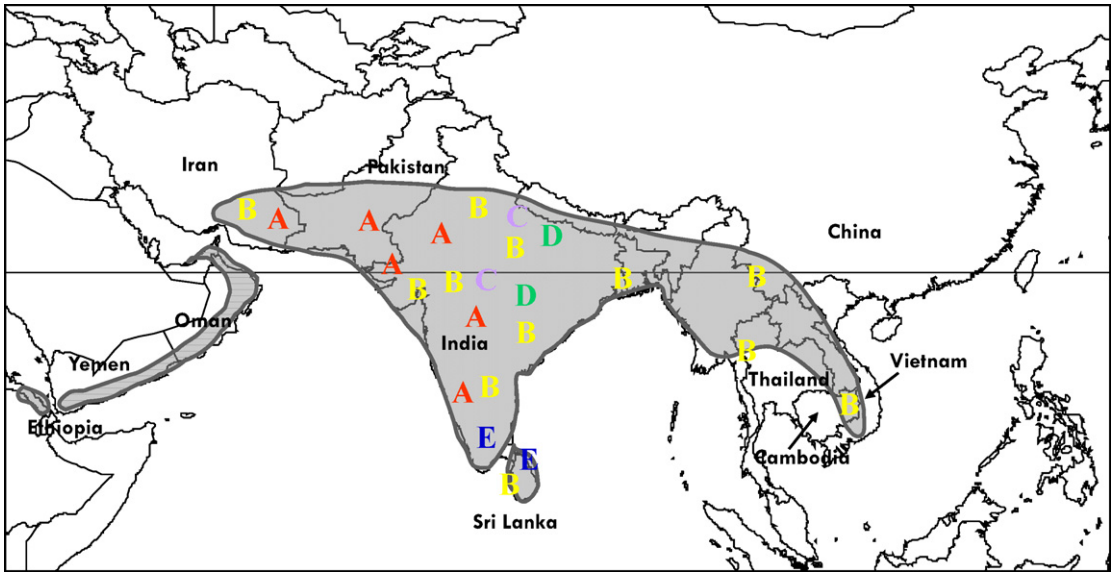


Fig. 4. Distribution of the Culicifacies Complex (grey) and each of the five species (A to E).

including southern China (Sichuan), Vietnam, Laos, Cambodia, and northwestern Thailand (Harrison, 1980; Harrison et al., 1990; Kobayashi et al., 1997; Van Bortel et al., 2002; Vatandoost et al., 2004; Zhang and Yang, 1996). Westwards, it occurs in sympatry with other species of the complex, especially species E, the most important malaria vector in southern India and Sri Lanka (Fig. 4). No data are available for the species that occur in the Arabian Peninsula and Ethiopia.

Anopheles culicifacies s.l. is responsible for the transmission of 60–70% of malaria cases in India, mainly due to species E, which is highly anthropophilic (90%) and endophilic (Subbarao, 1988). Species A, C, and D are mainly zoophilic, with a low anthropophilic index that does not exceed 3–4% (Subbarao and Sharma, 1997). Therefore, these three species play a minor role in malaria transmission (Sharma et al., 1995), although species C was found responsible for local malaria transmission in deforested riverine areas of central India (Nanda et al., 2000). Species B is highly zoophilic but it sometimes plays a role in sporadic epidemics in Myanmar, Laos, and Vietnam (Oo et al., 2004; Sucharit et al., 1988; Trung et al., 2004). *Anopheles culicifacies* s.l. occurs in different ecotypes, such as forests with perennial streams and deforested riverine ecosystems (hills, plains) or irrigated areas. Larval habitats include irrigated canals, rock pools, and sandy pools near paddies or quarries. A study in Sri Lanka showed that species E exploits a wide range of habitats, which reflects a greater environmental adaptability of this malaria vector than species B (Surendran and Ramasamy, 2005). The study by Nanda et al. (2000) in Orissa (India) that compared forested and deforested ecosystems showed that specimens of *An. culicifacies* s.l., unlike those of the Fluviatilis Complex (see above), are present in both ecosystems. In forested areas, *An. culicifacies* C (71%) outnumbered species B; in deforested areas species C (78%), B (21%), and species A (1%) were present. This also shows the ability of *An. culicifacies* C to adapt to environmental changes.

Even though the Fluviatilis and Culicifacies Complexes include some major malaria vectors, further studies are needed to resolve the taxonomic status of the individual species. Other complexes, especially the Dirus and Leucosphyrus Complexes and the Maculatus Group (see below) have been thoroughly studied and serve as models for the delineation of species and the development of molecular identification methods that provide important tools for improving our knowledge of the distribution and bionomics of the individual species.

5. The Dirus Complex (Leucosphyrus Group, Neomyzomyia Series)

The Dirus Complex includes seven species that vary from highly competent malaria vectors to non-vectors of human malaria in tropical evergreen rainforests, cultivated forests, and forest fringes throughout Southeast Asia (Baimai, 1998; Oo et al., 2004) (Fig. 5). The taxonomy of the complex was recently resolved and all the species now have morphological descriptions and formal Latin names (Sallum et al., 2005), and their distributions in Southeast Asia have been mapped (Baimai, 1998; Obsomer et al., 2007).

Anopheles dirus (=*An. dirus* species A) has a wide distribution in eastern Asia—it is known to occur in Myanmar, Thailand, Cambodia, Laos, Vietnam, and Hainan Island (China) (Fig. 6). *Anopheles cracens* (=*An. dirus* B) is known from southern (peninsular) Thailand, peninsular Malaysia, and

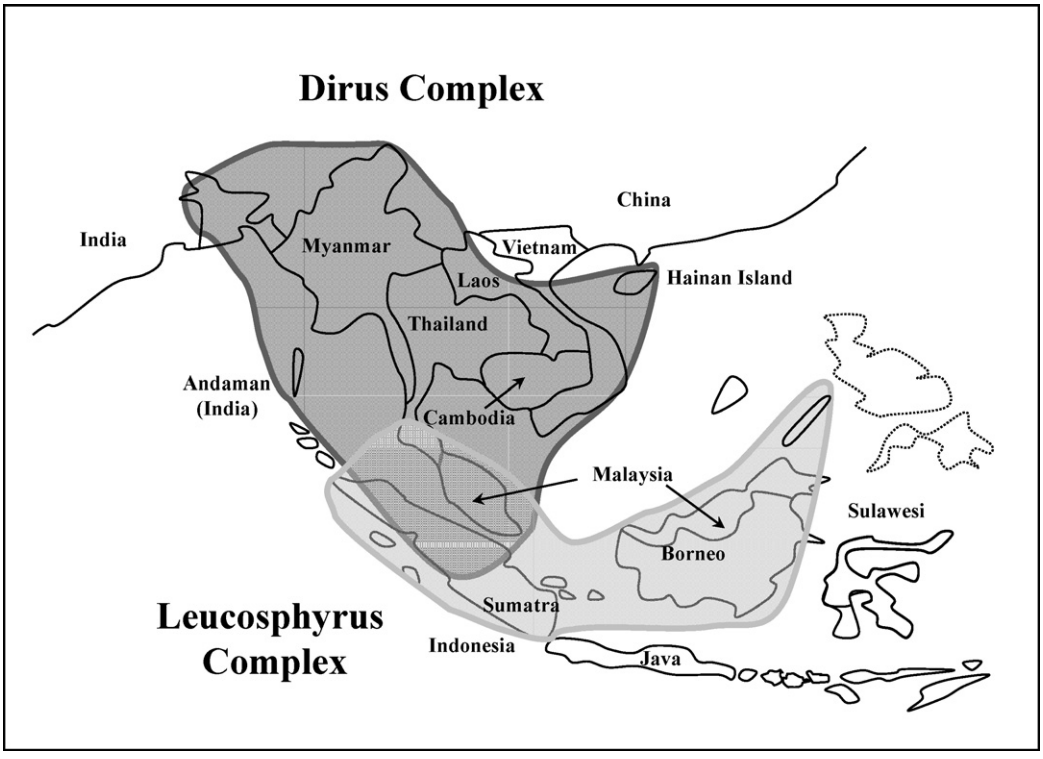


Fig. 5. Distribution of the Dirus (dark grey) and Leucosphyrus (light grey) Complexes showing the zone of overlap in the Malay Peninsula and Sumatra (Indonesia).

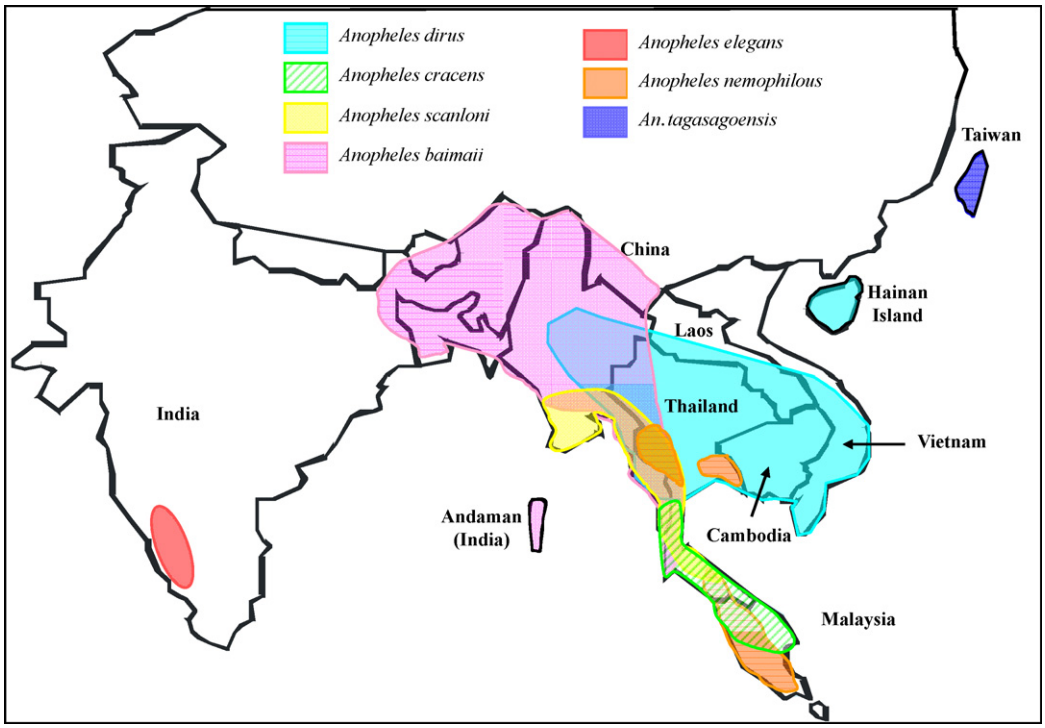


Fig. 6. Distribution of the seven members of the Dirus Complex (records of *An. cracens* in Sumatra are not shown).

Sumatra (Indonesia). *Anopheles scanloni* (=*An. dirus* C) occurs in a relatively narrow area along the borders of southern Myanmar and western and southern Thailand, and appears to be intimately linked to limestone environment. In Thailand, *An. scanloni* populations are restricted to ''islands'' of limestone karst habitat that support high levels of population structure (O'Loughlin et al., 2007). *Anopheles baimaii* (= *An. dirus* D) occurs in areas from southwestern China (Yunnan Province), western Thailand, Myanmar, and Bangladesh to northeastern India and the Andaman Islands (India) (Sallum et al., 2005). *Anopheles elegans* (=*An. dirus* E) is restricted to hilly forests of southwestern India. *Anopheles nemophilous* (=*An. dirus* F) has a patchy distribution along the Thai–Malay peninsula and Thai border areas with Myanmar and Cambodia. Finally, *An. takasagoensis* is restricted to Taiwan.

Initial recognition of the species was achieved primarily by cross-mating experiments and studies of polytene chromosome banding patterns, and subsequently by electrophoresis analyses of allozyme variation (Baimai et al., 1987; Green et al., 1992a; Hii, 1984). Since several species of the complex occur in sympatry, it was important to clearly identify the specimens. Therefore, two AS-PCR assays were developed (Table 1). Walton et al. (1999b) designed an AS-PCR based on ITS2 sequence to distinguish and unambiguously identify *An. dirus*, *An. cracens*, *An. scanloni*, *An. baimaii*, and *An. nemophilous*, and Manguin et al. (2002) developed a SCAR[6]-based PCR to identify the same five species. The recent revisionary study of the complex by Sallum et al. (2005) revealed that morphological characters are present in all life stages that distinguish the species, but the authors stated that ''Due to the variability of their elaborate ornamentations, separating the many species of this group [Leucosphyrus Group] will always be morphologically challenging.''

The Dirus Complex includes primary vectors of forest malaria, principally *An. dirus* and *An. baimaii*, which transmit both *Plasmodium falciparum* and *P. vivax*. Records show that these species are anthropophilic, mainly exophagic, and highly competent vectors (Baimai et al., 1988). *Anopheles cracens* is an anthropophilic species that may play a role in malaria transmission, and also *An. scanloni* or potentially *An. elegans*. The availability of morphological and molecular identification methods will allow researchers to investigate the degree to which these species may be involved in malaria transmission. The two other species of the complex, *An. nemophilous* and *An. takasagoensis*, appear to be non-vectors of human malaria due to their zoophilic behavior (Baimai, 1988; Peyton and Harrison, 1980).

Species of the Dirus Complex are forest mosquitoes (forested foothills, deep forests, cultivated forests), but are occasionally collected in open areas adjacent to forest (forest fringes). Larvae of the species typically inhabit small, usually temporary, mostly shaded bodies of fresh, stagnant water, such as pools, puddles, animal footprints, streams, and even wells in hilly or mountainous regions with primary, secondary evergreen or deciduous forests, bamboo forests, and fruit and rubber plantations (Baimai et al., 1988; Oo et al., 2002; Prakash et al., 2002).

Species of the Dirus Complex are closely related to members of the Leucosphyrus Complex, and this has been the cause of considerable confusion in published literature. Numerous studies, mainly based on crossing experiments, cytogenetics,

[6] SCAR: Sequence characterized amplified region.

allozyme data, and more recently molecular methods, have been necessary to recognize the individual species and to confirm their taxonomic status (Baimai, 1988, 1989; Baimai et al., 1987; Green et al., 1992a; Hii, 1984; Sallum et al., 2005).

6. The Leucosphyrus Complex (Leucosphyrus Group, Neomyzomyia Series)

The Leucosphyrus Complex of four species, recently revised by Sallum et al. (2005), includes *An. balabacensis*, *An. introlatus*, *An. latens* (= *An. leucosphyrus* A), and *An. leucosphyrus* (= *An. leucosphyrus* B). *Anopheles leucosphyrus* and *An. latens* are morphologically indistinguishable, but they can be differentiated from *An. balabacensis* and *An. introlatus* (Sallum et al., 2005). Various members of the complex occur in southern Thailand, Malaysia (Sabah, Sarawak, mainland), Indonesia (Java, Kalimantan, Sumatra), and Balabac Island of the Philippines. *Anopheles latens* and *An. introlatus* are sympatric with members of the closely related Dirus Complex in the Malay Peninsula, including southern Thailand (Fig. 5). *Anopheles latens* is also widely distributed in Borneo (Kalimantan, Sarawak, Sabah), together with *An. balabacensis* in the forested areas of eastern Borneo (Fig. 7) (Rattanarithikul and Harrison, 1973; Rattanarithikul et al., 2006). *Anopheles leucosphyrus* has only been found in Sumatra.

Two species, *An. balabacensis* and *An. latens*, are recognized as malaria vectors with sporozoite infection rates of 1.3% and 1%, respectively, and both species are reported to be exophagic and exophilic (Harbach et al., 1987). No information exists on the vectorial status of *An. leucosphyrus*. *Anopheles introlatus* is known to transmit simian malaria in Malaysia (Eyles et al., 1963). Overall, the importance of the species as vectors of human malarial parasites is not well established because the species have been largely misidentified.

Species of this complex are forest mosquitoes and share the same types of habitats as members of the Dirus Complex. Typical larval habitats are freshwater ground pools along stream margins, flood pools, seepage pools, sandy pools, wallows, small shallow streams, elephant footprints, and even large swamps (Leicester, 1903; Sallum et al., 2005). Water in the habitats may be stagnant or slow running, turbid or clear, and partially or heavily shaded. The species occur at elevation ranging from 70 to 500 m (Sallum et al., 2005).

7. The Maculatus Group (Neocellia Series)

The Maculatus Group includes eight formally named species (Harbach, 2004): *An. pseudowillmori* and *An. willmori* and six species assigned to subgroups, the Maculatus Subgroup, which includes *An. dispar*, *An. greeni*, *An. dravidicus*, and *An. maculatus*, and the Sawadwongporni Subgroup, which includes *An. notanandai* and *An. sawadwongporni* (Ma et al., 2006). Members of the group are variously distributed in areas from India to Indonesia and the Philippines. Two species, *An. dispar* and *An. greeni* (Fig. 8), are found exclusively in the Philippines (Rattanarithikul and Harbach, 1990; Torres et al., 1997). In addition, a recent genetic study of the Maculatus Group found that chromosomal form K in eastern Thailand has an unique ITS2 sequence that is 3.7% divergent from the closest taxon (*An. sawadwongporni*), which indicates it is a distinct species (Walton et al., 2007). Hence, the group would appear to include nine species with form K falling into the Sawadwongporni Subgroup (Ma et al., 2006).

Adults of the complex are difficult to identify to species using morphology because of overlapping characters. In fact, members of the group were first recognized using cytogenetics (Baimai et al., 1993; Green and Baimai, 1984; Green et al., 1985, 1992b). Eleven cytogenetic forms were described that represent eight genetic species (Green et al., 1991; Rattanarithikul and Green, 1986). The correspondence between the formally named species and the 11 chromosomal forms is given by Walton et al. (2007).

A reliable and easy RFLP-PCR assay (Table 1) was developed to distinguish *An. dispar* and *An. greeni* (Torres

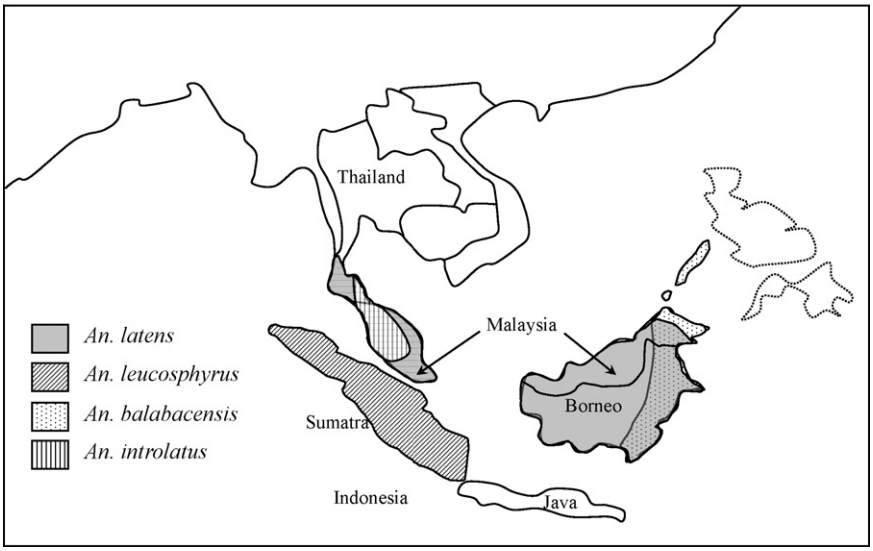


Fig. 7. Distribution of the four members of the Leucosphyrus Complex.

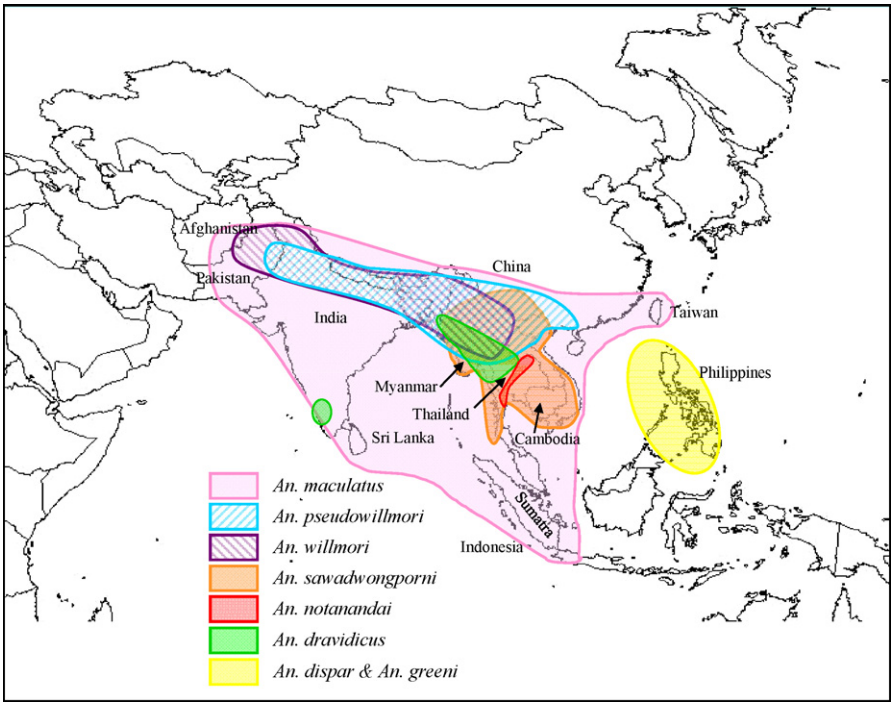


Fig. 8. Distribution of the eight members of the Maculatus Group.

et al., 2000). Use of this method should shed light on the vector status of these two species as previous data collected on *An. maculatus* s.l. in the Philippines is unreliable. Two allele-specific-PCR (AS-PCR) assays (Table 1) have also been developed to distinguish *An. dravidicus*, *An. maculatus*, *An. pseudowillmori*, *An. sawadwongporni*, and either *An. willmori* (Ma et al., 2006) or chromosomal form K (Walton et al., 2007).

Anopheles pseudowillmori occurs from northern India (Punjab, Assam, Kasauli) through northwestern Thailand and southern China (Yunnan), and *An. willmori* is found at higher altitudes in Afghanistan, Pakistan (Kashmir), northern India (Punjab to Assam), Nepal, northern Thailand (Chiang Mai), and southern China (Yunnan) (Green et al., 1992b; Li et al., 2003; Pradhan et al., 1970; Rao, 1984). *Anopheles maculatus* has the widest distribution, ranging from Afghanistan, Pakistan, India, and Sri Lanka eastward through Southeast Asia to Taiwan and Indonesia. *Anopheles dravidicus* has a very peculiar distribution as it is found in two separate areas, one in southwestern India (Nilgiri Hills) and the other in northwestern Myanmar (Kale Valley) and northern Thailand (Kanchanaburi, Tak, Chiang Mai, Chiang Rai, Mae Hong Son, Loei, Phrae) (Rattanarithikul and Green, 1986; Rattanarithikul et al., 2006). This disjunctive distribution is surprising and further investigation must be carried out to determine whether it is actually the same species in both areas and/or to find new populations in the intervening areas. The two members of the Sawadwongporni Subgroup are distributed in the Mekong Region: *An. sawadwongporni* occurs from Myanmar to China and Vietnam and *An. notanandai* is only known from west-central Thailand (Baimai et al., 1993; Green et al., 1992b; Oo et al., 2004; Rattanarithikul and Green, 1986).

Members of the Maculatus Group are variously involved in malaria transmission in the Oriental Region (Rahman et al., 1993; Rongnoparut et al., 1996; Upatham et al., 1988), but individual species may have quite different vectorial capacities. The precise role of each species is unknown due to misidentifications based on morphological characters. In addition, the vectorial capacity of an individual species seems to vary depending on location. *Anopheles maculatus* has a wide distribution (Fig. 8), but it is considered to be a major malaria vector only in eastern India, southern Thailand, peninsular Malaysia, and Java (Barcus et al., 2002; Green et al., 1991; Hodgkin, 1956; Rahman et al., 1993; Rattanarithikul et al., 1996b; Reid, 1968). Whereas *An. willmori* is one of the primary vectors in Nepal (Pradhan et al., 1970), it is rare in Thailand and not involved in malaria transmission. *Anopheles pseudowillmori* is a secondary vector in northwestern Thailand along the Myanmar border (Green et al., 1991, 1992b). *Anopheles sawadwongporni* has been found with sporozoite rates of 1–2% in Thailand where it is an important malaria vector along with *An. maculatus* (Rattanarithikul et al., 1996a; Somboon et al., 1998). The two species that occur in the Philippines, *An. dispar* and *An. greeni*, are regarded as secondary vectors but their specific involvement in malaria transmission has not been determined. These two species exhibit strong exophagic and zoophilic behavior, with a biting rate on water buffalo that is 50 times the human landing rate (Torres et al., 1997). *Anopheles notanandai* and *An. dravidicus* are not known to be involved in malaria transmission (Mouchet et al., 2004).

Members of the Maculatus Group are found in or near hilly and mountainous areas. Larvae of *An. maculatus* s.l. have been collected in a diversity of habitats, including ponds, lakes, swamps, ditches, wells, different pools (grassy, sandy, ground, flood, stream), stream margins, seepages springs, rice fields, foot or wheel prints, and occasionally tree holes or bamboo

stumps (Rattanarithikul et al., 1995, 2006, 1994). More specific studies have shown that the species have preferred habitats. For instance, larvae of *An. willmori* are found only along the margins of streams at altitudes between 990 and 1450 m in northern Thailand, and larvae of *An. pseudowillmori* have been collected primarily in rice fields, stream margins, ponds, pits, and wells (Rattanarithikul et al., 1995, 2006).

8. The Sundaicus Complex (Pyretophorus Series)

Behavioral and ecological differences, notably the occurrence of immature stages in brackish and freshwater, led Reid (1970) to suspect that *Anopheles sundaicus* was a species complex. Subsequently, the presence of three cytogenetic forms (A, B, and C) were detected in Sumatra, Java, and Thailand, and confirmed by allozyme analysis (Sukowati et al., 1996, 1999). *Anopheles sundaicus* species A was found in both Indonesia and Thailand; species B, was strongly linked to freshwater in northern Sumatra and central Java; and species C was only found at a single locality (Asahan) in northern Sumatra where all three cytotypes were collected in sympatry. A fourth cytotype D (Nanda et al., 2004) was later identified from the Nicobar and Andaman Islands of India (Fig. 9) and recently confirmed by molecular analysis of ITS2 and D3 sequences of rDNA (Alam et al., 2006). The identity of *An. sundaicus* s.s. was fixed with the designation of a neotype from the Lundu District of Sarawak (Fig. 9) in northern Borneo (Malaysia) based on morphology and sequences for the ITS2 rDNA and COI mtDNA loci (Linton et al., 2005). Finally, two allopatric species were verified based on two mitochondrial markers, COI and Cytochrome *b* (Cyt-b), and the ribosomal marker, ITS2 (Dusfour et al., 2004b). One of these two species, *An. epiroticus* (=*An. sundaicus* A), occurs in coastal brackish water sites from southern Vietnam to peninsular Malaysia (Fig. 9). The other species, *An. sundaicus* species E (Fig. 9), occurs in Sumatra and Java (Dusfour et al., 2007b). Neither ITS2 nor COI revealed a distinction between cytogenetic forms B and C from Asahan (Sumatra), rather the molecular data indicated the existence of only one species, which was informally designated *An. sundaicus* species E (Dusfour et al., 2007b). An allele-specific PCR was developed for the identification of three of the four species: *An. sundaicus*, species E, and *An. epiroticus* (Table 1) (Dusfour et al., 2007a).

Anopheles sundaicus s.l. is a malaria vector in coastal areas (Fig. 9) that extend from northeastern India eastwards to southern Vietnam (south of the 11th parallel) and southwards to the Andaman and Nicobar Islands (India), Malaysia (peninsular and Borneo), and Indonesia (Java, Sumatra, Sulawesi) (Dusfour et al., 2004a). Adult females are mainly anthropophilic and endophilic, and exhibit both endophagy and exophagy. This taxon is responsible for regular malaria outbreaks in certain areas where it occurs in great numbers (Oo et al., 2004). The availability of a reliable PCR identification method will allow future investigators to determine more precisely the behavior (and the distribution) of each species of the complex.

Due to its ecological and behavioral plasticity, *An. sundaicus* s.l. has adapted to a range of coastal and inland environmental situations. It is regarded mainly as a brackish water taxon, but larvae tolerate a wide range of salinity from freshwater to sea water (Nguyen Tang Am et al., 1993). The immature stages require sunlit habitats with stagnant fresh or brackish water, floating algae, and non-invasive vegetation. Filamentous floating algae and aquatic plants appear to be crucial for the development of the larvae. Aquatic flora provides food (micro-algae and bacteria) and protection against predators. Particularly favorable habitats are coastal shrimp/fish ponds or irrigated inland sea-water canals, but immature stages also inhabit ponds, swamps, mangrove, and rock pools (Chang et al., 2001; Dusfour et al., 2004a; Harinasuta et al., 1974; Ikemoto et al., 1986; Kalra, 1978; Nguyen Tang Am et al., 1993). The affinity of *An. epiroticus* with aquaculture, particularly shrimp and fish ponds in southern Vietnam (Nguyen Tang Am et al., 1993; Trung et al., 2004), needs to be monitored on a larger scale as this economic activity is growing throughout Southeast Asia with an increasing risk of malaria epidemics.

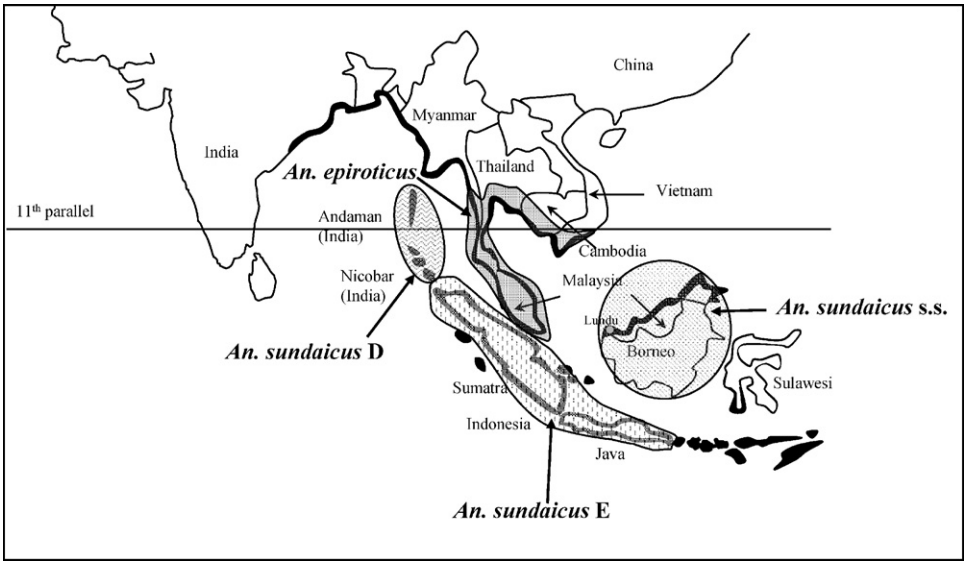


Fig. 9. Distribution of the four members of the Sundaicus Complex.

9. Implications for vector control

Malaria vector control programs in Southeast Asia are largely based on the use of insecticides for indoor residual spraying (IRS) and, to a lesser extent, insecticide treated nets (ITNs). IRS programs were first organized in the 1950s and implemented for decades in the Southeast Asian countries. Later, ITNs were required to supplement the IRS campaigns. Nowadays, IRS is mainly used to control focal epidemics (WHO, 2007c). Until recently, coverage of ITN and IRS only reached 10–20% of the populations at risk, but with the recent support of the Global Fund to the malaria endemic countries, the use of ITNs has increased significantly. However, more training and resources are required. The current trend in Southeast Asian is to rely on the increasing use of ITN based on several actions: (i) expansion to a maximum coverage of ITN, including re-treatments; (ii) promotion of long-lasting ITNs to avoid re-treatments; (iii) lowering the cost of bed nets to a minimum through tax exemption; and (iv) increased access to nets through the commercial sector and targeted subsidies for the poorest people (WHO, 2007d).

These actions will most likely improve the malaria situation in Southeast Asia but they should be closely associated with a better knowledge of the targeted malaria vectors. The success of vector control programs is highly dependant on a thorough knowledge of behavior and bionomics of the vector, which must be precisely identified to ensure the application of appropriate control measures (Trung et al., 2005). Precise information about ecological changes, the behavioral plasticity of most *Anopheles* species as mentioned above, and the accurate identification of the vector species is required for the development of effective vector control strategies that are based principally on IRS and ITN. For instance, IRS is more efficient where the vector is endophilic and bites late at night, such as *An. minimus* and *An. epiroticus*, whereas ITNs seem more effective than other techniques for the control of exophagic and exophilic mosquitoes such as *An. dirus* s.l. (Trung et al., 2005). For the protection of people at occupational risk, such as forest workers, the use of treated hammock nets needs to be implemented, although their efficacy against this vector is sometimes questionable (Trung et al., 2004). Rapid ecological changes, such as those occurring in Southeast Asia, especially deforestation, are modifying the cohort of malaria vector species, as noted for the Minimus and Dirus Complexes, and surveillance needs to be done on a regular basis because vector control programs are highly dependant on the vector species and its behavior. This is necessary for appropriately targeting the species involved in malaria transmission. Vector control should also be adapted to seasonal variations. In hilly forested areas of Southeast Asia, it is common knowledge that malaria transmission is perennial due to the presence of species of the Dirus and Minimus Complexes, the first being present mainly during the rainy season and the latter during the drier periods of the year (Harbach et al., 1987; Ismail et al., 1978; Phan, 1998; Rattanarithikul et al., 1996a).

Beside the use of chemical insecticides through IRS and ITN, efforts to minimize this dependency have been undertaken while searching and developing eco-friendly alternative methods for the control of vector mosquitoes. Instead of controlling adult mosquitoes, these alternative methods target immature stages, particularly larvae. Application of environmental management and biological control need to be utilized wherever it is cost effective and feasible (WHO, 2007d). Nowadays, biological control methods are once again receiving much research focus for malaria vector control. Larvivorous fish have been used for over 100 years for controlling mosquito densities and malaria incidence in many countries, including India, Malaysia, Papua New Guinea, and Thailand (Rozendaal, 1997). *Gambusia affinis* and *Poecilia reticulata* are the most successful and effective larvivorous fish for vector control (Ghosh and Dash, 2007). In India, remarkable results have been achieved for the control of malaria vectors like *An. culicifacies* species A that breed predominantly in ponds and wells (Ghosh et al., 2005). Larvivorous fish have also been used to control *An. sundaicus* s.l. in Indonesia (Ikemoto et al., 1986) and *An. dirus* s.l. in gem pits in Thailand (Kitthawee et al., 1993). In certain areas where *An. dirus* s.l. utilizes small habitats, for example wells in Myanmar, this biological control method should be applied and may give good results (Oo et al., 2004). However, the effectiveness of this strategy is questionable (Meek, 1995), particularly in large wetlands where its efficacy has not been demonstrated. Its implementation requires some baseline knowledge of vector biology, and should be included as part of an integrated malaria control program (Ghosh and Dash, 2007; Meek, 1995). Another potential biological control agent, tested in Japan, involves the use of copepod species as predators in rice fields during the summer (Dieng et al., 2003). Results showed that copepods are efficient biological control agents against mosquito larvae. However, the reduction of larval densities is temporary if not properly managed; hence, the method only has a limited effect on malaria transmission (Subbarao and Sharma, 1997).

Various biolarvicides have also been thoroughly investigated, especially strains of the bacteria *Bacillus sphaericus* and *B. thuringiensis* var. *israelensis* H-14 (*Bti*), which are highly effective against mosquito larvae at very low doses and safe to other non-target organisms. Formulations of *B. sphaericus* have been used against *An. stephensi*, *An. subpictus*, and *An. culicifacies* s.l., but repeated applications in the same habitat resulted in the development of resistance in the larvae of the targeted species (Mittal, 2003). Therefore, *B. sphaericus* has limited prospects for the control of malaria vectors. *Bti* formulations have a broader spectrum of activity against *Aedes*, *Culex*, and *Anopheles* species, but it was found less effective against *Anopheles* due to many limitations (exposure to sunlight reduces efficacy, weekly application required in most habitats, etc.) (Mittal, 2003).

Biological control may have an impact on malaria vectors in certain specific situations but in most cases it has proven to be too tedious for general use because the types of larval habitats of the main malaria vectors are not conducive of this kind of strategy (Meek, 1995). However, biological control can still be considered within an integrated vector management strategy based on selective application of various control measures determined by the eco-epidemiological situation of malaria. For instance, in certain rural communities, biological control

may be a helpful supplement to IRS or ITN, particularly during the dry season when larvae of vector species are concentrated in relatively few habitats (Walker and Lynch, 2007).

Environmental management for larval control is another option, especially against *An. sundaicus* s.l. Some success in controlling this taxon was achieved in Malaysia by building bunds and digging drains for excluding brackish water (Moorhouse and Wharton, 1965). More recently, larvicide by clearance of algae has been used successfully in Indonesia against this malaria vector species (Kirnowardoyo, 1988; Soekirno et al., 1983). In Malaysia, larvae of *An. maculatus* s.l. have been controlled by periodic flushing of streams using small dams fitted with siphons (Williamson and Scharff, 1936), or by drainage (Moorhouse and Wharton, 1965). These larvicidal methods are opportunities to complement adulticiding along with other components of integrated vector management, and have a direct bearing on concerns about insecticide resistance, environmental impact, rising costs of IRS, and logistical constraints (Walker and Lynch, 2007).

The lesson learned over the years is that malaria control is too complex to be addressed by a single approach (Shiff, 2002). It is important to tailor the strategy to the prevailing malaria vector species, as well as ecological and epidemiological conditions (Mouchet and Carnevale, 1998). We now understand the ecological conditions that affect and regulate the distribution and abundance of mosquito populations (Gillies, 2001), and reliable and easy molecular methods have been developed to supplement morphological identification of closely related and isomorphic species (Table 1). Therefore, combined sustainable and appropriate vector control measures in relation to the targeted vector species and prophylaxis must be implemented to achieve the goal of the revised strategy of the Southeast Asian Regional Committee, which aims to reduce the level of malaria morbidity and mortality recorded in 2000 by 50% before 2010 (WHO, 2007a).

References

Adak, T., Subbarao, S.K., Sharma, V.P., Rao, S.R., 1994. Lactate dehydrogenase allozyme differentiation of species in the *Anopheles culicifacies* complex. Med. Vet. Entomol. 8, 137–140.

Alam, M.T., Das, M.K., Ansari, M.A., Sharma, Y.D., 2006. Molecular identification of *Anopheles* (*Cellia*) *sundaicus* from the Andaman and Nicobar islands of India. Acta Trop. 97, 10–18.

Amerasinghe, P.H., Amerasinghe, F.P., Konradsen, F., Fonseka, K.T., Wirtz, R.A., 1999. Malaria vectors in a traditional dry zone village in Sri Lanka. Am. J. Trop. Med. Hyg. 60, 421–429.

Baimai, V., 1988. Population cytogenetics of the malaria vector *Anopheles leucosphyrus* group. Southeast Asian J. Trop. Med. Public Health 19, 667–680.

Baimai, V., 1989. Speciation and species complexes of the *Anopheles* malaria vector in Thailand. In: Proceeding of the 3rd conference on malaria research, Thailand, 18–20 October, pp. 146–162.

Baimai, V., 1998. Heterochromatin accumulation and karyotypic evolution in some Dipteran insects. Zoolog. Stud. 37, 75–88.

Baimai, V., Andre, R.G., Harrison, B.A., Kijchalao, U., Panthusiri, L., 1987. Crossing and chromosomal evidence for two additional sibling species within the taxon *Anopheles dirus* Peyton and Harrison (Diptera: Culicidae) in Thailand. In: Proceedings of the Entomological Society of Washington, vol. 89. pp. 157–166.

Baimai, V., Kijchalao, U., Sawadwongporn, P., Green, C.A., 1988. Geographic distribution and biting behaviour of four species of the *Anopheles dirus* complex (Diptera: Culicidae) in Thailand. Southeast Asian J. Trop. Med. Public Health 19, 151–161.

Baimai, V., Kijchalao, U., Rattanarithikul, R., Green, C.A., 1993. Metaphase karyotypes of *Anopheles* of Thailand and Southeast Asia. II. Maculatus Group, Neocellia Series, subgenus *Cellia*. Mosq. Syst. 25, 116–123.

Barcus, M.J., Laihad, F., Sururi, M., Sismadi, P., Marwoto, H., Bangs, M.J., Baird, K.J., 2002. Epidemic malaria in the Menoreh hills of Central Java. Am. J. Trop. Med. Hyg. 66, 287–292.

Bhatt, R.M., Kohli, V.K., 1996. Biting rhythms of some anophelines in Central Gujarat. Indian J. Malariol. 33, 180–190.

Chang, M.S., Linton, Y.M., Harbach, R.E., 2001. The *Anopheles sundaicus* species complex—a preliminary study of larval ecology from two sites in Sarawak, Malaysia. Mekong Malaria Forum 8, 91–95.

Chen, B., Harbach, R.E., Butlin, R.K., 2002. Molecular and morphological studies on the *Anopheles minimus* group of mosquitoes in southern China: taxonomic review, distribution and malaria vector status. Med. Vet. Entomol. 16, 253–265.

Chen, B., Butlin, R.K., Harbach, R.E., 2003. Molecular phylogenetics of the Oriental members of the Myzomyia Series of *Anopheles* subgenus *Cellia* (Diptera: Culicidae) inferred from nuclear and mitochondrial DNA sequences. Syst. Entomol. 28, 1–13.

Chen, B., Butlin, R.K., Pedro, P.M., Wang, X.Z., Harbach, R.E., 2006. Molecular variation, systematics and distribution of the *Anopheles fluviatilis* complex in southern Asia. Med. Vet. Entomol. 20, 33–43.

Coosemans, M., Thuan, L.K., Socheat, D., Phompida, S., Baimai, V., Manguin, S., Harbach, R.E., Hemingway, J. 2006. Monitoring insecticide resistance and mapping malaria vectors in Southeast Asia: a prerequisite for sustainable malaria vector control. (INCO report, IC4-CT-2002-10041), p. 202.

Dieng, H., Boots, M., Tuno, N., Tsuda, Y., Takagi, M., 2003. Life history effects of prey choice by copepods: implications for biocontrol of vector mosquitoes. J. Am. Mosq. Control Assoc. 19, 67–73.

Duc, N.H., Huu, H.V., 1973. Distribution of *Anopheles minimus* in Northern Vietnam. Ann. Sci. Works 134–140.

Dusfour, I., Harbach, R.E., Manguin, S., 2004a. Bionomics and systematics of the oriental *Anopheles sundaicus* complex in relation to malaria transmission and vector control. Am. J. Trop. Med. Hyg. 71, 518–524.

Dusfour, I., Linton, Y.M., Cohuet, A., Harbach, R.E., Baimai, V., Trung, H.D., Seng, C.M., Matusop, A., Manguin, S., 2004b. Molecular evidence of speciation between island and continental populations of *Anopheles* (*Cellia*) *sundaicus* (Diptera: Culicidae), a principal malaria vector taxon in Southeast Asia. J. Med. Entomol. 41, 287–295.

Dusfour, I., Blondeau, J., Harbach, R.E., Vythilingham, I., Baimai, V., Trung, H.D., Sochanta, T., Bangs, M.J., Manguin, S., 2007a. Polymerase chain reaction identification of three members of the *Anopheles sundaicus* (Diptera: Culicidae) complex, malaria vectors in Southeast Asia. J. Med. Entomol. 44, 723–731.

Dusfour, I., Michaux, J.R., Harbach, R.E., Manguin, S., 2007b. Speciation and phylogeography of the Southeast Asian *Anopheles sundaicus* complex. Infect. Genet. Evol. 7, 484–493.

Eyles, D.E., Warren, M., Guinn, E., Wharton, R.H., Ramachandran, C.P., 1963. Identification of *Anopheles balabacensis introlatus* as a vector of monkey Malaria in Malaya. Bull. World Health Organ. 28, 134–135.

Foley, D.H., Rueda, L.M., Wilkerson, R.C., 2007. Insight into global mosquito biogeography from country species records. J. Med. Entomol. 44, 554–567.

Garros, C., Koekemoer, L.L., Coetzee, M., Coosemans, M., Manguin, S., 2004a. A single multiplex assay to identify major malaria vectors within the African *Anopheles funestus* and the Oriental *An. minimus* groups. Am. J. Trop. Med. Hyg. 70, 583–590.

Garros, C., Koekemoer, L.L., Kamau, L., Awolola, T.S., Van Bortel, W., Coetzee, M., Coosemans, M., Manguin, S., 2004b. Restriction fragment length polymorphism method for the identification of major African and Asian malaria vectors within the *Anopheles funestus* and *An. minimus* groups. Am. J. Trop. Med. Hyg. 70, 260–265.

Garros, C., Harbach, R.E., Manguin, S., 2005a. Morphological assessment and molecular phylogenetics of the Funestus and Minimus Groups of *Anopheles* (*Cellia*). J. Med. Entomol. 42, 522–536.

Garros, C., Harbach, R.E., Manguin, S., 2005b. Systematics and biogeographical implications of the phylogenetic relationships between members of the Funestus and Minimus Groups of *Anopheles* (Diptera: Culicidae). J. Med. Entomol. 42, 7–16.

Garros, C., Van Bortel, W., Trung, H.D., Coosemans, M., Manguin, S., 2006. Review of the Minimus Complex of *Anopheles*, main malaria vector in Southeast Asia: from taxonomic issues to vector control strategies. Trop. Med. Int. Health 11, 102–114.

Ghosh, S.K., Dash, A.P., 2007. Larvivorous fish against malaria vectors: a new outlook. Trans. R. Soc. Trop. Med. Hyg. 101, 1063–1064.

Ghosh, S.K., Tiwari, S.N., Sathyanarayan, T.S., Sampath, T.R., Sharma, V.P., Nanda, N., Joshi, H., Adak, T., Subbarao, S.K., 2005. Larvivorous fish in wells target the malaria vector sibling species of the *Anopheles culicifacies* complex in villages in Karnataka, India. Trans. R. Soc. Trop. Med. Hyg. 99, 101–105.

Gillies, M.T., 2001. In: Wernsdorfer, W., Mc Gregor, I. (Eds.), Vector Behaviour and Bionomics, Malaria. Churchill Livingstone, Edinburgh, UK, pp. 453–485.

Goswani, G., Raghavendra, K., Nanda, N., Gakhar, S.K., Subbarao, S.K., 2005. PCR-RFLP of mitochondrial cytochrome oxidase subunit II and ITS2 of ribosomal DNA: markers for the identification of members of the *Anopheles culicifacies* complex (Diptera: Culicidae). Acta Trop. 95, 92–99.

Green, C.A., Baimai, V., 1984. Polytene chromosomes and their use in species studies of malaria vectors as exemplified by the *Anopheles maculatus* complex. In: Chopra, V.L., Joshi, B.C., Sharma, R.P., Bansal, H.C. (Eds.), Genetics: New Frontiers Proceedings of the XV International Congress of Genetics. Oxford and IBH Publ. Comp., New Delhi.

Green, C.A., Baimai, V., Harrison, B.A., Andre, R.G., 1985. Cytogenetic evidence for a complex of species within the taxon *Anopheles maculatus* (Diptera Culicidae). Biol. J. Linnean. Soc. 24, 321–328.

Green, C.A., Gass, R.F., Munstermann, L.E., Baimai, V., 1990. Population-genetic evidence for two species in *Anopheles minimus* in Thailand. Med. Vet. Entomol. 4, 25–34.

Green, C.A., Rattanarithikul, R., Ponparit, S., Sawadwongporn, P., Baimai, V., 1991. A newly recognised vector of human malarial parasites in the Oriental region, *Anopheles* (*Cellia*) *pseudowillmori* (Theobald 1910). Trans. R. Soc. Trop. Med. Hyg. 85, 35–36.

Green, C.A., Munstermann, L.E., Tan, S.G., Panyim, S., Baimai, V., 1992a. Population genetic evidences for species A, B, C, and D of the *Anopheles dirus* complex in Thailand and enzyme electromorphs for their identification. Med. Vet. Entomol. 6, 29–36.

Green, C.A., Rattanarithikul, R., Charoensub, A., 1992b. Population genetic confirmation of species status of the malaria vectors *Anopheles willmori* and *An. pseudowillmori* in Thailand and chromosome phylogeny of the Maculatus Group of mosquitoes. Med. Vet. Entomol. 6, 335–341.

Harbach, R.E., 2004. The classification of genus *Anopheles* (Diptera: Culicidae): a working hypothesis of phylogenetic relationships. Bull. Entomol. Res. 94, 537–553.

Harbach, R.E., Gingrich, J.B., Pang, L.W., 1987. Some entomological observations on malaria transmission in a remote village in northwestern Thailand. J. Am. Mosq. Control Assoc. 3, 296–301.

Harbach, R.E., Parkin, E., Chen, B., Butlin, R.K., 2006. *Anopheles* (*Cellia*) *minimus* Theobald (Diptera: Culicidae): neotype designation, characterization, and systematics. Proc. Entomol. Soc. Wash. 108, 198–209.

Harbach, R.E., Garros, C., Manh, N.D., Manguin, S., 2007. Formal taxonomy of species C of the *Anopheles minimus* sibling species complex (Diptera: Culicidae). Zootaxa 1654, 41–54.

Harinasuta, C., Guptavanij, P., Vasuvat, C., 1974. Studies on the medical ecological epidemiology in mangrove areas in Thailand. Southeast Asian J. Trop. Med. Hyg. 32, 588–594.

Harrison, B.A., 1980. The Myzomyia Series of *Anopheles* (*Cellia*) in Thailand, with emphasis on intra-interspecific variations (Diptera: Culicidae) Medical entomology studies-XIII. Contr. Am. Entomol. Inst. 17, 1–195.

Harrison, B.A., Rattanarithikul, R., Peyton, E.L., Mongkolpanya, K., 1990. Taxonomic changes, revised occurrence records and notes on the Culicidae of Thailand and neighboring countries. Mosq. Syst. 22, 196–227.

Hii, J.L.K., 1984. Involvement of the X-chromosome in hybrid male sterility from crosses between species A and species B of the taxon *Anopheles dirus*. Mosq. News 44, 192–196.

Hodgkin, E.P., 1956. The transmission of malaria in Malaya. Stud. Inst. Med. Res. Malaya 27 (viii+), 1–98.

Ikemoto, T., Sumitro, S., Panjaitan, W., Shibuya, T., 1986. Laboratory and small-scale field tests of larvicides and larvivorous fishes against *Anopheles sundaicus* larvae at Perupuk village, North Sumatra. Indonesia. Jpn. J. Sanit. Zool. 37, 105–112.

Ismail, I.A., Phinichpongse, S., Boonrassri, P., 1978. Responses of *Anopheles minimus* to DDT residual spraying in a cleared forested foothill area in central Thailand. Acta Trop. 35, 69–82.

Jaichapor, B., Kengluecha, A., Rongnoparut, P., Rueda, L.M., Sithiprasasna, R., 2005. Morphological variations of *Anopheles minimus* A in Tak province, Thailand. Southeast Asian J. Trop. Med. Hyg. 36, 609–615.

Jambulingam, P., Sahu, S.S., Manonmani, A., 2005. Reappearance of *Anopheles minimus* in Singhbum hills of East-Central India. Acta Trop. 96, 31–35.

Kalra, N.L., 1978. National malaria eradication programme, India—its problems, management and research needs. J. Commun. Dis. 10, 1–20.

Kar, I., Subbarao, S.K., Eapen, A., Ravindran, J., Satyanarayana, T.S., Raghavendra, K., Nanda, N., Sharma, V.P., 1999. Evidence for a new malaria species, species E, within the *Anopheles culicifacies* complex (Diptera: Culicidae). J. Med. Entomol. 36, 595–600.

Kengne, P., Trung, H.D., Baimai, V., Coosemans, M., Manguin, S., 2001. A multiplex PCR-based method derived from random amplified polymorphic DNA (RAPD) markers for the identification of species of the *Anopheles minimus* group in Southeast Asia. Insect Mol. Biol. 10, 427–435.

Kirnowardoyo, S., 1988. *Anopheles* malaria vector and control measures applied in Indonesia. Southeast Asian J. Trop. Med. Public Health 19, 713–716.

Kitthawee, S., Edman, J.D., Upatham, E.S., 1993. Mosquito larvae and associated macroorganisms occurring in gem pits in southern Tha Mai District, Chanthaburi Province Thailand. Southeast Asian J. Trop. Med. Public Health 24, 143–151.

Klinkenberg, E., Konradsen, F., Herrel, N., Mukhtar, M., van der Hoek, W., Amerasinghe, F.P., 2004. Malaria vectors in the changing environment of the southern Punjab. Pakistan. Trans. R. Soc. Trop. Med. Hyg. 98, 442–449.

Kobayashi, J., Nambanya, S., Miyagi, I., Vanachone, B., Manivong, K., Koubouchan, T., Amano, H., Nozaki, H., Inthakone, S., Sato, Y., 1997. Collection of anopheline mosquitoes in three villages endemic for malaria in Khammouane, Lao PDR. Southeast Asian J. Trop. Med. Hyg. 28, 20.

Leicester, G.F., 1903. A breeding place of certain forest mosquitoes in Malaya. J. Trop. Med. 6, 291–292.

Li, S.Z., Ma, Y.J., Zheng, Z.M., 2003. Molecular identification of *Anopheles maculatus* complex (Diptera: Culicidae) from China: single stranded conformation polymorphism (SSCP) analysis assay. Entomotaxonomia 25, 125–130.

Linton, Y.M., Dusfour, I., Howard, T.M., Ruiz, L.F., Duc Manh, N., Ho Dinh, T., Sochanta, T., Coosemans, M., Harbach, R.E., 2005. *Anopheles* (*Cellia*) *epiroticus* (Diptera: Culicidae), a new malaria vector species in the Southeast Asian Sundaicus Complex. Bull. Entomol. Res. 95, 329–339.

Ma, Y., Li, S., Xu, J., 2006. Molecular identification and phylogeny of the Maculatus Group of *Anopheles* mosquitoes (Diptera: Culicidae) based on nuclear and mitochondrial DNA sequences. Acta Trop. 99, 272–280.

Malakar, P., Das, S., Saha, G.K., Dasgupta, B., Hati, A.K., 1995. Indoor resting anophelines of north Bengal. Indian J. Malariol. 32, 24–31.

Manguin, S., Kengne, P., Sonnier, L., Harbach, R.E., Baimai, V., Trung, H.D., Coosemans, M., 2002. SCAR markers and multiplex PCR-based identification of isomorphic species in the *Anopheles dirus* complex in Southeast Asia. Med. Vet. Entomol. 16, 46–54.

Manonmani, A., Townson, H., Adeniran, T., Jambulingam, P., Sahu, S.S., Vijayakumar, T., 2001. rDNA-ITS2 polymerase chain reaction assay for the sibling species of *Anopheles fluviatilis*. Acta Trop. 78, 3–9.

Manonmani, A., Nanda, N., Jambulingam, P., Sahu, S., Vijayakumar, T., Ramya Vani, J., Subbarao, S.K., 2003. Comparison of polymerase chain reaction assay and cytotaxonomy for identification of sibling species of *Anopheles fluviatilis* (Diptera: Culicidae). Bull. Entomol. Res. 93, 169–171.

Meek, S.R., 1995. Vector control in some countries of Southeast of Asia: comparing the vectors and the strategies. Ann. Trop. Med. Parasitol. 89, 135–147.

Mittal, P.K., 2003. Biolarvicides in vector control: challenges and prospects. J. Vector Borne Dis. 40, 20–32.

Moorhouse, D.E., Wharton, R.H., 1965. Studies on Malayan vectors of malaria; methods of trapping and observations on biting cycles. J. Med. Entomol. 1, 359–370.

Mouchet, J., Carnevale, P., 1998. Entomological biodiversity of malaria in the World. Res. Rev. Parasitol. 58, 189–195.

Mouchet, J., Carnevale, P., Coosemans, M., Julvez, J., Manguin, S., Richard-Lenoble, D., Sircoulon, J., 2004. Biodiversité du paludisme dans le monde. John Libbey Eurotext, Paris, France, 428 pages.

Naddaf, S.R., Oshagi, M.A., Vatandoost, H., Djavadian, E., Telmadarei, Z., Assmar, M., 2002. Use of random amplified polymorphic DNA polymerase chain reaction (RAPD-PCR) and ITS2 PCR assays for differentiation of populations and putative sibling species of *Anopheles fluviatilis* (Diptera: Culicidae) in Iran. Iranian J. Public Health 32, 137.

Naddaf, S.R., Oshagi, M.A., Vatandoost, H., Assmar, M., 2003. Molecular characterization of *Anopheles fluviatilis* species complex in the Islamic Republic of Iran. Eastern Mediterranean Health J. 9, 257–265.

Nanda, N., Yadav, R.S., Subbarao, S.K., Joshi, H., Sharma, V.P., 2000. Studies on *Anopheles fluviatilis* and *Anopheles culicifacies* sibling species in relation to malaria in forested hilly and deforested riverine ecosystems in northern Orissa, India. J. Am. Mosq. Control Assoc. 16, 199–205.

Nanda, N., Das, M.K., Wattal, S., Adak, T., Subbarao, S.K., 2004. Cytogenetic characterization of *Anopheles sundaicus* (Diptera: Culicidae) population from Car Nicobar island, India. Ann. Entomol. Soc. Am. 97, 171–176.

Nanda, N., Yadav, R.S., Sahu, S.S., Sharma, S.K., Malhotra, M.S., 2005. *Anopheles fluviatilis* species S-related malaria in hyperendemic districts of Orissa state, India. In: International Conference on Malaria, New Delhi, India, 4–6 November.

Nandi, J., Kaul, S.M., Sharma, S.N., Lal, S., 2000. Anthropophily of anophelines in Duars of west Bengal and other regions of India. J. Commun. Dis. 32, 95–99.

Nguyen Tang Am, Le Quy, R., Vu Thi, H., Ngyuen Bieh, L., 1993. Entomo-epidemiological studies of malaria in the coastal zone of Ho Chi Minh City, 1990–1992. Santé 3, 464–473.

Obsomer, V., Defourny, P., Coosemans, M., 2007. The *Anopheles dirus* complex: spatial distribution and environmental drivers. Malaria J. 6, http://www.malariajournal.com/content/6/1/26.

O'Loughlin, S.M., Somboon, P., Walton, C., 2007. High levels of population structure caused by habitat islands in the malarial vector *Anopheles scanloni*. Heredity 99, 31–40.

Oo, T.T., Storch, V., Becker, N., 2002. Studies on the bionomics of *Anopheles dirus* (Culicidae: Diptera) in Mudon, Mon State, Myanmar. J. Vector Ecol. 27, 44–54.

Oo, T.T., Storch, V., Becker, N., 2004. Review of the anopheline mosquitoes of Myanmar. J. Vector Ecol. 29, 21–40.

Peyton, E.L., Harrison, B.A., 1980. *Anopheles* (*Cellia*) *takasagoensis* Morishita 1946, an additional species in the Balabacensis complex of Southeast Asia (Diptera: Culicidae). Mosq. Syst. 12, 335–347.

Phan, V.T., 1998. Epidémiologie du paludisme et lutte antipaludique au Vietnam. Editions Médicales Vietnam, Hanoi, 241 pages.

Phuc, H.K., Ball, A.J., Son, L., Hanh, N.V., Tu, N.D., Lien, N.G., Verardi, A., Townson, H., 2003. Multiplex PCR assay for malaria vector *Anopheles minimus* and four related species in the Myzomyia Series from Southeast Asia. Med. Vet. Entomol. 17, 423–428.

Pradhan, S.P., Shrestha, S.L., Vaidya, R.G., 1970. Malaria transmission in high mountain valleys of west Nepal including the first record of *Anopheles willmori* (James) as a third vector of malaria. J. Nepal Med. Assoc. 8, 89–97.

Prakash, A., Bhattacharyya, D.R., Mohapatra, P.K., Mahanta, J., 2002. Physico-chemical characteristics of breeding habitats of *Anopheles dirus* (Diptera: Culicidae) in Assam, India. J. Environ. Biol. 23, 95–100.

Raeisi, A., Vatandoost, H., Mousakazemi, S.H., 2005. Understanding bionomics of main malaria vectors in irrigated areas in southern Baluchistan, for implementation of impregnated bednets in Iran. In: International Conference on Malaria, New Delhi, India 4–6 November.

Rahman, W.A., Abu Hassan, A., Adanan, C.R., Rashid, M.R., 1993. The prevalence of *Plasmodium falciparum* and *P. vivax* in relation to *Anopheles maculatus* densities in a Malaysian village. Acta Trop. 55, 231–235.

Rao, T.R., 1984. The Anophelines of India. Malaria Center Research. Indian Council of Medical Research, New Delhi, India.

Rattanarithikul, R., Green, C.A., 1986. Formal recognition of the species of the *Anopheles maculatus* group (Diptera: Culicidae) occurring in Thailand, including the descriptions of two new species and a preliminary key to females. Mosq. Syst. 18, 246–278.

Rattanarithikul, R., Harbach, R.E., 1990. *Anopheles maculatus* (Diptera: Culicidae) from the type locality of Hong Kong and two new species of the Maculatus Complex from the Philippines. Mosq. Syst. 22, 160–183.

Rattanarithikul, R., Harrison, B.A., 1973. An Illustrated Key to the *Anopheles* larvae of Thailand. Bangkok, 42 pages.

Rattanarithikul, R., Mongkolpanya, K., Noigamol, C., Chanaimongkol, S., Mahapibul, P., Nakngen, S., 1994. Dry-season distribution of mosquito larvae in the bed of the Mekong River, northwestern Thailand. J. Am. Mosq. Control Assoc. 10, 197–201.

Rattanarithikul, R., Green, C.A., Panyim, S., Noigamol, C., Chanaimongkol, S., Mahapibul, P., 1995. Larval habitats of malaria vectors and other *Anopheles* mosquitoes around a transmission focus in northwestern Thailand. J. Am. Mosq. Control Assoc. 11, 428–433.

Rattanarithikul, R., Konishi, E., Linthicum, K., 1996a. Observations on nocturnal biting activity and host preference of Anophelines collected in southern Thailand. J. Am. Mosq. Control Assoc. 12, 52–57.

Rattanarithikul, R., Konishi, E., Linthicum, K.J., 1996b. Detection of *Plasmodium vivax* and *Plasmodium falciparum* circumsporozoite antigen in anopheline mosquitoes collected in southern Thailand. Am. J. Trop. Med. Hyg. 54, 114–121.

Rattanarithikul, R., Harrison, B.A., Harbach, R.E., Panthusiri, P., Coleman, R.E., 2006. Illustrated keys to the mosquitoes of Thailand. IV. Anopheles. Southeast Asian J. Trop. Med. Public Health 37 (Suppl 2), 1–128.

Reid, J.A., 1968. Anopheline mosquitoes of Malaya and Borneo. Stud. Inst.Med. Res. Malaya 31.

Reid, J.A., 1970. Systematics of malaria vectors - anopheline systematics and malaria control, with special reference to Southeast Asia. Misc. Publ. Entomol. Soc. Am. 7, 52–62.

Rongnoparut, P., Yaicharoen, S., Sirichotpakorn, N., Rattanarithikul, R., Lanzaro, G.C., Linthicum, K.J., 1996. Microsatellite polymorphism in *Anopheles maculatus*, a malaria vector in Thailand. Am. J. Trop. Med. Hyg. 55, 589–594.

Rongnoparut, P., Ugsang, D.M., Baimai, V., Honda, K., Sithiprasasna, R., 2005. Use of a remote sensing-based geographic information system in the characterizing spatial patterns for *Anopheles minimus* A and C breeding habitats in western Thailand. Southeast Asian J. Trop. Med. Public Health 36, 1145–1152.

Rowland, M., Mohammed, N., Rehman, H., Hewitt, S., Mendis, C., Ahmad, M., Kamal, M., Wirtz, R.A., 2002. Anopheline vectors and malaria transmission in eastern Afghanistan. Trans. R. Soc. Trop. Med. Hyg. 96, 620–626.

Rozendaal, J.A., 1997. Vector Control: Methods for Use by Individuals and Communities. WHO, Geneva, 412 pages.

Sallum, M.A., Peyton, E.L., Wilkerson, R.C., 2005. Six new species of the *Anopheles leucosphyrus* group, reinterpretation of *An. elegans* and vector implications. Med. Vet. Entomol. 19, 158–199.

Sharma, S.K., Nanda, N., Dutta, V.K., 1995. Studies on the bionomics of *Anopheles fluviatilis* s.l. and the sibling species composition in the foothills of Shiwalik Range India. Southeast Asian J. Trop. Med. Public Health 26, 566.

Sharpe, R.G., Harbach, R.E., Butlin, R.K., 2000. Molecular variation and phylogeny of members of the Minimus group of *Anopheles* subgenus *Cellia* (Diptera: Culicidae). Syst. Entomol. 25, 263–272.

Shiff, C., 2002. Integrated approach to malaria control. Clin. Microbiol. Rev. 15, 278–293.

Singh, N., Nagpal, A.C., Saxena, A., Singh, M.P., 2004a. Changing scenario of malaria in central India, the replacement of *Plasmodium vivax* by *Plasmodium falciparum* (1986–2000). Trop. Med. Int. Health 9, 364–371.

Singh, O.P., Chandra, D., Nanda, N., Raghavendra, K., Sunil, S., Sharma, S.K., Dua, V.K., Subbarao, S.K., 2004b. Differentiation of members of the

Anopheles fluviatilis species complex by an allele-specific polymerase chain reaction based on 28S ribosomal DNA sequences. Am. J. Trop. Med. Hyg. 70, 27–32.

Singh, O.P., Goswami, G., Nanda, N., Raghavendra, K., Chandra, D., Subbarao, S.K., 2004c. An allele-specific polymerase chain reaction assay for the differentiation of members of the *Anopheles culicifacies* complex. J. Biosci. 29, 275–280.

Singh, O.P., Chandra, D., Nanda, N., Sharma, S.K., Htun, P.T., Adak, T., Subbarao, S.K., Dash, A.P., 2006. On the conspecificity of *Anopheles fluviatilis* species S with *Anopheles minimus* species C. J. Biosci. 31, 671–677.

Soekirno, M., Bang, Y.H., Sudomo, M., Pemayun, T.P., Fleming, G.A., 1983. Bionomics of *Anopheles sundaicus* and other anophelines associated with malaria in coastal areas of Bali, Indonesia. Document. WHO/VBC/83.885, Geneva.

Somboon, P., Aramrattana, A., Lines, J., Webber, R., 1998. Entomological and epidemiological investigations of malaria transmission in relation to population movements in forest areas of north-west Thailand. Southeast Asian J. Trop. Med. Public Health 29, 3–9.

Somboon, P., Walton, C., Sharpe, R.G., Higa, Y., Tuno, N., Tsuda, Y., Takagi, M., 2001. Evidence for a new sibling species of *Anopheles minimus* from the Ryukyu Archipelago, Japan. J. Am. Mosq. Control Assoc. 17, 98–113.

Subbarao, S.K., 1988. The *Anopheles culicifacies* complex and control of malaria. Parasitol. Today 4, 72–75.

Subbarao, S.K., Sharma, V.P., 1997. *Anopheles* species complexes and malaria control. Indian J. Med. Res. 106, 164–173.

Subbarao, S.K., Vasantha, K., Raghavendra, K., Sharma, V.P., Sharma, G.K., 1988. *Anopheles culicifacies*: sibling species composition and its relationship to malaria incidence. J. Am. Mosq. Control Assoc. 4, 29–33.

Subbarao, S.K., Nanda, N., Vasantha, K., Dua, V.K., Malhotra, M.S., Yadav, R.S., Sharma, V.P., 1994. Cytogenetic evidence for three sibling species in *Anopheles fluviatilis*. Ann. Entomol. Soc. Am. 87, 116–121.

Sucharit, S., Komalamisra, N., Leemingsawat, S., Apiwathnasorn, C., Thongrungkiat, S., 1988. Population genetic studies on the *Anopheles minimus* complex in Thailand. Southeast Asian J. Trop. Med. Public Health 19, 717–723.

Sukowati, S., Baimai, V., Andris, H., 1996. Sex chromosome variation in natural populations of the *Anopheles sundaicus* complex from Thailand and Indonesia. Mosq. Borne Dis. Bull. 13, 8–13.

Sukowati, S., Baimai, V., Harun, S., Dasuki, Y., Andris, H., Efriwati, M., 1999. Isozyme evidence for three sibling species in the *Anopheles sundaicus* complex from Indonesia. Med. Vet. Entomol. 13, 408–414.

Sungvornyothin, S., Garros, C., Chareonviryaphap, T., Manguin, S., 2006a. How reliable is the humeral pale spot for identification of cryptic species of the Minimus Complex? J. Am. Mosq. Control Assoc. 20, 185–191.

Sungvornyothin, S., Muenvorn, V., Garros, C., Prabaripai, A., Bangs, M.J., Manguin, S., Chareonviryaphap, T., 2006b. Trophic behavior and biting activity of the two sibling species of the *Anopheles minimus* Complex in western Thailand. J. Vector Ecol. 31, 252–261.

Surendran, S.N., Ramasamy, R., 2005. Some characteristics of the larval breeding sites of *Anopheles culicifacies* species B and E in Sri Lanka. J. Vec. Borne Dis. 42, 39–44.

Torres, E.P., Salazar, N.P., Belizario, V.Y., Saul, A., 1997. Vector abundance and behaviour in an area of low malaria endemicity in Bataan, the Philippines. Acta Trop. 63, 209–220.

Torres, E.P., Foley, D.H., Saul, A., 2000. Ribosomal DNA sequence markers differentiate two species of the *Anopheles maculatus* (Diptera: Culicidae) complex in the Philippines. J. Med. Entomol. 37, 933–937.

Trung, H.D., Van Bortel, W., Sochanta, T., Keokenchanh, K., Quang, N.T., Cong, L.D., Coosemans, M., 2004. Malaria transmission and major malaria vectors in different geographical areas of Southeast Asia. Trop. Med. Int. Health 9, 230–237.

Trung, H.D., Van Bortel, W., Sochanta, T., Keokenchanh, K., Briet, O., Coosemans, M., 2005. Behavioural heterogeneity of *Anopheles* species in ecologically different localities in Southeast Asia: a challenge for vector control. Trop. Med. Int. Health 10, 251–262.

Upatham, E.S., Prasittisuk, C., Ratanatham, S., Green, C.A., Rojanasunan, W., Setakana, P., Theerasilp, N., Tremongkol, A., Viyanant, V., Pantuwatana, S., Andre, R.G., 1988. Bionomics of *Anopheles maculatus* complex and their role in malaria transmission in Thailand. Southeast Asian J. Trop. Med. Public Health 19, 259–269.

Van Bortel, W., Trung, H.D., Manh, N.D., Roelants, P., Verle, P., Coosemans, M., 1999. Identification of two species within the *Anopheles minimus* complex in northern Vietnam and their behavioural divergences. Trop. Med. Int. Health 4, 257–265.

Van Bortel, W., Trung, H.D., Roelants, P., Harbach, R.E., Backeljau, T., Coosemans, M., 2000. Molecular identification of *Anopheles minimus s.l.* beyond distinguishing the members of the species complex. Insect Mol. Biol. 9, 335–340.

Van Bortel, W., Sochanta, T., Harbach, R.E., Socheat, D., Roelants, P., Backeljau, T., Coosemans, M., 2002. Presence of *Anopheles culicifacies* B in Cambodia established by the PCR-RFLP assay developed for the identification of *Anopheles minimus* species A and C and four related species. Med. Vet. Entomol. 16, 329–334.

Van Bortel, W., Trung, H.D., Roelants, P., Backeljau, T., Coosemans, M., 2003. Population genetic structure of the malaria vector *Anopheles minimus* A in Vietnam. Heredity 91, 487–493.

Vatandoost, H., Shahi, B., Abai, M.R., Hanafi-Bojd, A.A., Oshaghi, M.A., Zamani, G., 2004. Larval habitats of main malaria vectors in Hormozgan province and their susceptibility to different larvicides. Southeast Asian J. Trop. Med. Hyg. 35, 1–4.

Walker, K., Lynch, M., 2007. Contributions of *Anopheles* larval control to malaria suppression in tropical Africa: review of achievements and potential. Med. Vet. Entomol. 21, 2–21.

Walsh, J.F., Molyneux, D.H., Birley, M.H., 1993. Deforestation: effects on vector-borne disease. Parasitology 106 (Suppl), S55–S75.

Walton, C., Handley, J.M., Kuvangkadilok, C., Collins, F.H., Harbach, R.E., Baimai, V., Butlin, R.K., 1999a. Identification of five species of the *Anopheles dirus* complex from Thailand, using allele-specific polymerase chain reaction. Med. Vet. Entomol. 13, 24–32.

Walton, C., Sharpe, R.G., Pritchard, S.J., Thelwell, N.J., Butlin, R.K., 1999b. Molecular identification of mosquito species. Biol. J. Linnean Soc. 68, 241–256.

Walton, C., Somboon, P., O'Loughlin, S.M., Zhang, S., Harbach, R.E., Linton, Y.M., Chen, B., Nolan, K., Duong, S., Fong, M.Y., Vythilingam, I., Mohammed, Z.D., Trung, H.D., Butlin, R.K., 2007. Genetic diversity and molecular identification of mosquito species in the *Anopheles maculatus* group using the ITS2 region of rDNA. Infect. Genet. Evol. 7, 93–102.

WHO, 2007a. Agenda items to be discussed at the Regional Committee: Note for the Media. SEARO, WHO Southeast Asia Region SEA/07-02, http://www.searo.who.int/EN/Section316/Section503/Section2358_13492.htm.

WHO, 2007b. Malaria: Disease burden in SEA Region. SEARO, WHO Southeast Asia Region, http://www.searo.who.int/EN/Section10/Section21/Section340_4018.htm.

WHO, 2007c. Malaria: Goals, objectives and strategies. SEARO, WHO Southeast Asia Region, http://searo.who.int/EN/Section10/Section21/Section335.htm.

WHO, 2007d. Malaria: The Mekong Roll Back Malaria initiative (Proceedings). SEARO, WHO Southeast Asia Region, http://www.searo.who.int/EN/Section10/Section21/Section342_1046.htm.

Williamson, K.B., Scharff, J.W., 1936. Anti-larval sluicing. Malaya Med. J. 11, 123–151.

Zhang, Z., Yang, C., 1996. Application of deltamethrin-impregnated bednets for mosquito and malaria control in Yunnan, China. Southeast Asian J. Trop. Med. Public Health 27, 367–371.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 504–519
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Discussion

Behind the scene, something else is pulling the strings: Emphasizing parasitic manipulation in vector-borne diseases

Thierry Lefèvre *, Frédéric Thomas

GEMI, UMR CNRS-IRD 2724, IRD, 911, av. Agropolis BP 64501, 34394 Montpellier Cedex 5, France

Received 22 March 2007; received in revised form 13 May 2007; accepted 17 May 2007
Available online 21 May 2007

Abstract

Merging the field of epidemiology with those of evolutionary and behavioural ecology can generate considerable fundamental knowledge, as well as help to guide public health policies. An attempt is made here to integrate these disciplines by focusing on parasitic manipulation in vector-borne diseases. Parasitic manipulation is a fascinating strategy of transmission which occurs when a parasite alters phenotypic trait(s) of its host in a way that enhances its probability of transmission. Vector-borne parasites are responsible for many of the most harmful diseases affecting humans, and thus represent public health priority. It has been shown for several decades that viruses, bacteria and protozoa can alter important features of their arthropod vector and vertebrate host in a way that increases their probability of transmission. Here, we review these changes, including, the feeding behaviour, survival and immune system of the vector, as well as attraction, defensive behaviour, blood characterictics and immune system of the vertebrate host. Based on the classic measure of vector-borne disease transmission R_0, additional changes, such as, vertebrate host choice by infected vectors or parasite development duration in the vector are expected. Reported or expected phenotypic changes are discussed in terms of costs and benefits to the parasite, its vector, and the vertebrate host. Introducing the parasitic manipulation concept into vector-borne diseases clearly highlights fruitful avenues not only for fundamental research, but also for developing strategies for disease control.

Keywords: Parasitic manipulation; Vector-borne diseases

1. Introduction

Sit down on the bank of a river in the south of France during the summer and you may witness a spectacular scene: a cricket going straight towards the river and jumping in the water, even tough all his conspecifics remain in the shady wood. Wait 10 s and you may well observe a Gordian worm bursting out of the body of the ''diver cricket'' and swimming into the water, where it will continue its life cycle. The suicidal behaviour of the insect illustrates a case of behavioural manipulation: water seeking is not part of the normal behavioural repertoire of crickets; it is induced by the parasite. The ability of parasites to cause such changes has usually been considered to be adaptive as it enhances host-to-host transmission, or ensures the parasite (or its propagules) gets released in an appropriate location. The scientific literature is now rife with examples of behavioural, morphological and physiological alterations and such parasitic manipulations have been well documented in a variety of host–parasite systems (Fig. 1, Moore, 2002).

To date, many vector-borne parasites have been shown to alter phenotypic traits of their insect vectors and vertebrate hosts in a way that increases contact between them and hence increases the probability of parasite transmission. Emerging and re-emerging infectious vector-borne diseases present one of the most pressing issues facing public health systems. In response, there has been progress in understanding disease transmission, but many challenges remain to be overcome. In this context, studying this intriguing strategy of transmission may help to understand pathogen dynamics and hence bring crucial information on disease control measures.

The first part of this review deals with the concept of parasitic manipulation and discusses the adaptiveness of host behavioural alteration. The subsequent section adresses parasitic manipulation in vector-borne diseases and reviews reported cases of host behavioural and physiological alterations that result, or appear to result, in increased transmission of the parasite. Finally, we end by proposing future avenues of

* Corresponding author. Tel.: +33 467416318; fax: +33 467416299.
E-mail address: thierry.lefevre@mpl.ird.fr (T. Lefèvre).

1567-1348/$ – see front matter
doi:10.1016/j.meegid.2007.05.008

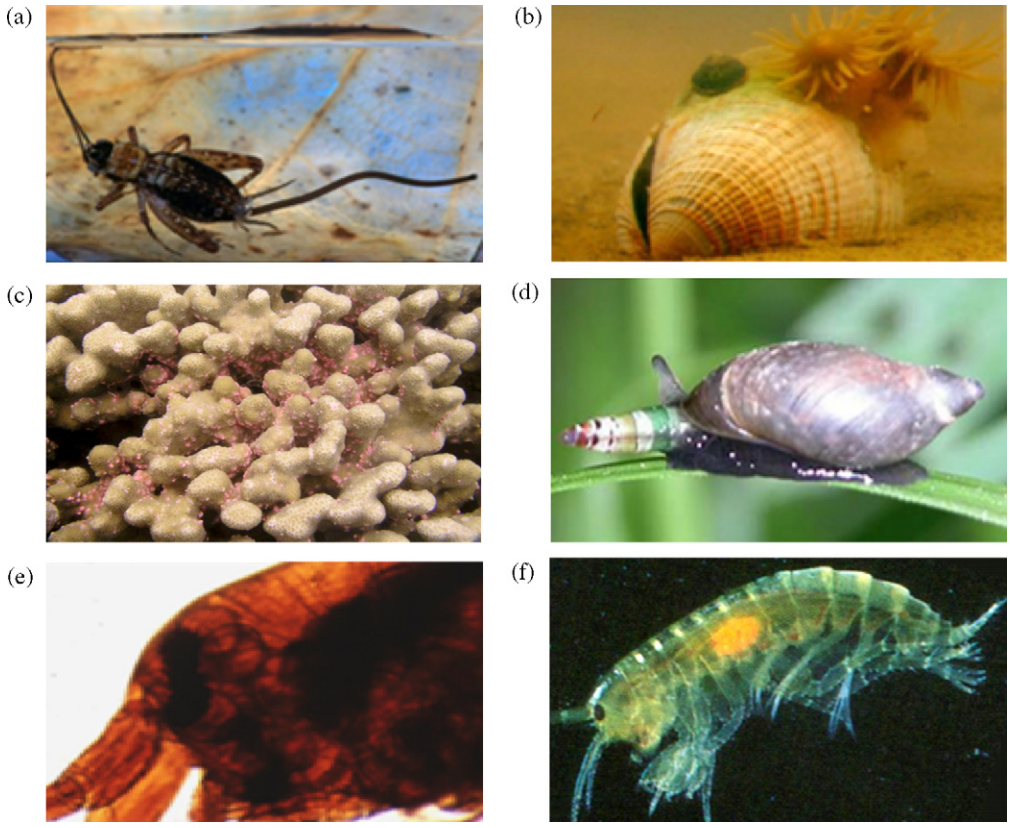


Fig. 1. Examples of parasitic manipulation. (a) The Gordian worm (*Paragordius tricuspidatus*) exiting the body of a cricket (*Nemobius sylvestris*) (Photo P. Goetgheluck). (b) A cockle (*Austrovenus stutchburyi*) parasitized by a trematode (*Curtuteria australis*). The parasites impair the natural burrowing behaviour of their cockle intermediate hosts, making them more likely to be eaten by aquatic bird (final host) (Photo F. Thomas). (c) Coral polyps infected with a trematode (*Podocotyloides stenometra*). The parasite induces pink, swollen nodules on the coral colony and impairs their retraction ability. Infected polyps are, therefore, both conspicuous and vulnerable to predation by the coral-feeding butterflyfish, *Chaetodon multicinctus* (parasite's definitive host) (Photo G. Aeby). (d) The sporocysts of *Leucochloridium paradoxum* develops in the snail's tentacles, where they can be seen conspicuously pulsating. This make the tentacles look to a bird, definitive host, like caterpillars (Photo P. Vogel). (e) Metacercariae of *Microphallus papillorobustus* encyst in the brain of the intermediate host, *Gammarus insensibilis* (Photo F. Thomas). (f) Cystacanth of *Polymorphus minutus* encysts in the body cavity of *Gammarus pulex* (Photo F. Cézilly). Infected *G. pulex* and infected *G. insensibilis* show similar behavioural changes (i.e. negative geotaxis, aberrant evasive behaviour and in infected *G. insensibilis* a positive phototaxis). Since both parasites are phylogenetically distant, these changes can be considered as a case of convergence.

research in order for this topic to prosper in the emerging field of evolutionary epidemiology.

2. The 'spread' of parasitic manipulation

The field of host manipulation by parasites acquired a conceptual framework with the Richard Dawkins' book entitled 'The Extended Phenotype' published in 1982. In this book, Dawkins proposes a vision of life where the unit of natural selection is the gene and where its influence can extend outside its 'vehicle', i.e. the body in which the gene resides. For instance, a bird's nest is a phenotype like any other, under the control of the bird's genes in the same way as the shape of its beak is under the control of its genes. This vision has helped scientists to establish a framework for research on parasitic manipulation of infected hosts. Henceforth, host alteration may be regarded as the expression of the genes of the parasite in the host phenotype. In this view, genes of the parasite are selected for their effect on host phenotype. For example, in behavioural manipulation, one could expect that the parasite secretes chemicals (or induces their production in the host) acting on the central nervous system of the host to induce altered behaviour. Recently, it has been shown that the Gordian worm mentioned earlier alters the normal functions of the grasshopper's central nervous system by producing certain ''effective'' molecules. In particular, a protein from the Wnt family was found that act directly on the development of the central nervous system. These proteins show important similarities with those known in insects, suggesting a case of molecular mimicry (Biron et al., 2005).

Dawkins' theoretical explanation, namely, that host behavioural changes can be parasite adaptations, has led researchers to consider all behavioural changes observed in an infected organism as beneficial for the parasite. However, other explanations are possible. First it has been proposed that the host may benefit from these changes. For instance, suicidal behaviour may reduce the risk of infection for the host's kin. Similarly, behavioural fever and elevated body temperature may help the host to eliminate its parasite (Hart, 1988) in raising its body temperature. Alternatively, changes might be pathological consequences of infection, adaptive to neither host nor parasite, and that not all behavioural modifications

following infection are mediated by parasites. Indeed, not all behavioural modifications have evolved to increase the reproductive success of the parasite.

Since it is difficult to distinguish between these different possibilities, four criteria have been proposed by Poulin (1995) to consider changes as adaptive in the context of transmission: complexity; purposiveness of design (i.e. conformity between a priori design and the alterations); convergence (similar changes in several independent lineages) and fitness consequences. This important paper has helped to highlight the need for a novel and more rigorous approach to studying and interpreting potential cases of parasitic manipulation. The first criterion, i.e. complexity, is perhaps the least useful since it is difficult to judge if a trait is complex or not. Simple host behavioural changes may indeed rely on complex physiological modifications induced by the parasite. Natural selection does not always favour complex solutions over simple ones; it favours the solutions that work. Concerning the second criterion (i.e. purposiveness of design), the host behavioural changes are indeed often exactly those we could expect to enhance parasite transmission. This aspect is well illustrated when looking at different types of parasite life cycles. First, in direct life cycles, host phenotypic alterations leading to increased contacts among conspecifics are likely to be selected. For instance, the rabies virus, Hantavirus, and borna disease virus induce increased aggression and physical contacts in the host population. Second, in complex life cycles involving trophic transmission, many parasites alter the behaviour of their intermediate host in a way that increases the risk of them being preyed upon by final hosts, by, for example, changing intermediate host microhabitat or inhibiting anti-predatory behaviour. The third criterion proposed to assess the adaptiveness of behavioural alteration is convergence. When a similar phenotypic trait in several organisms has not been inherited from their most recent common ancestor but conversely has independently evolved, the trait is said to be a homoplasy or a convergence. A convergent trait is likely to be an adaptation since it results from the same selective pressures acting on phylogenetically distant organisms. For instance, three species of parasites (i.e. two acanthocephalans, *Polymorphus minutus* and *P. paradoxus* and a trematode, *Microphalus papillorobustus*) are known to induce in their crustacean gammarids (intermediate hosts) similar aberrant evasive behaviour making them more prone to be eaten by aquatic birds (definitive hosts) (Bethel and Holmes, 1973; Helluy, 1984; Cézilly et al., 2000). As parasitism has independently evolved in *Trematoda* and *Acanthocephala*, the induced aberrant evasive behaviour can be seen as convergent (Fig. 1). Conversely, between the two acanthocephalan species it is likely that the ability to induce this aberrant evasive behaviour was inherited from the common ancestor; it is thus less likely to be convergence than in the latter case. The last criterion proposed, undoubtedly the most convincing evidence in favour of adaptation, is the demonstration of fitness benefits. For several decades researchers have attempted to show that behavioural modifications result in more successful transmission. Linking behavioural modification with increased probability of transmission is difficult, that is why the case of trophic transmission by intermediate hosts is surely the easiest to investigate. The first study considering this conjecture was that carried out by Holmes and Bethel (1972). In this study, a predator (definitive host) was experimentally offered known numbers of infected and uninfected gammarids. At the end of the experiment, numbers of infected and uninfected individuals that survived were counted. Holmes and Bethel showed that the behavioural alterations induced by the parasite lead to a significantly increased predation rate of infected gammarids. Since results obtained in the laboratory may not reflect precisely that which occurs naturally, a second method has been devised by researchers in attempt to show that behavioural alteration actually increases transmission rates. In this approach, the prevalence of infected individuals in the intermediate host population is compared to the prevalence of infected individuals in the final host's diet. An over-representation of infected preys in the gut contents of an appropriate definitive host is considered as evidence for an increased rate of transmission. For instance, Moore (1983) showed that terrestrial isopods parasitized by an acanthocephalan were more active compared to uninfected ones, and were consequently more likely to be found in the diet of the definitive host, the starling.

The publication of Poulin's critical review (1995) has marked the start of a new period during which many studies taking into account the previous recommendations appeared in the scientific literature. This paper has nonetheless contains one obscure point: should we consider the changes due to pathological consequences of infection and that are coincidentally beneficial for the parasite as adaptations? Poulin answers no: ''Experiments can also serve to distinguish between behavioural modifications that are truly adaptive and those that are coincidentally beneficial. For instance, parasitized hosts can make themselves more visible to predators because they have increased energy requirements and must forage more, not because they are manipulated by the parasite''. In this case, even if the link between behavioural alterations and increased rate of transmission is demonstrated, we cannot consider the changes as adaptive. Another situation illustrating this restriction corresponds to what is called ''fortuitous payoff of other adaptations''. For example, the eye fluke trematode inducing blindness in his intermediate host fish could have been selected to encyst in the eye of the fish to avoid the host immune system, affording later the ability to alter the host evasive behaviour. In this case, the manipulation can be regarded as the consequence of traits that have evolved for other purposes. This definition clearly looks like the definition of exaptation sensu Gould and Vrba (1982): an exaptation is a trait that evolved for other functions, or no function at all, but which has been co-opted for a new use. Concerning this point, it is also possible that more than one advantage to the eye fluke trematode could immediately occur from encysting in the eyes and that selection pressures for this phenotype will be much stronger if several benefits ensue (i.e. increased transmission and immune system avoidance). Unfortunately, we often take a simplistic approach by only looking at a single consequence of an action at a time.

As for the confusion between these ''by-products'' (i.e. changes coincidentally beneficial and fortuitous payoff) and adaptation, a review written by Reeve and Sherman (1993)

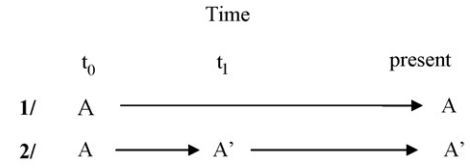


Fig. 2. Two evolutionary sequences of a trait illustrating the confusion concerning the label "exaptation". (modified after Reeve and Sherman, 1993).

helps to clarify the issue. This paper deals with the slippery concept of adaptation. Reeve and Sherman stressed the fact that it is dangerous to confuse the product of the selective process and the process itself. For instance, exaptation and adaptation are actually similar, because both refer to traits that exist because they have been, and continue to be favoured over alternative traits. Do host behavioural modifications result from energy depletion, from the consequences of another adaptation, or from a true manipulation by parasites? Let us imagine two evolutionary sequences for a trait (Fig. 2). In the first sequence, the present trait with function A would be interpreted as an adaptation, i.e. as a trait selected for the purpose it has served in the past and continues to serve. In the second sequence, a trait having a new function *A*' would be an adaptation if we examine the trait between t_1 and the present time, but it would be considered as an exaptation if we examine the trait between t_0 and t_1. Thus, since the label "exaptation" depends on the point in evolutionary history at which we investigate the trait's function, this concept generates ambiguity, rather than clarification. Moreover, concerning our trait of interest, that is, host behavioural modification, it is virtually impossible to identify its original role (no fossil records). To circumvent these problems, Reeve and Sherman (1993) proposed an operational, non-historical definition of adaptation, particularly "adapted" to behavioural ecologists interested in questions of phenotype existence in the field of parasitic manipulation: "an adaptation is a phenotypic variant that results in the highest fitness among specified set of variants in a given environment". Considering this definition, all hosts' behavioural modifications that lead to an enhancement of parasite transmission are maintained by natural selection and thus are a parasite adaptation. To return to the "by product" hypothesis, being more visible to predators because of an increased energy requirement would be an adaptation under this definition if it increases transmission towards a definitive host and the parasite's reproductive success, when compared to an infected host no more visible to predators. For the moment no studies provide information concerning the link between variability in manipulative effort and an increased rate of transmission (and subsequently the parasite's reproductive success), but one could expect that, all else being equal, a parasite that manipulates more than other parasites increases the probability of being transmitted. It could be possible to experimentally select over several generations the phenotypes of parasites with high, medium and low manipulative effort and to examine the rates of transmission between the selected phenotypes.

But why does the host not resist and why does it act in a way that favours the parasite? Behavioural modification must be seen as a co-evolved trait, i.e., a phenotype shared by the host and his parasite. The host is under strong selective pressures to resist manipulation and the parasite is under strong selective pressure to manipulate. The changes might thus be interpreted as an equilibrium state, nonetheless often biased toward the parasite (e.g. in the case of trophic transmission). Understanding why the benefits are often balanced in favour of parasites is easy when highlighted by the life—dinner principle proposed by Dawkins and Krebs (1979). This principle states that the rabbit runs faster than the fox because the rabbit is running for his life, while the fox is only running for his dinner. Thus, the consequences of failure are asymmetric for those involved: one has more to lose than the other. We can apply this principle to the case of parasitic manipulation (Barnard, 1990; Poulin et al., 1994). Transmission is always a matter of life and death for the parasite, while parasitic manipulation does not necessarily have such drastic consequences for the host. The host is descended from a long line of ancestors, only a proportion of which ever met a manipulator parasite. Moreover, some of the individuals subjected to parasitic manipulation may have already reproduced and others may survive to reproduce (depending on the type of host–parasite interaction). Host genes coding for a lack of resistance can thus be transmitted to the next generation. In contrast, every single parasite alive today is descended from a long line of ancestors who have all negotiated successfully the passage between two suitable successive hosts. Therefore parasite genes coding for an effective transmission, e.g. manipulation, have a greater chance to be passed on to the next generations. This reasoning helps us to understand why the host is more likely to capitulate. The stable compromise resulting from this arms race is dynamic and depends on many parameters, such as the prevalence of the parasite, the cost of host resistance, the cost of being manipulated, the cost of parasite manipulation and the cost of the parasite's failure.

3. Parasitic manipulation and vector-borne diseases

Haematophageous insects, when feeding on their host can transmit numerous blood pathogens. An increasing number of studies demonstrate, or suggest, that vector-borne parasites are able to manipulate several phenotypic traits of their vertebrate hosts and vectors in ways that render parasite transmission more probable (Molyneux and Jefferies, 1986; Moore, 1993, 2002; Hurd, 2003; Lacroix et al., 2005). The basic case reproduction number, R_0, the fundamental and classic measure of infectious disease transmission (Macdonald, 1957) can help establish the framework for research on parasitic manipulation in vector-borne diseases. R_0 describes the average number of new infections arising from a single current infection, for vector-borne disease: $R_0 = ma^2p^n/-r \ln p$ (Dye, 1992; Garrett-Jones and Shidrawi, 1969).

Thus, the parasite transmission depends on the number of insect vectors per individual vertebrate host (m), the daily biting rate of an individual vector on vertebrate hosts (a) (once to become infected, and once to transmit, hence a^2), the daily survival rate (p), the number of days required for the parasite to develop (n, i.e. the extrinsic latent period), and r the daily

recovery rate of infected hosts (note: the value $ma^2p^n/-\ln p$ is the vectorial capacity (C), thus $R_0 = C/r$ and the value $p^n/-\ln p$ describes the expectation of infective life of the vector population; Garrett-Jones and Shidrawi, 1969). Generally, methods of estimating these parameters assume that uninfected and infected vectors are similar. However, evidence suggests that this assumption may not be valid and the influence of parasitism on host phenotype is well documented (Combes, 2001; Moore, 2002). Since the R_0 equation gives us a checklist of the key transmission components, it can serve as a basis to determine which traits are likely to be parasitically modified in vertebrate and insect hosts. In this section, we briefly review reported cases of physiological and behavioural modifications that indeed increase, or appear to increase, parasite transmission. Examples are drawn from field and laboratory studies and include a wide range of parasites (viruses, protozoa, bacteria), arthropod vectors (mosquitoes, sandflies, tsetse flies, fleas, bugs, ticks) and vertebrate host species (humans, mice, oxen, lizards, birds).

3.1. The biting rate: 'be aggressive'

According to this equation, parasite transmission is particularly sensitive to biting rate (a^2). It is then not so surprising that in many systems where pathogens are transmitted by bloodfeeding insects, an increased biting rate has been often reported (Table 1). Such a behavioural alteration

Table 1
Reported cases of behavioural and physiological alterations in the arthropod vector

Physiological and behavioural alterations	Host–parasite systems	References
Feeding behaviour: increased biting rate (BR) and increased biting duration (BD)	Virus–Mosquito	
	LaCrosse virus–Aedes triseriatus (BR)	Grimstad et al. (1980)
	Dengue 3–Aedes aegypti (BD)	Platt et al. (1997)
	Malaria–Mosquito	
	P. gallinaceum–A. aegypti (BR) (BD)	Rossignol et al. (1984, 1986), Koella et al. (2002)
	P. yoelii nigeriensis–Anopheles stephensi (BR)	Anderson et al. (1999)
	P. falciparum–An. gambiae (BR) (BD)	Wekesa et al. (1992)
	P. falciparum–An. gambiae (BR)	Koella et al. (1998)
	Plasmodium spp.–*An. punctulatus* (BR) (BD)	Koella and Packer (1996)
	Trypanosomes–Bugs	
	T. rangeli–Rhodnius prolixus (BR) (BD)	D'Alessandro and Mandel (1969), Anez and East (1984), Garcia et al. (1994)
	T. cruzi–Mepraia spinolai (BR)	Botto-Mahan et al. (2006)
	Bacteria–Flea	
	Yersina pestis–Xenopsylla cheopis (BR)	Bacot and Martin (1914)
	Leishmania–Sand fly	
	Leishmania mexicana–Lutzomyia longipalpis (BR)	Killick-Kendrick et al. (1977)
	Leishmania major–Phlebotomus duboscqi (BR) (BD)	Beach and Leeuwenberg (1985)
	African Trypanosomes–Glossina	
	Trypanosoma brucei–G. morsitans morsitans (BR) (BD)	Jenni et al. (1980)
	T. congolense–G. morsitans morsitans (BR) (BD)	Roberts (1981)
	Mite–Nematode	
	Litomosoides carinii–Ornithonyssus bacoti (BD)	Jefferies (1984)
Lifespan: increased (IL) and same (SL)	Malaria–Mosquitoes	
	Plasmodium falciparum–Anopheles. gambiae (SL)	Chege and Beier (1990), Robert et al. (1990), Hogg and Hurd, 1997
	P. falciparum–An. funestus (SL)	Chege and Beier (1990)
	P. falciparum, *P. vivax–An. tessellatus* (SL)	Gamage-Mendis et al. (1993)
	P. vivax–An. maculipennis (SL)	Sinton and Shute (1938)
	P. gallinaceum–A. aegypti (SL)	Freier and Friedman (1987), Rossignol et al. (1986)
	African Trypanosomes–Glossina	
	Trypanosoma gambiense–Glossina palpalis (IL)	Duke (1928)
	T. Rhodesiense, *T.brucei–G. morsitans* (IL) (SL)	Baker and Robertson (1957) Maudlin et al. (1998)
	T. vivax, *T. congolense*, *T. brucei–G. morsitans morsitans* (SL)	Moloo and Kutuza (1985)
	Protozoa–Ticks	
	Babesia microti–Ixodes trianguliceps (IL)	Randolph (1991)
Temperature preference	Malaria–Sand fly: *Plasmodium mexicanum–Lutzomyia vexator*	Fiahlo and Schall (1995)
Immune response	Malaria–Mosquito	
	P. gallinaceum–Aedes. aegypti	Boëte et al. (2004)
	P. falciparum–An. gambiae	Lambrechts et al. (2007)

has indeed been shown for tsetse flies infected with African trypanosomes, in bugs infected with *Trypanosoma* spp., in sandflies infected with *Leishmania* spp., in fleas infected with the plague bacterium, in mosquitoes infected with *Plasmodium* spp. and viruses (Molyneux and Jefferies, 1986; Moore, 2002; Hurd, 2003). Interestingly, in almost all cases, the infective stages of the parasites appear to interfere with the ingestion process of the insects by, for example, obscuring phagoreceptors (e.g. trypanosome-infected tsetse flies), blocking the foregut (e.g. leishmania-infected sandflies) and reducing apyrase activity in salivary glands (e.g. malaria-infected mosquitoes). These different mechanisms seem to impair the vector's ability to fully engorge and therefore induce them to bite vertebrate hosts several times.

An increased biting rate was first observed by Bacot and Martin (1914) in fleas parasitized with the plague bacillus, *Yersina pestis*. In this system, infected fleas probe more often and sometimes without ingesting blood (Bibikova, 1977). The bacteria multiply in the midgut and proventriculus causing occlusion of the proventriculus and blockage of the gut. The block prevents the flow of host blood from the foregut into the midgut. As a consequence, to obtain a blood meal, a blocked flea regurgitates the plug and injects it into the host (Hinnebusch et al., 1998).

Infected sandflies have also difficulty in feeding and hence will often bite a host several times before a blood meal is taken (Table 1, Killick-Kendrick et al., 1977; Beach and Leeuwenberg, 1985). In this system, the parasite produces a gel-like plug, the promastigote secretory gel (PSG) that blocks the foregut (Rogers et al., 2002; see also Volf et al., 2004). Rogers et al. (2004) demonstrated that the parasite, *Leishmania mexicana* accumulates behind the plug and are actively regurgitated during biting. An average of 1086 parasites was egested per bite, and 86–98% of them were infective metacyclic promastigotes (Rogers et al., 2004). In addition, this study sheds light on an important structural component of the PSG, the filamentous proteophosphoglycan (fPPG), which beyond its role in blocking the foregut, actively participates in the success of vertebrate host infection. First, by comparing infections caused by a single fly bite with those caused by a syringe inoculation, Rogers et al. (2004) found that an exacerbation factor, facilitating mice infections, was egested along with the plug containing metacyclic promastigotes. Then, they demonstrated that this factor was of parasite origin, and finally that egested fPPG was responsible for this infection exacerbation.

In tsetse flies parasitized with trypanosomes, Jenni et al. (1980) showed that *Glossina morsitans morsitans* and *G. austeni* infected with *Trypanosoma brucei* probed three times more often and fed more voraciously than uninfected flies. Working on *G. m. morsitans* infected with *T. congolense*, Roberts (1981) reported that infected flies probed significantly more frequently and took longer to engorge than uninfected flies. The increased probing may be caused by physical interference of the parasite with phagoreceptors in the tsetse fly labrum (Thevenaz and Hecker, 1980; Livesey et al., 1980; Molyneux and Jenni, 1981). However, such findings have not been confirmed by studies carried out by Moloo's group (Moloo, 1983; Moloo and Dar, 1985; Makumi and Moloo, 1991). Combinations of vectors and parasites from different locations and/or parasite intensities may help explain these contradictory results.

Malaria-infected mosquitoes bite their vertebrate hosts longer (Rossignol et al., 1984, 1986; Wekesa et al., 1992), more often (Rossignol et al., 1986; Wekesa et al., 1992; Koella et al., 1998, 2002, but see Li et al., 1992) and are more persistent in seeking out blood meals (Koella and Packer, 1996; Anderson et al., 1999) (Table 1). These three aspects of feeding appear to be crucial for parasite transmission since it arouses vertebrate host defensive behaviour, leading to interrupted feeding, which in turn, could multiply host contacts per gonotrophic cycle (Rossignol et al., 1986; Wekesa et al., 1992). Disruption of feeding activity, and hence multiple biting, occur in this system probably as a result of altered apyrase activity. Apyrase is a salivary enzyme that inhibits platelet aggregation, facilitates blood vessel location and thus promotes blood-feeding by the insect (Ribeiro et al., 1984). For instance, in malaria-infected *Aedes aegypti*, apyrase activity is reduced to a third following maturation of sporozoites (i.e. the mature transmissible stage of *Plasmodium* spp.) and results in a longer biting period (Rossignol et al., 1984, see also Ribeiro et al., 1985).

Triatomines infected with *Trypanosoma* spp. also experience difficulties in engorging (Tobie, 1965; D'Alessandro and Mandel, 1969; Schaub, 1989; Garcia et al., 1994). It has been shown that the biting rate of *Rhodnius prolixus* and *R. robustus* was increased by infection with *Trypanosoma rangeli* (D'Alessandro and Mandel, 1969; Anez and East, 1984; Garcia et al., 1994) and that it probably results from reduced apyrase activity (Garcia et al., 1994; Azambuja and Garcia, 2005). In this system, it has been also proposed that the sluggish movements observed in infected bugs might be parasitic manipulation of locomotory activity aiming at facilitate the parasite transmission by predation and/or cannibalism (Schaub, 2006). Recently, Botto-Mahan et al. (2006) demonstrated that *Trypanosoma cruzi*, transmitted through defecation and responsible for Chagas disease, manipulates the behaviour of the kissing bug vector, *Mepraia spinolai*, in two ways. First, it induces an increased biting rate and second it reduces the defecation time after feeding. As a consequence, parasite transmission may be enhanced because more bites induce more wounds for parasite contamination from faeces deposited shortly afterwards (Botto-Mahan et al., 2006).

Reduced feeding success, increased biting rates and their duration have also been reported in *Aedes triseriatus* infected with La Crosse virus (Grimstad et al., 1980) and in *A. aegypti* infected with dengue virus (Platt et al., 1997, but see Putnam and Scott, 1995). Although no mechanism has been demonstrated, Grimstad et al. (1980) and Platt et al. (1997) suggested that infection of the salivary gland coupled with heavy infection of the nervous system, eyes and abdominal ganglion (all involved in host seeking and blood feeding processes) may provide proximal explanations of altered feeding behaviour in virus-infected mosquitoes.

Following the examples of feeding alterations described above, two striking features must be emphasized. First, behavioural changes often occur with precise timing. Altered

feeding behaviour indeed often takes place only when parasites are fully developed, i.e. once the parasites reach the infective stage for the next host. For instance, Koella et al. (2002) (see also Anderson et al., 1999; but see Ferguson and Read, 2004) showed that *P. gallinaceum* manipulates the mosquito vector, *A. aegypti,* in two different ways and in a stage-specific manner: when ready to be transmitted to the vertebrate hosts (sporozoite stage), the parasite increases the biting rate of his vector. In contrast, at an earlier developmental stage (oocyst) which is not transmissible to the vertebrate hosts, the parasite decreases the contact between vector and vertebrate hosts, by decreasing the natural host seeking behaviour of the insect. Since biting is risky and could lead to the death of the insect, this change seems to be beneficial for the parasite. Beach and Leeuwenberg (1985) also observed that uninfected sandflies and sandflies with immature infections probe only once or twice and are fully engorged within 10 min, whereas mature-infected flies probe more often and never result in complete engorgement.

Second, it appears that vector-borne parasites have evolved slightly different mechanisms (i.e. blocking the foregut, reducing apyrase, obscuring phagoreceptors) aimed at inducing similar feeding alterations (i.e. interference in ingestion process leading to new feeding attempts). Poulin (1995) (see also Thomas and Poulin, 1998) pointed out that convergence in manipulative processes is a likely scenario when host alterations bare on the same function, but are derived from different proximal mechanisms. Both the timing of feeding alteration and its independent evolution in phylogenetically distant organisms (i.e. convergence) are strong arguments for parasitic manipulation. Unfortunately empirical evidence in support of increased transmission induced by increased biting rate is still lacking.

3.2. Longevity and survival rate: 'have a long life'

The R_0 equation describes that the longer vectors live, the more the parasite can be transmitted. Life history theory suggests that a reduction or suppression of reproductive effort will result in increased lifespan (Stearns, 1992). Thus, it has been suggested that vector-borne parasites should be able to manipulate resource allocation of their insect vectors in a way that changes the optimum trade-off between reproduction and longevity, which in turn, could favour longer vector survivorship and hence the parasite's overall transmission (Hamilton and Hurd, 2002). Globally, studies on the effects of infection on survival reported no significant lifespan differences between infected and uninfected vectors. However, studies on the effects of infection on fecundity (i.e. number of produced gametes) often report a reduction in fecundity. Since parasites use vector resources to develop and reproduce, it has been suggested that the observed fecundity reduction in infected vectors does not induce a longer lifespan but the same lifespan as uninfected insects (Hamilton and Hurd, 2002).

Parasite-induced fecundity reduction has been reported in many insect vector—parasite associations (Hurd et al., 1995; Hurd, 2003). The association between malaria parasites and mosquitoes provides the most reported case of parasite-induced fecundity reduction. Potential proximate causes for this phenomenon in malaria have been extensively reviewed in Hurd (2001, 2003) and will not be developed here.

Whereas parasite-induced fecundity reduction of insect hosts seems to be a widespread phenomenon, lifespan comparisons of infected and uninfected insect vectors have provided much more conflicting results. Some studies found reduced vector lifespan, while others showed no effect of infection, and even increased lifespan (Table 1). For instance, Ferguson and Read (2002) used a meta-analysis on several published laboratory studies to demonstrate that overall, malaria does reduce mosquito survival suggesting that no such manipulation is occurring in this system. However, they also showed that morbidity effects are more likely to be found in unnatural vector–parasite combinations and in studies of longer duration. As a consequence, when considering studies carried out on natural associations, no effect of infection on longevity is found. Concerning tsetse flies, sandflies and bugs, the picture that emerges is also far from conclusive. For instance, greater longevity of trypanosome-infected tsetse flies has been reported by Duke (1928) and Baker and Robertson (1957). In contrast, Moloo and Kutuza (1985) found no difference in longevity and fecundity between infected and uninfected individuals, while Makumi and Moloo (1991) reported a higher longevity for tsetse males but the reverse for females, and no effect of infection on fecundity. Comparison of the survival distributions of uninfected glossina with those exposed to infection with either *Trypanosoma congolense* or *T. brucei* showed that the first significantly reduced glossina survival, while the second had little or no effect on the survival (Maudlin et al., 1998). In the sandflies, *Phlebotomus papatasi* and *P. langeroni*, infection with *Leishmania major* and *L. infantum* results not only in a significant reduction in fecundity but also in longevity (El Sawaf et al., 1994). Such findings are consistent with those reported for Rocky Mountain wood ticks, *Dermacentor andersoni*, when infected with *Rickettsia rickettsii*, the causative agent of Rocky Mountain spotted fever (Niebylski et al., 1999). Reduced survival has also been found for ticks infected with several protozoa (e.g. Gray, 1982; Watt and Walker, 2000; but see Randolph, 1991), in mosquitoes infected with several viruses (e.g. Faran et al., 1987; McGaw et al., 1998; Scott and Lorenz, 1998), and finally in mosquitoes infected with filariasis (Krishnamoorthy et al., 2004).

As conditions in the laboratory may not reflect precisely those which occur in the field, the influence of parasitism on survival and longevity has not yet been well addressed. Insect vectors are often naturally exposed to adverse conditions which are rarely considered in laboratory studies, and might act synergistically with parasitism. In addition, one might assume that the optimal survival rate of the insect may differ between parasite developmental stages: immature stages are expected to increase the vector's survival to increase their chance of becoming mature, while mature stages are potentially confronted with a trade-off between increasing the vector's lifespan and thus survival rate and increasing biting rate (see Section 3.1) that may decrease the survival rate since biting is risky *in natura* (Koella, 1999; Schwartz and Koella, 2001). For

instance, it has been shown under natural conditions that sporozoite-infected mosquitoes have higher feeding associated mortality than those without sporozoites (Anderson et al., 2000). This finding suggests that malaria parasites might harm their mosquito host more than usually thought.

Unfortunately, many studies exploring the influence of parasitism have concentrated only on either vector reproduction or vector survival/longevity. As a consequence, studies on the links between infection, vector reproduction and vector longevity/survival are still rare (but see Makumi and Moloo, 1991; Hogg and Hurd, 1997) and clearly more investigation is needed to fully explore this interesting hypothesis. In addition, a better understanding of the effects of infection on longevity and fecundity will be achieved after more investigations of naturally infected vectors.

3.3. The parasite development duration: 'become infective early'

In vector-borne parasites, one might expect that natural selection will favour an optimal developmental schedule for each parasite stage that ensures transmission between successive hosts. Once in the insect vector, a major challenge facing the parasite is to reach its infective stage before the insect takes its last blood meal and dies. This idea is particularly valid in systems involving short-lived vectors, such as, mosquitoes and sandflies. In this context, shortening the duration of development should permit the parasite to become infective earlier and hence to increase its probability of transmission. So, why is this period so long? In many vector-borne pathogens, this period is indeed as long as their insect vector's average lifespan. For instance, estimates for the probability that a 'newly' infected mosquito will live long enough to transmit *Plasmodium* spp. range from 80% (Macdonald, 1956) to less than 10% (e.g. Rodriguez et al., 1992). In fact, this duration probably results from constraints and/or trade-offs aimed at increasing the overall parasite transmission.

In malaria–mosquito systems, the vector injects a very small number of sporozoites (i.e. 10–20) into the host compared to the huge quantity available in the salivary gland (Paul et al., 2003). Koella (1999) thus asked: ''if only a few (sporozoites) are necessary, should evolutionary pressures not lead the parasite to produce the sporozoites earlier, if doing so would increase the likelihood that mosquito survives the developmental period?'' At a first glance, we can suggest that it results from a constraint, that is, a large number of sporozoites are needed in salivary glands to enable the injection of only a few. However, it has been demonstrated that this large number is neither correlated to the number of those injected nor to the probability of successfully infecting the host (Ponnudurai et al., 1991; Beier et al., 1991). As a consequence, knowing that the vector's biting rate increases with the sporozoite load (Koella et al., 1998), Koella (1999) proposed a model to investigate whether the developmental pattern could be, in fact, a mechanism that increases the overall parasite's transmission. The model showed that if mortality increases no more than linearly with biting rate, maximal parasite transmission success can occur when all sporozoites are released from the oocyst simultaneously. Consequently, Koella (1999) concluded that the development time of *Plasmodium* within its insect may be a mechanism for manipulating vector-biting behaviour to increase its transmission. Atlhough interesting, further work is necessary to fully clarify this point.

Another argument in favour of an adaptive development period over constrained development period is the case of *Plasmodium mexicanum*. In Central America, this parasite is transmitted to the fence lizard, *Sceloporus occidentalis*, by the sandfly, *Lutzomyia vexator*. Sandfly species have the particularity of having very short lifespans; that is, only about 2% live long enough to take a second blood meal (Fiahlo and Schall, 1995). As a consequence, compared to other *Plasmodium* species, *P. mexicanum* has evolved toward rapid development that ensures the matching with the high vector's mortality. In addition, Fiahlo and Schall (1995) showed that infected sandflies were attracted by higher temperatures compared to uninfected counterparts. It turned out that this higher temperature is optimal for the rapid development of the parasite, but sub-optimal for the development of sandfly eggs, suggesting a parasitic manipulation of host temperature preference (Table 1).

It is well known that the development rate of vector-borne parasites is sensitive to temperature (e.g. Garnham, 1964). However studies like the one performed by Fiahlo and Schall (1995) are scarce. Extending such investigations to other insect vector–parasite systems, particularly those affecting human welfare, is promising and could shed light on new insights into parasitic manipulation of vector-borne diseases.

3.4. The recovery rate: 'evading the immune response'

Escaping the immune system of both the vertebrate and insect host is a *sine qua non* condition to the success of parasite transmission. As a consequence, vector-borne parasites evolved the ability to manipulate the immune response of their hosts (Tables 1 and 2).

In human for instance, malaria parasites, *Giardia* and African trypanosomes evolved an evasion mechanism known as antigenic variation. These parasites have a large multigene family of proteic variants and the evasion involves a sequential dominance of antigenic variants. Any one-parasite individual expresses only a single variant and only a few individuals exhibit a new variant per generation. As a consequence, parasites that exhibit the rare variant avoid the immune system and those that exhibit the frequent variant are killed. This enables the parasites to persist in the host. In addition, other strategies have been developed by protozoan parasites (i.e. malaria, African trypanosomes, *T. cruzi* and *Leishmania*) to avoid and/or suppress both adaptive and innate immune responses of their vertebrate host, such as eliminating their protein coat, induction of blocking antibodies, molecular mimicry, modulation of dendritic cells maturation and alterations of memory T-cell, macrophages, and cytokines functions (Zambrano-Villa et al., 2002; Sacks and Sher, 2002).

Table 2
Reported cases of behavioural and physiological alterations in the vertebrate host

Physiological and behavioural alterations	Host–parasite systems	References
Attraction (A), defensive behaviour (D), blood characteristics (B)	Malaria–Mosquito–Vertebrate	
	P. chabaudi, P. berghei, P. yoelii–A. aegypti–mouse (D) (B)	Day and Edman (1983), Day et al. (1983), Rossignol et al. (1985)
	P. chabaudi, *P. berghei, P. yoelii–Culex quinquefasciatus–mouse* (D)	Day and Edman (1983)
	P. chabaudi, *P. berghei*, *P. yoelii–An. stephensi–mouse* (D)	Day and Edman (1983), Taylor (2001), Ferguson et al. (2003), Ferguson and Read (2004).
	P. falciparum–An. gambiae–human (A)	Lacroix et al. (2005)
	Virus–Mosquito–Vertebrate	
	Rift Valley fever virus–A. aegypti–mouse (B)	Rossignol et al. (1985)
	Rift Valley fever virus–Culex pipiens–hens (A)	Mahon and Gibbs (1982)
	Rift Valley fever virus–Culex pipiens–lamb (A)	Turell et al. (1984)
	African trypanosomes–Glossina–Vertebrate	
	T. congolense–Glossina pallidipes–boran steer (A) (B)	Baylis and Nambiro (1993), Baylis and Mbwabi (1995), Moloo et al. (2000)
	Leishmania–Sandfly–Vertebrate	
	Leishmania infantum–Lutzomyia longipalpis–hamster (A)	Rebollar-Téllez (1999), O'Shea et al. (2002)
	Protozoa–Ticks–Rodents	
	Babesia microti–Ixodes trianguliceps–rodents (B)	Randolph (1991)
Immune response	Malaria–Vertebrate	
	P. falciparum–human	Holder et al. (1999), Riley et al. (1989), Urban et al. (1999)
	P. berghei–mouse	Xu et al. (2001)
	Trypanosomes–Vertebrate	
	T. cruzi–human	Norris (1998), Hall et al. (1992), Brodskyn et al. (2002)
	Leishmania–Vertebrate	
	Leishmania–human	Belkaid et al. (2001), Brittingham et al. (1995), Desjardins and Descoteaux (1997), Piedrafita et al. (1999).
	African Trypanosomes–human	
	T. brucei–human	Raper et al. (2001)

In the insect vector, it has been shown that although *Anopheles gambiae* was able to encapsulate and melanize foreign bodies, it failed to encapsulate *P. falciparum* (Schwartz and Koella, 2002). This failure probably results from the suppresssion of the insect immune encapsulation response both directly by the parasite and indirectly via the effects of the blood stage infection on the host vertebrate immune response (Boëte et al., 2004). It has been indeed demonstrated that in *P. gallinaceum*-infected *A. aegypti*, ookinete stages first actively suppress the insect immune melanization response and second the parasite appears to suppress the immune response indirectly by changing blood quality in the vertebrate host. It has been suggested that this indirect suppression results from complex interactions between the vertebrate and the mosquito immune responses (Boëte et al., 2004). However this hypothesis has been recently questioned. *Plasmodium falciparum* may rely in fact on a different strategy to avoid the immune response of *An. gambiae*. By comparing the ability to melanize a Sephadex bead of infected mosquitoes, of mosquitoes that had fed on infectious blood without becoming infected, and of control mosquitoes fed on uninfected blood, Lambrechts et al. (2007) found that infected mosquitoes had a stronger melanisation response than uninfected counterparts and mosquitoes in which infection failed. This result contrasts with the previous example and seems to indicate that *P. falciparum* relies on immune evasion rather than immuno-suppression to escape the immune response of *An. gambiae* (Lambrechts et al., 2007). A note of caution is however warranted, immune evasion of insect vectors by malaria parasites does not systematically occur. It has been indeed evidenced that *P. falciparum* can be melanized by *An. gambiae* after selection tratments in the laboratory (e.g. Collins et al., 1986) but also in nature (e.g. Riehle et al., 2006). However, the extend to which such resistance occurs in natural populations remains somehow obscure (Schwartz and Koella, 2002; Riehle et al., 2006).

Although parasite's strategies to overcome the vertebrate host's immune response have been widely explored and demonstrated (Zambrano-Villa et al., 2002; Sacks and Sher, 2002), little is known concerning such phenomena in the infected insect vector. It is also important to recognize that immune evasion could, in some cases, simply result from the absence of selective pressures on the vector to develop costly immune responses against malaria parasites. In other case, inability of insect to resist infection is unlikely to result from a lack of selective pressures, since evidence of fecundity and sometimes longevity reduction continue to accumulate (see Section 3.2). As a consequence, one might imagine that insect immune avoidance and suppression is also a common strategy used by parasite and more investigations are needed.

3.5. The vertebrate host attraction: 'be quiet and rapturous'

Once the parasite has reached maturity in the vertebrate host, any traits that could render its transmission from the vertebrate to the insect more probable should have a selective advantage. In this context, several malaria parasites, filarial worms and trypanosomes have evolved synchronicity between the peak number of infective stages and the peak biting time of their vector species (Lehane, 2005). Another way to optimize transmission would be for the parasite to respond positively to the presence of vectors. Recently, it has been shown that malaria parasites accelerate the growth of asexual stages and the development of infective gametocytes (i.e. gametocytogenesis) when exposed to a high number of mosquitoes (Billingsley et al., 2005). A last way to optimize transmission from vertebrate to insect would be to induce a bias for vectors toward hosts with mature infections. This can be achieved by alterations of vertebrate host (i) attraction, (ii) behavioural defences and (iii) blood characteristics (Moore, 2002).

Increased attractiveness of infected vertebrate host has been reported in many (e.g. Mahon and Gibbs, 1982; Baylis and Mbwabi, 1995; O'Shea et al., 2002; Ferguson et al., 2003; Lacroix et al., 2005; Table 2) but not all studies (Freier and Friedman, 1976; Burkot et al., 1989; Kruppa and Burchard, 1999). It has been suggested that fever and its symptoms (i.e. increase in body temperature, in lactate production and in hyperventilation) that accompany infections (e.g. arboviruses, malaria) might increase attractiveness of infected individuals (Turell et al., 1984; Nacher, 2005). However, in many cases, no consistent differences in body temperature are found between infected and uninfected individuals (e.g. Mahon and Gibbs, 1982; Day and Edman, 1984). Recently, through an elegant semi-natural study, Lacroix et al. (2005) demonstrated that people infected with transmissible stages of malaria appear to produce something attractive to mosquitoes. First, they found that mosquitoes are more attracted to humans harbouring the transmissible gametocyte stages of the parasite than to uninfected individuals and individuals harbouring non-transmissible asexual stages of the parasite. As the infection was asymptomatic, a raise in body temperature could not explain this differential attractiveness. Second, they showed that after the clearance of the parasite, previously infected individuals were no more attractive than other individuals. Therefore, the observed increased attractiveness was not the consequence of intrinsic attractiveness (and thus making these individuals infected frequently) but rather the consequence of infection with transmissible stages of the parasite.

At shorter distances, reducing host defensive behaviour is a good way to facilitate the vector blood meal and hence the transmission. In this context, Day et al. (1983) and Day and Edman (1983) found that infection with *Plasmodium chabaudi*, *P. yoelli* and *P. berghei* make mice less defensive and consequently preferentially bitten and fed upon by mosquito vectors. In addition, they showed that reduced defensive behaviours coincided with peaks in gametocyte numbers and consequently would result in increased transmission (see also Turell et al., 1984). In this case, the general lethargy associated with infection may be the cause of reduced vector-repellent behaviour (Rossignol et al., 1985).

Hosts infected with malaria but also with other vector-borne diseases (e.g. dengue, African trypanosomiasis, babesiosis, etc.) often exhibit altered blood characteristics, including anaemia (i.e. loss of erythrocytes), and thrombocytopenia (i.e. loss of platelets). These modifications seem to reduce blood viscosity that may facilitate blood ingestion, and also decrease the vector's mortality during the blood meal (Ribeiro et al., 1985; Taylor and Hurd, 2001). For instance, Rossignol et al. (1985) found that the time taken by mosquitoes to feed on mice and hamsters experimentally infected with *P. chabaudi* or Rift Valley fever was reduced by at least one minute. In the field, tsetse flies feed more successfully on cattle infected with *T. congolense* than on uninfected cattle (Baylis and Nambiro, 1993; Baylis and Mbwabi, 1995; Moloo et al., 2000). The increased feeding success of tsetse flies may not be related to changes in the level of anaemia, but rather because of vasodilatation in infected cattles (Moloo et al., 2000).

3.6. The vertebrate host selection: 'make the good choice'

Vector host choice is a very important key predictor for the transmission intensity of vector-borne diseases. This choice may be influenced by genetic and environmental factors such as the innate host preference of the bloodsucking insect and the host availability. Beside this aspect, vector-borne parasites show some degree of vertebrate host specificity and one might imagine that parasites acquired, during the course of evolution, the ability to target appropriate host and/or avoid unsuitable ones. In the context of parasitic manipulation, studying specificity is one way to estimate the costs associated with the changes and thus is one way to address the adaptiveness of this strategy. The costs for a manipulative parasite can include energetic expenses that are necessary to induce the changes and also mortality associated with these changes. For example in the case of trophically transmitted parasites, increased susceptibility to predation by unsuitable hosts should be very costly (e.g. Mouritsen and Poulin, 2003). In vector-borne diseases, the ''qualitative manipulation'' hypothesis has been recently proposed, according to which generalist blood-feeding insects, once infected, would develop a feeding preference for hosts targeted by the parasite (Lefèvre et al., 2006). Alteration of host choice could theoretically occur at inter- and/or intra-specific level with infected vectors preferentialy feeding on host species or on host individuals that are suitable for the parasite.

First, maximising transmission towards a suitable host could be achieved by parasites by inducing in the vector a sensory bias for host traits that are correlated with optimal suitability for the parasite. Second, the parasite may induce changes in the generalist vector such as an alteration of microhabitat choice, in order to spatially match the microhabitat of the suitable host. Finally, the parasite may induce changes in the vector such as an alteration of time activity in order to temporally match the time rest or activity of the suitable host. These scenarios can be

easily tested, by for example, sampling at distinct period and comparing the proportion of infected vectors between samples. However, if we observe a difference in habits such as microhabitat choice, temporal activity and host choice between infected and uninfected vectors, one might argue that infection can be more a consequence than a cause of these habits. In other words, an infected vector may prefer a host species that is suitable for parasite not because of being infected but just because of an innate or acquired preference for this host, and thus making these vectors infected. To test between these two alternatives, comparison of habits (e.g. host choice) between uninfected vectors, vectors with mature infection and vectors with immature infection could be performed in the field. For instance, if vectors with mature infection do not exhibit the same habits than uninfected vectors and those with immature infection, then the observed habits are more likely to result from a manipulative process rather than from innate or acquired processes. In addition, comparisons between uninfected vectors and experimentally infected ones could be also performed at laboratory.

Qualitative manipulation deserves consideration and its investigation would permit to adress the complexity and multidimensionality in host manipulation by parasites. Indeed, many studies concentrated on only one character whereas manipulative parasites are known to modify more than a single dimension in their host phenotype (Thomas et al., 2005; Cezilly and Perrot-Minnot, 2005). Studies exploring the qualitative manipulation hypothesis may yield considerable information about the diversity of transmission strategies used by vector-borne parasites and find, for instance, that insect vectors infected with transmissible stages of the parasite bite more but also bite better.

4. Conclusion and future directions

Taken together the examples presented above suggest that many vector-borne parasites manipulate their hosts to increase their probability of transmission. Unfortunately, we have little evidence that such changes indeed affect transmission, although it is reasonable to expect that it would (Moore, 2002). For example, evidence have been accumulating showing that vector-borne parasites alter their vertebrate hosts to improve the attraction and feeding success of their vectors (see Section 3.5) and important epidemiological consequences of such parasitic influences have been demonstrated by models of vector-borne disease transmission (Dye and Hasibeder, 1986; Kingsolver, 1987; Rossignol and Rossignol, 1988; Burkot, 1988; Randolph and Nuttall, 1994).

Whereas some authors consider these changes as examples of parasitic manipulation (e.g. Lehane, 2005), others interpreted these as non-adaptive side effects of infection or modifications that are coincidentally beneficial for the parasite (e.g. Clements, 1999). As discussed in the first section of this review, such notion emanates from Poulin (1995) and can sometimes bring confusions that may lead to sentence such as: ‘‘All these potential benefits are a direct result of infection pathology rather than any direct manipulation of the host. Although it is feasible that increasingly virulent pathology is selected for increased transmission to the insect vector, it is more likely that these are simply side-effect, rather than selected trait’’ (Paul et al., 2003). We fully agree that many of these changes may result from indirect processes (see Thomas et al., 2005 for discussion on indirect/direct mechanisms in manipulative process), however if the trait that affects parasite transmission has a genetic basis, and presents some variation, one could expect that natural selection has not been blind, and that consequently this trait has been selected over the course of evolution. Could a side-effect not be a selected trait?

It has been also suggested that such changes might be considered as examples of mutualism (Rossignol et al., 1985; Clements, 1999): the vector benefits from the pathological conditions induced by the parasites in the vertebrate host (e.g. reduction of blood viscosity facilitating blood ingestion) while the parasite gains from the choice of the vector. However, one might argue that it is not clear whether this short-term benefit for the vector would outweigh the long-term costs of becoming parasitized. Indeed, it appears that parasites harm their vectors in many ways (Lehane, 2005).

In the future, it will be useful to perfect the empirical and theoretical study of parasitic manipulation in vector-borne diseases to move from documenting the existence of parasitic manipulation to asking how and why parasitic manipulation is expressed in some instances and not others. For this we will need a sharper quantification of the benefits and costs of manipulating and being manipulated. Beyond potential energetic expenses to induce phenotypic changes, costs for a manipulative parasite can also include mortality associated to these changes. For example, the balance between the transmission benefit of an increased biting rate and the costs of mortality associated with this behavioural change may partly explain why this change is expressed in many but not all cases. From the insect vector’s perspective, increased biting rate is also very risky. It has been indeed shown that sporozoite-infected mosquitoes have higher feeding associated mortality than those without sporozoites (Anderson et al., 2000). However theoretical analysis showed that the optimal biting rate is higher for infective stages of malaria parasites than for their mosquito vectors (Koella, 1999; Schwartz and Koella, 2001). Usually, selection to reduce the effects of manipulation is expected, when the extent with which the parasite reduces its own fitness by manipulating its host reaches a certain threshold. Additional research is requiered to fill gaps in our knowledge of the balance between costs and benefits of parasitic manipulation.

As pointed out by Paul et al. (2003), the strategies used by vector-borne parasites to optimize interaction with one host may also be subject to selection for optimization in the other. For instance, the ‘‘transmission enhancer’’ fPPG produced by leishmania parasite (see Section 3.1) appears to participate in both increased feeding rate of infected sandflies and the success of infection within the host (Rogers et al., 2004). Recasting the parasite in its two host-life cycle may clearly offer new research avenues in manipulative process. For instance in malaria, conflict of interest between immature stages (i.e. oocyst) and

mature ones (i.e. sporozoite) in the mosquito have been suggested concerning the optimal biting rate of the vector (Koella, 1999; Koella et al., 2002). Such conflicts may also occur between different stages in the mosquito and in the vertebrate host: when a mosquito infected with malaria feeds on an infected vertebrate host, sporozoite stages within the mosquito would prefer a higher blood viscosity than what is optimal for gametocyte stages within the host. Indeed, infection in the host may cause anemia to facilitate the insect bloodmeal, whereas infection in the insect may cause apyrase reduction to impair the insect bloodmeal (see Sections 3.1 and 3.5).

Recently, emphasis has been placed on the role of co-occuring parasite species within manipulated hosts. What should we expect if the same insect-vector harbours different parasite species that may or may not have the same vertebrate host? This within-host interaction creates opportunities for either synergistic or conflicting interests between different parasite species and has thus the potential to affect the result of manipulation. In trophically transmitted parasites, several scenarii have been proposed to illustrate this phenomenon and empirical evidence begins to accumulate (Lafferty, 1999; Thomas et al., 2005). For example, when different parasite species share both vectors and vertebrate hosts, all parasites may benefit from a higher transmission success if one of them is a manipulative parasite. Non-manipulative parasites can be simply 'lucky passenger' (in the vector and/or in the host) when randomly associated with a manipulator. In trophic transmission it has been shown that selection may even favour non-manipulative parasites able to preferentially infect intermediate hosts already manipulated, a strategy named 'hitchhicking'. Mixed trypanosome infections in tsetse flies seem to be common and hitchhicking has been hypothesized to explain the non-random associations between the savannah and riverine forest taxonomic groups of *T. congolense* (Solano et al., 2001). The riverine-forest type of *T. congolense* appears to benefit from the presence of the savannah type within the glossina both for its development and transmission. Parasite species may also have conflicting interests in the use of the insect vector shared with manipulative species when it requires different definitive hosts. Such a situation may theoretically occur in *Glossina* (e.g. *palpalis* or *morsitans* group) that are important vectors for both human and non-human trypanosomes. For example, conflicts may be solved by the sabotage of the manipulation. However, the prevalence of the parasites should be high in order to exert a selection pressure strong enough for such a strategy to evolve. Co-occuring parasites within manipulated host is a fruitful area for further research and we therefore encourage to examine the entire community of parasites in manipulated insect vectors and vertebrate hosts.

Another interesting avenue is the investigation of the variation in the intensity of alterations. Several studies on host manipulation by parasites indeed indicated that the level of manipulative abilities in parasites and/or the level of resistance in hosts show some degree of variability (Thomas et al., 2005). Another similar situation occurs in a much more investigated topic, that is, the variation in virulence and resistance in malaria parasite—rodent systems (e.g. Grech et al., 2006). Exploring variability in manipulative processes in vector-borne diseases and integrating the approach developed to understand the evolution of virulence would certainly provide novel information to understand the evolution of parasitic manipulation and may offer new insights into control measures. Intra-specific (host and/or parasite) variation may be responsible for the variation in manipulative processes; in an ecological context, another cause that warrants investigations is the variation due to seasonality. As for some trophically transmitted parasites (Helluy and Holmes, 2005), seasonal patterns can also occur in manipulative vector-borne parasites. For instance, the optimal manipulative effort (sensu Poulin, 1994) to make the vertebrate host attractive (see Section 3.5) may decrease when the vector abundance increases.

Future research emphasizing the molecular and physiological mechanisms of altered behaviour in vector-borne diseases will probably provide a considerable gain to knowledge. Fundamentally, exploring the proximate mechanisms underlying behavioural manipulation could partly address the existence and magnitude of the costs of manipulation. Do manipulative parasites produce themself the molecules (e.g. neuromodulators) leading to the altered behaviours or do they use instead the products of the immunity responses they induce in the hosts? (Adamo, 2002; Thomas et al., 2005). Such investigations require significant input from field biologists, evolutionary ecologists, neurobiologists, biochemists and epidemiologists. Combining the different approaches, techniques, and backgrounds of such disciplines is likely to produce information of wide interest.

Vector-borne parasites induce many changes in their insect and vertebrate hosts in a way that renders their transmission more probable. Studying these strategies and the interplay with the hosts has the potential to bring important amount of fundamental knowledge as well as define potential ''Achille-s'heels'' for control measures.

Acknowledgements

We thank Simone Helluy, Phil Agnew, Manue Rival and two anonymous referees for constructive comments on earlier drafts of this paper.

References

Adamo, S., 2002. Modulating the modulators: parasites, neuromodulators and host behavioral change. Brain Behav. Evol. 60, 370–377.

Anderson, R.A., Koella, J.C., Hurd, H., 1999. The effect of *Plasmodium yoelii nigeriensis* infection on the feeding persistence of *Anopheles stephensi* Liston throughout the sporogonic cycle. Proc. R. Soc. Lond. B 266, 1729–1733.

Anderson, R.A., Knols, B.G., Koella, J.C., 2000. *Plasmodium falciparum* sporozoites increase feeding-associated mortality of their mosquito hosts *Anopheles gambiae* s.l. Parasitology 120, 329–333.

Anez, East, 1984. Studies on *Trypanosoma rangeli* Tejera, 1920. II. Its effect on feeding behaviour of triatomine bugs. Acta Trop. 41, 93–95.

Azambuja, P., Garcia, E.S., 2005. *Trypanosoma rangeli* interactions within the vector *Rhodnius prolixus*: a mini review. Mem. Inst. Oswaldo Cruz 100, 567–572.

Bacot, A.W., Martin, C.J., 1914. Observations of the mechanism of the transmission of plague by fleas. J. Hyg., Plague Supplement III 423–429.

Baker, J.R., Robertson, D.H.H., 1957. An experiment on the infectivity to *Glossina morsitans* of a strain of *Trypanosoma rhodesiense* and of a strain of *T. brucei* with some observations on the longevity of infected flies. Ann. Trop. Med. Parasitol. 51, 121–135.

Barnard, C.J., 1990. Parasitic relationship. In: Barnard, C.J., Behnke, J.M. (Eds.), Parasitism and Host Behaviour. Taylor and Francis, London, pp. 1–33.

Baylis, M., Nambiro, C.O., 1993. The effect of cattle infection by *Trypanosoma congolense* on the attraction, and feeding success, of the tsetse fly *Glossina pallidipes*. Parasitology 106, 357–361.

Baylis, M., Mbwabi, A.L., 1995. Feeding behaviour of tsetse flies (*Glossina pallidipes Austen*) on Trypanosoma-infected oxen in Kenya. Parasitology 110, 297–305.

Beach, R., Leeuwenberg, J., 1985. Modification of sandfly biting behavior by *Leishmania* leads to increased parasite transmission. Am. J. Trop. Med. Hyg. 34, 278–282.

Beier, J.C., Davis, J.R., Vaughan, J.A., Noden, B.H., Beier, M.S., 1991. Quantitation of *Plasmodium falciparum* sporozoites transmitted in vitro by experimentally infected *Anopheles gambiae* and *Anopheles stephensi*. Am. J. Trop. Med. Hyg. 44, 564–570.

Belkaid, Y., Hoffmann, K.F., Mendez, S., Kamhawi, S., Udey, M.C., Wynn, T.A., Sacks, D.L., 2001. The role of interleukin (IL)-10 in the persistence of *Leishmania major* in the skin after healing and the therapeutic potential of anti-IL-10 receptor antibody for sterile cure. J Exp. Med. 19, 1497–1506.

Bethel, W.M., Holmes, J.C., 1973. Altered evasive behavior and responses to light in amphipods harboring acanthocpehalan cystacanths. J. Parasitol. 59, 945–956.

Bibikova, V.A., 1977. Contemporary views on the interrelationships between fleas and the pathogens of human and animal diseases. Annu. Rev. Entomol. 22, 23–32.

Billingsley, P.F., Snook, L.S., Johnston, V.J., 2005. Malaria parasite growth is stimulated by mosquito probing. Biol. Lett. 1, 185–189.

Biron, D.G., Marché, L., Ponton, F., Loxdale, H., Galéotti, N., Renault, L., Joly, C., Thomas, F., 2005. Behavioural manipulation in a grasshopper harbouring hairworm: a proteomics approach. Proc. R. Soc. Lond. B 272, 2117–2126.

Boëte, C., Paul, R.E.L., Koella, J.C., 2004. Direct and indirect immunosuppression by a malaria parasite in its mosquito vector. Proc. R. Soc. Lond. B 271, 1611–1615.

Botto-Mahan, C., Cattan, P.E., Medel, R., 2006. Chagas disease parasite induces behavioural changes in the kissing bug *Mepraia spinolai*. Acta Trop. 98, 219–223.

Brittingham, A., Morrison, C.J., McMaster, W.R., McGwire, B.S., Chang, K.P., Mosser, D.M., 1995. Role of the Leishmania surface protease gp63 in complement fixation, cell adhesion, and resistance to complement-mediated lysis. J. Immunol. 155, 3102–3111.

Brodskyn, C., Patricio, J., Oliveira, R., Lobo, L., Arnholdt, A., Mendonca-Previato, L., Barral, A., Barral-Netto, M., 2002. Glycoinositolphospholipids from *Trypanosoma cruzi* interfere with macrophages and dendritic cell responses. Infect. Immun. 70, 3736–3737.

Burkot, T.R., 1988. Non-random host selection by anopheline mosquitoes. Parasitol. Today 4, 156–162.

Burkot, T., Narara, A., Paru, R., Graves, P., Garner, P., 1989. Human host selection by anophelines: no evidence for preferential selection of malaria or microfilariae-infected individuals in a hyperendemic area. Parasitology 98, 337–342.

Cézilly, F., Grégoire, A., Bertin, A., 2000. Conflict between co-occurring manipulative parasites? An experimental study of the joint influence of two acanthocephalan parasites on the behaviour of *Gammarus pulex*. Parasitology 120, 625–630.

Cezilly, F., Perrot-Minnot, M.-J., 2005. Studying adaptive changes in the behaviour of infected hosts: a long and winding road. Behav. Proc. 68, 223–228.

Chege, G.M.M., Beier, J.C., 1990. Effect of *Plasmodium falciparum* on the survival of naturally infected Afrotropical *Anopheles* (Diptera: Culicidae). J. Med. Entomol. 27, 454–458.

Clements, A.N., 1999. The Biology of Mosquitoes. Sensory Reception and Behaviour. Cambridge University Press, Cambridge.

Collins, F.H., Sakai, R.K., Vernick, K.D., Paskewitz, S., Seeley, D.C., Miller, L.H., Collins, W.E., Campbell, C.C., Gwadz, R.W., 1986. Genetic selection of a *Plasmodium*-refractory strain of the malaria vector *Anopheles gambiae*. Science 234, 607–610.

Combes, C., 2001. Parasitism. The Ecology and Evolution of Intimate Interactions. Chicago University Press.

D'Alessandro, A., Mandel, S., 1969. Natural Infections and Behavior of *Trypanosoma rangeli* and *Trypanosoma cruzi* in the Vector *Rhodnius prolixus* in Colombia. J. Parasitol. 55, 846–852.

Dawkins, R., Krebs, J.R., 1979. Arms races between and within species. Proc. R. Soc. Lond. B 205, 489–511.

Dawkins, R., 1982. The Extended Phenotype. Oxford University Press, Oxford.

Day, J.F., Edman, J.D., 1983. Malaria renders mice susceptible to mosquito feeding when gametocytes are most infective. J. Parasitol. 69, 163–170.

Day, J.F., Ebert, K.M., Edman, J.D., 1983. Feeding patterns of mosquitoes (Diptera: Culicidae) simultaneously exposed to malarious and healthy mice, including a method for separating blood meals from conspecific hosts. J. Med. Entomol. 20, 120–127.

Day, J.F., Edman, J.D., 1984. Mosquito engorgement on normally defensive hosts depends on host activity patterns. J. Med. Entomol. 21, 732–740.

Desjardins, M., Descoteaux, A., 1997. Inhibition of phagolysosomal biogenesis by the *Leishmania* lipophosphoglycan. J. Exp. Med. 185, 2061–2068.

Duke, A.L., 1928. On the effect on the longevity of *G. palpalis* of trypanosome infections. Ann. Trop. Med. Parasitol. 22, 25–32.

Dye, C., Hasibeder, G., 1986. Population dynamics of mosquito-borne disease: effects of flies which bite some people more frequently than others. Trans. R. Soc. Trop. Med. Hyg. 80, 69–77.

Dye, C., 1992. The analysis of parasite transmission by blood-sucking insect. Annu. Rev. Entomol. 37, 1–19.

El Sawaf, B.M., El Sattar, S.A., Shehata, M.G., Lane, R.P., Morsy, T.A., 1994. Reduced longevity and fecundity in *Leishmania*-infected sand flies. Am. J. Trop. Med. Hyg. 51, 767–770.

Faran, M.E., Turell, M.J., Romoser, W.S., Routier, R.G., Gibbs, P.H., Cannon, T.L., Bailey, C.L., 1987. Reduced survival of adult *Culex pipiens* infected with Rift Valley fever virus. Am. J. Trop. Med. Hyg. 37, 403–409.

Ferguson, H.M., Read, A.F., 2002. Why is the effect of malaria parasites on mosquito survival still unresolved? Trends Parasitol. 18, 256–261.

Ferguson, H.M., Rivero, A., Read, A.F., 2003. The influence of malaria parasite genetic diversity and anaemia on mosquito feeding and fecundity. Parasitology 127, 9–19.

Ferguson, H.M., Read, A.F., 2004. Mosquito appetite for blood is stimulated by *Plasmodium chabaudi* infection in themselves and their vertebrate hosts. Malaria J. 3, 12.

Fiahlo, R.F., Schall, J.J., 1995. Thermal ecology of a malarial parasite and its insect vector: consequences for the parasite's transmission success. J. Anim. Ecol. 64, 553–562.

Freier, J.E., Friedman, S., 1976. Effect on host infection with *Plasmodium gallinaceum* on the reproductive capacity of *Aedes aegypti*. J. Invertebr. Pathol. 28, 161–166.

Freier, J.E., Friedman, S., 1987. Effect of *Plasmodium gallinaceum* infection on the mortality and body weight of *Aedes aegypti* (Diptera: Culicidae). J. Med. Entomol. 24, 6–10.

Gamage-Mendis, A.C., Rajakaruna, J., Weerasinghe, S., Mendis, C., Carter, R., Mendis, K.N., 1993. Infectivity of *Plasmodium vivax* and *P. Falciparum* to *Anopheles tesselatus*: relationship between oocyst and sporozoite development. Trans. R. Soc. Trop. Med. Hyg. 87, 3–6.

Garcia, E.S., Mello, C.B., Azambuja, P., Ribeiro, J.M.C., 1994. *Rhodnius prolixus*: salivary anti-hemostatic components decrease with *Trypanosoma rangeli* infection. Exp. Parasitol. 78, 287–293.

Garnham, P.C.C., 1964. In: Taylor, A.E.R. (Ed.), Factors Influencing the Development of Protozoa in Their Arthropodan Hosts. Host–Parasite Relationships in Invertebrate Hosts. Blackwell, Oxford, pp. 33–50.

Garrett-Jones, C., Shidrawi, G.R., 1969. Malaria vectorial capacity of a population of *Anopheles gambiae*: an exercise in epidemiological entomology. Bull. WHO 40, 531–545.

Gould, S.J., Vrba, E.S., 1982. Exaptation—a missing term in the science of form. Paleobiology 8, 4–15.

Gray, J.S., 1982. The effects of the piroplasm *Babesia bigemina* on the survival and reproduction of the blue tick, *Boophilus decoloratus*. J. Invertebr. Pathol. 39, 413–415.

Grech, K., Watt, K., Read, A.F., 2006. Host–parasite interactions for virulence and resistance in a malaria model system. J. Evol. Biol. 19, 1620–1630.

Grimstad, P.R., Ross, Q.E., Graiq, J.B., 1980. *Aedes triseriatus* (Diptera: Culicidae) and La Crosse virus. II. Modification of mosquito feeding behavior by virus infection. J. Med. Entomol. 17, 1–7.

Hall, B.F., Webster, P., Ma, A.K., Joiner, K.A., Andrews, N.W., 1992. Desialylation of lysosomal membrane glycoproteins by *Trypanosoma cruzi*: a role for the surface neuraminidase in facilitating parasite entry into the host cell cytoplasm. J. Exp. Med. 1, 313–325.

Hamilton, J.G.C., Hurd, H., 2002. In: Lewis, E.E., Campbell, J.F., Sukhdeo, M.V.K. (Eds.), The Behavioural Ecology of Parasites. CABI Publishing, pp. 259–281.

Hart, B.L., 1988. Biological basis of the behavior of sick animals. Neurosci. Behav. Rev. 12, 123–137.

Helluy, S., 1984. Relations hôtes-parasites du trématode *Microphallus papillorobustus* (Rankin, 1940). III. Facteurs impliqués dans les modifications du comportement des Gammarus hôtes intermediaires et tests de prédation. Ann. Parasitol. Hum. Comp. 59, 41–56.

Helluy, S., Holmes, J.C., 2005. Parasitic manipulation: further considerations. Behav. Proc. 68, 205–210.

Hinnebusch, B.J., Fischer, E.R., Schwan, T.G., 1998. Evaluation of the role of the *Yersinia pestis* plasminogen activator and other plasmid-encoded factors in temperature-dependent blockage of the flea. J. Infect. Dis. 178, 1406–1415.

Hogg, J.C., Hurd, H., 1997. The effects of natural *Plasmodium falciparum* infection on the fecundity and mortality of *Anopheles gambiae* s.l. in north east Tanzania. Parasitology 114, 325–331.

Holder, A.A., Guevara Patino, J.A., Uthaipibull, C., Syed, S.E., Ling, I.T., Scott-Finnigan, T., Blackman, M.J., 1999. Merozoite surface protein 1, immune evasion, and vaccines against asexual blood stage malaria. Parassitologia 41, 409–414.

Holmes, J.C., Bethel, W.M., 1972. Modification of intermediate host behaviour by parasites. In: Canning, E.U., Wright, C.A. (Eds.), Behavioural Aspects of Parasite Transmission. Academic Press, London, pp. 123–149.

Hurd, H., Hogg, J.C., Renshaw, M., 1995. Interaction between blood-feeding, fecundity and infection in mosquitoes. Parasitol. Today 11, 411–416.

Hurd, H., 2001. Host fecundity reduction: a strategy for damage limitation? Trends Parasitol. 17, 363–368.

Hurd, H., 2003. Manipulation of medically important insect vectors by their parasites. Annu. Rev. Entomol. 48, 141–161.

Jefferies, D., 1984. Transmission of Disease by Haematophagous Arthropods: Host–Parasite Interactions and Fluid Mechanics. PhD Thesis. University of Salford.

Jenni, L., Molyneux, D.H., Livesey, J.L., Galun, R., 1980. Feeding behavior of tsetse flies infected with salivarian trypanosomes. Nature 283, 383–385.

Killick-Kendrick, R., Leaney, A.J., Ready, P.D., Molyneux, D.H., 1977. Leishmania in phlebotomid sandflies. IV. The transmission of *Leishmania mexicana amazonensis* to hamsters by the bite of experimentally infected *Lutzomyia longipalpis*. Proc. R. Soc. B 196, 105–115.

Kingsolver, J.G., 1987. Mosquito host choice and the epidemiology of malaria. Am. Nat. 130, 811–827.

Koella, J.C., Packer, M.J., 1996. Malaria parasites enhance blood-feeding of their naturally infected vector *Anopheles punctulatus*. Parasitology 113, 105–109.

Koella, J.C., Sørensen, F.L., Anderson, R., 1998. The malaria parasite *Plasmodium falciparum* increases the frequency of multiple feeding of its mosquito vector *Anopheles gambiae*. Proc. R. Soc. Lond. B 265, 763–768.

Koella, J.C., 1999. An evolutionary view of the interactions between anopheline mosquitoes and malaria parasites. Microbes Infect. 1, 303–308.

Koella, J.C., Rieu, L., Paul, R.E.L., 2002. Stage-specific manipulation of a mosquito's host-seeking behavior by the malaria parasite *Plasmodium gallinaceum*. Behav. Ecol. 13, 816–820.

Krishnamoorthy, K., Subramanian, S., Van Oortmarssen, G.J., Habbema, J.D.F., Das, P.K., 2004. Vector survival and parasite infection: the effect of *Wuchereria bancrofti* on its vector *Culex quinquefasciatus*. Parasitology 129, 43–50.

Kruppa, T.F., Burchard, G.D., 1999. Similar blackfly attraction by onchocerciasis patients and individuals putatively immune to *Onchocerca volvulus*. Trans. R. Soc. Trop. Med. Hyg. 93, 365–367.

Lacroix, R., Mukabana, W.R., Gouagna, L.C., Koella, J.C., 2005. Malaria infection increases attractiveness of human to mosquitoes. PLoS Biol. 3.

Lafferty, K.D., 1999. The evolution of trophic transmission. Parasitol. Today 15, 111–115.

Lambrechts, L., Morlais, I., Awono-Ambene, P.H., Cohuet, A., Simard, F., Jacques, J.-C., Bourgouin, C., Koella, J.C., 2007. Effect of infection by *Plasmodium falciparum* on the melanisation immune response of *Anopheles gambiae*. Am. J. Trop. Med. Hyg. 76, 475–480.

Lefèvre, T., Koella, J.C., Renaud, F., Hurd, H., Biron, D.G., Thomas, F., 2006. New prospectes for research on manipulation of insect vectors by pathogens. PLoS Pathogens 2, 0633–0635.

Lehane, M., 2005. The Biology of Bloodsucking in Insects, 2nd ed. Cambridge University Press, Cambridge.

Li, X., Sina, B., Rossignol, P.A., 1992. Probing behavior and sporozoite delivery by *Anopheles stephensi* infected with *Plasmodium berghei*. Med. Vet. Entomol. 6, 57–61.

Livesey, J.L., Molyneux, D.H., Jenni, L., 1980. Mechanoreceptor–trypanosome interactions in the labrum of *Glossina*: fluid mechanics. Acta Trop. 37, 151–161.

Macdonald, G., 1956. Epidemiological basis of malaria control. Bull. WHO 15, 613–626.

Macdonald, G., 1957. The Epidemiology and Control of Malaria. Oxford University Press, 201 pp.

Mahon, R., Gibbs, A., 1982. Arbovirus-infected hens attract more mosquitoes. In: Mackenzie, J.D. (Ed.), Viral Diseases in Southeast Asia and the Western Pacific. Academic Press, Sydney, pp. 502–504.

Makumi, J.N., Moloo, S.K., 1991. *Trypanosoma vivax* in *Glossina palpalis gambiensis* do not appear to affect feeding behaviour, longevity or reproductive performance of the vector. Med. Vet. Entomol. 5, 35–42.

Maudlin, I., Welburn, S.C., Milligan, P.J., 1998. Trypanosome infections and survival in tsetse. Parasitology 116, S23–S28.

McGaw, M.M., Chandler, L.J., Wasieloski, L.P., Blair, C.D., Beaty, B.J., 1998. Effect of La Crosse virus infection on overwintering *Aedes triseriatus*. Am. J. Trop. Med. Hyg. 58, 168–175.

Moloo, S.K., 1983. Feeding behaviour of *Glossina morsitans morsitans* infected with *Trypanosoma vivax, T. congolense* or *T. brucei*. Parasitology 86, 51–56.

Moloo, S.K., Dar, F., 1985. Probing by *Glossina morsitans centralis* infected with pathogenic *Trypanosoma* species. Trans. R. Soc. Trop. Med. Hyg. 79, 119.

Moloo, S.K., Kutuza, S.B., 1985. Survival and reproductive performance of female *Glossina morsitans morsitans* when maintained on livestock infected with salivarian trypanosomes. Ann. Trop. Med. Parasitol. 79, 223–224.

Moloo, S.K., Sabwa, C.L., Baylis, M., 2000. Feeding behaviour of *Glossina pallidipes* and *G. Morsitans centralis* on Boran cattle infected with *Trypanosoma congolense* or *T. Vivax* under laboratory conditions. Med. Vet. Entomol. 14, 290–299.

Molyneux, D.H., Jenni, L., 1981. Mechanoreceptors, feeding behaviour and trypanosomes transmission in Glossina. Trans. R. Soc. Trop. Med. Hyg. 75, 160–163.

Molyneux, D.H., Jefferies, D., 1986. Feeding behaviour of pathogen-infected vectors. Parasitology 92, 721–736.

Moore, J., 1983. Responses of an avian predator and its isopod prey to an acanthocephalan parasite. Ecology 64, 1000–1015.

Moore, J., 1993. Parasites and the behaviour of biting flies. J. parasitol. 79, 1–16.

Moore, J., 2002. Parasites and the Behavior of Animals. Oxford Series in Ecology and Evolution. Oxford University Press, Oxford, UK.

Mouritsen, K.N., Poulin, R., 2003. Parasite-induced trophic facilitation exploited by a non-host predator: a manipulator's nightmare. Int. J. Parasitol. 33, 1043–1050.

Nacher, M., 2005. Charming the mosquito: do malaria symptoms increase the attractiveness of the host for the vector? Med. Hypotheses 64, 788–791.

Niebylski, M.L., Peacock, M.G., Schwan, T.G., 1999. Lethal effect of *Rickettsia rickettsii* on its tick vector (*Dermacentor andersoni*). Appl. Environ. Microbiol. 65, 773–778.

Norris, K.A., 1998. Stable transfection of *Trypanosoma cruzi* epimastigotes with the trypomastigote-specific complement regulatory protein cDNA confers complement resistance. Infect. Immun. 66, 2460–2465.

O'Shea, B., Rebollar-Tellez, E., Ward, R.D., Hamilton, J.G.C., El Naiem, D., Polwart, A., 2002. Enhanced sandfly attraction to Leishmania-infected hosts. Trans. R. Soc. Trop. Med. Hyg. 96, 117–118.

Paul, R.E.L., Ariey, F., Robert, V., 2003. The evolutionary ecology of *Plasmodium*. Ecol. Lett. 6, 866–880.

Platt, K.B., Linthicum, K.J., Myint, K.S., Innis, B.L., Lerdthusnee, K., Vaughn, D.W., 1997. Impact of dengue virus infection on feeding behavior of *Aedes aegypti*. Am. J. Trop. Med. Hyg. 57, 119–125.

Piedrafita, D., Proudfoot, L., Nikolaev, A.V., Xu, D., Sands, W., Feng, G.J., Thomas, E., Brewer, J., Ferguson, M.A., Alexander, J., Liew, F.Y., 1999. Regulation of macrophage IL-12 synthesis by *Leishmania* phosphoglycans. Eur. J. Immunol. 29, 235–244.

Ponnudurai, T., Lensen, A.H.W., Van Gemert, G.F.A., Bolmer, J.H.E., Meuwissen, J.H.E.T., 1991. Feeding behaviour and sporozoite ejection by infected *Anopheles stephensi*. Trans. R. Soc. Trop. Med. Hyg. 85, 175–180.

Poulin, R., 1994. The evolution of parasite manipulation of host behaviour: a theoretical analysis. Parasitology 109, S109–S118.

Poulin, R., Brodeur, J., Moore, J., 1994. Parasite manipulation of host behaviour: should host always lose? Oikos 70, 479–484.

Poulin, R., 1995. "Adaptive" change in the behaviour of parasitized animals: a critical review. Int. J. Parasitol. 25, 1371–1383.

Putnam, J.P., Scott, T.W., 1995. Blood feeding behavior of dengue-2 virus infected Aedes aegypti. Am. J. Trop. Med. Hyg. 52, 225–227.

Randolph, S.E., 1991. The effect of *Babesia microti* on feeding and survival in its tick vector, Ixodes trianguliceps. Parasitology 102, 9–16.

Randolph, S.E., Nuttall, P.A., 1994. Nearly right or precisely wrong? Natural versus laboratory studies of vector-borne diseases. Parasitol. Today 10, 458–462.

Raper, J., Portela, M.P., Lugli, E., Frevert, U., Tomlinson, S., 2001. Trypanosome lytic factors: novel mediators of human innate immunity. Curr. Opin. Microbiol. 4, 402–408.

Rebollar-Téllez, E.A., 1999. Kairomone Mediated Behaviour of Members of the Lutzomyia Longipalpis Complex (Diptera: Psychodidae). PhD Thesis. Keele University, Keele, UK.

Reeve, H.K., Sherman, P.W., 1993. Adaptation and the goals of evolutionary research. Q. Rev. Biol. 68, 1–32.

Ribeiro, J.M.C., Rossignol, P.A., Spielman, A., 1984. Role of mosquito saliva in blood vessel location. J. Exp. Biol. 108, 1–7.

Ribeiro, J.M.C., Rossignol, P.A., Spielman, A., 1985. *Aedes aegypti*: model for blood finding strategy and prediction of parasite manipulation. Exp. Parasitol. 60, 118–132.

Riley, E.M., Jobe, O., Blackman, M., Whittle, H.C., Greenwood, B.M., 1989. *Plasmodium falciparum* schizont sonicates suppress the lymphoproliferative response to mitogens and antigens in malaria-immune adults. Infect. Immun. 57, 3181–3188.

Riehle, M.M., Markianos, K., Niare, O., Xu, J., Li, J., Toure, A.M., Podiougou, B., Oduol, F., Diawara, S., Diallo, M., Coulibaly, B., Ouatara, A., Kruglyak, L., Traore, S.F., Vernick, K.D., 2006. Natural malaria infection in *Anopheles gambiae* is regulated by a single genomic control region. Science 312, 577–579.

Robert, V., Verhave, J.P., Carnevale, P., 1990. *Plasmodium falciparum* does not increase the precocious mortality rate of *Anopheles gambiae*. Trans. R. Soc. Trop. Med. Hyg. 84, 346–347.

Roberts, W., 1981. Probing by *Glossina morsitans morsitans* and transmission of *Trypanosoma (Nannomonas) congolense*. Am. J. Trop. Med. Hyg. 30, 948–951.

Rodriguez, M.H., Brown, D.N., Arrendondo-Jimenez, J.J., Villarreal, C., Loyola, E., Frederickson, C.E., 1992. Gonotrophic cycle and survivorship of *Anopheles albimanus* (Diptera: Culicidae) in southern Mexico. J. Med. Entomol. 29, 395–399.

Rogers, M.E., Chance, M.L., Bates, P.A., 2002. The role of promastigote secretory gel in the origin and transmission of the infective stage of *Leishmania mexicana* by the sandfly *Lutzomyia longipalpis*. Parasitology 124, 495–507.

Rogers, M.E., Ilg, T., Nikolaev, A.V., Ferguson, M.A., Bates, P.A., 2004. Transmission of cutaneous leishmaniasis by sand flies is enhanced by regurgitation of fPPG. Nature 430, 463–467.

Rossignol, P.A., Ribeiro, J.M.C., Spielman, A., 1984. Increased intradermal probing time in sporozoite-infected mosquitoes. Am. J. Trop. Med. Hyg. 33, 17–20.

Rossignol, P.A., Ribeiro, J.M.C., Jungery, M., Turell, M.J., Spielman, A., Bailey, C.L., 1985. Enhanced mosquito blood-finding success on parasitemic hosts—evidence for vector parasite mutualism. Proc. Natl. Acad. Sci. U.S.A. 82, 7725–7727.

Rossignol, P.A., Ribeiro, J.M.C., Spielman, A., 1986. Increased biting-rate and reduced fertility in sporozoite-infected mosquitoes. Am. J. Trop. Med. Hyg. 35, 277–279.

Rossignol, P.A., Rossignol, A.M., 1988. Simulations of enhanced malaria transmission and host bias induced by modified vector blood location behaviour. Parasitology 97, 363–372.

Sacks, D., Sher, A., 2002. Evasion of innate immunity by parasitic protozoa. Nat. Immunol. 3, 1041–1047.

Schaub, G.A., 1989. Does *Trypanosoma cruzi* stress its vector? Parasitol. Today 10, 463–468.

Schaub, G.A., 2006. Parasitogenic alterations of vector behaviour. Int. J. Med. Microbiol. 296, 37–40.

Scott, T.W., Lorenz, L.H., 1998. Reduction of *Culiseta melanura* fitness by eastern equine encephalomyelitis virus. Am. J. Trop. Med. Hyg. 59, 341–346.

Schwartz, A., Koella, J.C., 2001. Trade-offs, conflicts of interest and manipulation in *Plasmodium*–mosquito interactions. Trends Parasitol. 17, 189–194.

Schwartz, A., Koella, J.C., 2002. Melanization of *Plasmodium falciparum* and C-25 Sephadex Beads by Field-Caught *Anopheles gambiae* (Diptera: Culicidae) from Southern Tanzania. J. Med. Entomol. 39, 84–88.

Sinton, J.A., Shute, P.G., 1938. A report on the longevity of mosquitoes in relation to the transmission of malaria in nature. Rep. Public Health Med. Subj. 85, 1–45.

Solano, P., Gueguan, J.F., Reifenberg, J.M., Thomas, F., 2001. Trying to predict and explain the presence of African trypanosomes in tsetse flies. J. Parasitol. 87, 1058–1063.

Stearns, S.C., 1992. The Evolution of Life Histories. Oxford University Press, New York.

Taylor, P., Hurd, H., 2001. The influence of host haematocrit on the blood feeding success of *Anopheles stephensi*: implications for enhanced malaria transmission. Parasitology 122, 491–496.

Taylor, P., 2001. The Effects of Malaria Infection on the Blood Feeding Behaviour of Anopheline Mosquitoes. PhD Thesis. Keele University, Keele, UK.

Thevenaz, P.H., Hecker, H., 1980. Distribution and attachment of *Trypanosoma (Nannomonas) congolense* in the proximal part of the proboscis of *Glossina morsitans morsitans*. Acta Trop. 37, 163–175.

Thomas, F., Poulin, R., 1998. Manipulation of a mollusc by a trophically transmitted parasite: convergent evolution or phylogenetic inheritance? Parasitology 116, 431–436.

Thomas, F., Adamo, S., Moore, J., 2005. Parasitic manipulation: where are we and where should we go? Behav. Proc. 68, 185–199.

Tobie, E.J., 1965. Biological factors influencing transmission of *Trypanosoma rangeli* by *Rhodnius prolixus*. J. Parasitol. 51, 837–841.

Turell, M.J., Bailey, C.L., Rossi, C., 1984. Increased mosquito feeding on Rift Valley fever virus-infected lambs. Am. J. Trop. Med. Hyg. 33, 1232–1238.

Urban, B.C., Ferguson, D.J., Pain, A., Willcox, N., Plebanski, M., Austyn, J.M., Roberts, D.J., 1999. *Plasmodium falciparum* infected erythrocytes modulate the maturation of dendritic cells. Nature 400, 73–77.

Volf, P., Hajmova, M., Sadlova, J., Votypka, J., 2004. Blocked stomodeal valve of the insect vector: similar mechanismof transmission in two trypanosomatid models. Int. J. Parasitol. 34, 1221–1237.

Watt, D.M., Walker, A.R., 2000. Pathological effects and reduced survival in *Rhipicephalus appendiculatus* ticks infected with *Theileria parva* protozoa. Parasitol. Res. 86, 207–214.

Wekesa, J.W., Copeland, R.S., Mwangi, R.W., 1992. Effect of *Plasmodium falciparum* on blood feeding behavior of naturally infected Anopheles mosquitoes in western kenya. Am. J. Trop. Med. Hyg. 47, 484–488.

Xu, X., Kohsuke, S., Chiguang, F., Xiaozhong, X., Hua, S., Saho Ma, Makoto, K., Yoshihiro, A., 2001. Down-regulation of IL-12 p40 gene in *Plasmodium berghei*-infected mice. J. Immunol. 167, 235–241.

Zambrano-Villa, S., Rosales-Borjas, D., Carrero, J.C., Ortiz-Ortiz, L., 2002. How protozoan parasites evade the immune response. Trends Parasitol. 18, 272–278.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) 520–525
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Discussion

From population structure to genetically-engineered vectors: New ways to control vector-borne diseases?☆

O.A.E. Sparagano *, C.J. De Luna

School of Agriculture, Food, and Rural Development, Agriculture Building, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK

Received 27 March 2007; received in revised form 1 May 2007
Available online 6 May 2007

Abstract

Epidemiological studies on vectors and the pathogens they can carry (such as *Borrelia burgdorferi*) are showing some correlations between infection rates and biodiversity highlighting the ''dilution'' effects on potential vectors. Meanwhile other studies comparing sympatric small rodent species demonstrated that rodent species transmitting more pathogens are parasitized by more ectoparasite species. Studies on population structure and size have also proven a difference on the intensity of the parasitic infection. Furthermore, preliminary results in genetic improvement in mosquitoes (genetic markers, sexing, and genetic sterilization) will also increase performance as it has already been shown in field applications in developing countries. Recent results have greatly improved the fitness of genetically-modified insects compared to wild type populations with new approaches such as the post-integration elimination of transposon sequences, stabilising any insertion in genetically-modified insects. Encouraging results using the Sterile Insect Technique highlighted some metabolism manipulation to avoid the viability of offspring from released parent insect in the wild. Recent studies on vector symbionts would also bring a new angle in vector control capabilities, while complete DNA sequencing of some arthropods could point out ways to block the deadly impact on animal and human populations. These new potential approaches will improve the levels of control or even in some cases would eradicate vector species and consequently the vector-borne diseases they can transmit. In this paper we review some of the population biology theories, biological control methods, and the genetic techniques that have been published in the last years that are recommended to control for vector-borne diseases.

Keywords: Ticks; Mosquitoes; Mites; Biodiversity; Genetic engineering

1. Introduction

Vector-borne infectious diseases are a serious cause for socio-economical losses within the agricultural industry, especially in the topical latitudes. Each year, East Coast fever (ECF) causes US$168 million in losses to Africa, and as a result of the disease 1.1 million cattle annually die. Ticks and tsetse-borne diseases cost a further US$4–5 billion per year (Eisler et al., 2003). Forty percent of the world population is exposed to malaria, and every 30 s a child dies from the disease in Africa. Examples like these abound on every day life.

The most widely used control method for insects that act as vectors for both animal and plant populations are based on insecticides and pesticides. However, chemical resistance, animal welfare, environmental safety, national legislations, and the increased control by governmental agencies of these substances have put pressure in finding alternative ways to kill or decrease the population of vectors and in this way to control for the spread of diseases they can transmit.

Despite every effort to control vector-borne diseases, there are some factors that make the task even more difficult. Examples of these factors include wildlife acting as reservoirs, changes to the environment, insecticide resistance, vector movement between countries, problems with vaccines in some countries (interruption of the cold chain, funding, lack of protection, political and social aspects), pathogen resistance to drugs and vaccines, and variation of pathogenicity under field conditions. Global changes that involve the movement of animals (as pets or livestock) and human travelling have changed the exposure, sensitivity and susceptibility between hosts and pathogens in the presence of a constantly changing physical, biological and socio-economical environment (Sutherst, 2004). In this paper we discuss the population

☆ Related to an oral talk presented at MEEGID VIII, Bangkok, Thailand, December 2006.

* Corresponding author. Tel.: +44 191 2225071; fax: +44 191 2226720.
E-mail address: Olivier.Sparagano@ncl.ac.uk (O.A.E. Sparagano).

1567-1348/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2007.05.002

biology effect on transmission of vector-borne diseases and we also review the work on different biological and molecular approaches that the scientific community has been reporting over the last few years as alternative control methods.

2. The population biology effect

Keesing et al. (2006) stated that infectious diseases are ecological systems that involve the interactions among small to large networks of species. Therefore, there is a direct effect on species diversity and disease risk as a high host diversity is more likely to decrease disease risk. At the same time species diversity may impact on the reduction of an encounter, transmission reduction, susceptible host regulation, infected host mortality and the recovery of infected individuals (Keesing et al., 2006). Ostfeld and Keesing (2000) using a conceptual model explained the role of biodiversity on the transmission of Lyme disease (LD). The higher the species number that the ixodid tick vector parasite on, the less likely it may become infected with the spirochete (*Borrelia burgdorferi*) that produces LD. This is because many of the ticks never become infected because some hosts are inefficient to spirochete infections. This has been called as the ''dilution effect'' (Keesing et al., 2006).

The size of the population is another factor that affects the transmission of vectors to its hosts. Whiteman and Parker (2004) proved that the Galapagos' hawk, *Buteo galapagensis*, was parasitized by a larger number of two lice species when living in larger groups, when compared against birds that lived in smaller groups. Moreover, Durden et al. (2004) proposed that if a species acts as a reservoir of more vector-borne zoonotic pathogens when compared with another, it will also be parasitized by a higher number of ectoparasites. The rodent species, *Peromyscus gossypinus*, or cotton mouse, is known to be a reservoir of more vector-borne zoonotic diseases than the golden mice *Ochrotomys nuttalli*. The above paper confirms the hypothesis that the former rodent species is epidemiologically more important for the transmission of vector-borne disease than the latter (Durden et al., 2004).

3. New approaches against vectors

3.1. Arthropod vaccines

Evidence that vaccination with protein antigens is able to induce significant immunity against some arthropods has been gathered in recent years. Arthropods that feed on host blood have been targeted for vaccination, specifically; components of their saliva and midgut are the candidates to become antigens for vaccination (Mejia et al., 2006; Titus et al., 2006). Arthropod vaccination may affect the transmission of disease by affecting vector numbers, thus they could directly influence the incidence of the disease (Willadsen, 2004). Nevertheless, it must be taken into account that some vectors produce multiple infections, e.g. *Amblyomma variegatum* could transmit heartwater, Crimean-Congo hemorrhagic fever (CCHF), theileriosis and be associated with dermatophilosis. *Ixodes scapularis* ticks could infect its host with *Borrelia burgdorferi*, *Babesia microti,* and also produce human granulocytic encephalitis (now within the *A. phagocytophilum* complex), while *Rhipicephalus appendiculatus* ticks can be associated with the transmission of *Babesia bigemina*, *Theileria parva* (agent of East Coast fever) and the Thogoto and Nairobi sheep viruses.

Concealed antigens have been mentioned as one of the pathways for vaccine development, the other is the duplication of naturally acquired immunity (Willadsen, 1999).

The most successful anti-tick vaccine work has been the one against *Rhipicephalus* (*Boophilus*) *microplus*. Vaccination with recombinant *Rhipicephalus B. microplus* gut antigens has shown success in controlling tick infestations (de la Fuente et al., 2000). The use of the gut antigens Bm86, Bm91, and Bm95 from *Rhipicephalus B. microplus* is the basis of the commercial vaccines TickGARD and GAVAC used in Australia and Cuba, respectively (Willadsen, 2001). Another open possibility is the development of a vaccine that protects against several species of ticks, by using a highly conserved antigenic molecule; therefore, cross-reactions against different species of ticks are expected to be effective when immunising against a single species of tick (de la Fuente and Kocan, 2006). Even further, Mejia et al. (2006) have suggested the possibility of the creation of a pan-arthropod vaccine that targets some glycans attached to arthropod glycoproteins as potential antigens synthesised from *Drosophila melanogaster* cells.

The sheep scab mite *Psoroptes ovis* has also been subject to research in order to develop a vaccine to control for sheep scab. *P. ovis* allergens produce pathogenicity and also some immunological response (Huntley et al., 2004). This has created the opportunity of targeting the allergen, *Pso o* 1, as a candidate for the production of anti-*P. ovis* vaccine (Nisbet et al., 2007). Similar to *Pso o* 1, the allergens produced by the house dust mite, *Dermatophagoides pteronyssinus*, *Der p* 1 and *Der p* 2, have also been involved in the production of a DNA vaccine against the allergic reaction produced by the house dust mite (Kwon et al., 2003; HuangFu et al., 2006).

3.2. Biological control

Biocontrol agents are an area that has not been intensively researched and it has a lot of potential for the advancement of vector control. Entomopathogenic fungi have proven an effective control against vectors. *Metarhizium anisopliae* has been used against adult *Anopheles gambiae* mosquitoes for malaria control (Scholte et al., 2005). Fungus-infected *A. gambiae* had significantly shorter life spans compared with those of non-infected mosquitoes. Among the results reported by Scholte et al. (2005) they found that *M. anisopliae*-infected males and females lived on average for 3.70 and 3.49 days, respectively, meanwhile uninfected males and females had longer life spans of 5.88 and 9.30 days, respectively. Moreover, as a direct consequence of the shorter life span of the infected mosquitoes, up to a 75% reduction of the transmission intensity was reported, this of course reduced the risk of contracting malaria (Scholte et al., 2005).

Isolates of *Beauveria bassiana* have been used to control the poultry red mite *Dermanyssus gallinae* (Steenberg et al., 2006). *D. gallinae* is responsible for economic losses to the poultry industry, and it seriously affects the welfare of the bird (Chauve, 1998). Infections of *B. bassiana* proved to be persistent on the mites over time, and they also proved to be virulent, reducing the population growth of the mites where assessed on semi-controlled experiments (Steenberg et al., 2006).

Another recent approach is the use of vector-symbionts. Symbionts living in the alimentary tract of insects help them to provide the nutrients that are not taken from the diet. Genetically modified symbionts have been reintroduced into the insects, which are refractory to the parasites they used to transmit. An example of this technique is the work of Beard et al. (2001). They genetically modified various triatomine species that act as symbionts for *Rhodnius prolixus* that hosts the agent of the Chagas disease, *Trypanosoma cruzi*. Another example is the use of intracellular bacteria of the genus *Wolbachia*. They are often associated with the reproductive tract of arthropods. Therefore, the use of genetically modified *Wolbachia* spp. as a vector-symbionts method to block the transovarial transmission of pathogens has been described (Riehle and Jacobs-Lorena, 2005), and this technology may also be applied to the pathogens that reside or reproduce in other tissues (Beard et al., 1998). Other organisms used as entomopathogenic are some species of nematods; however, its value as potential biological control remains to be established (Samish et al., 2004).

Repellents, natural and synthetic, have also been used for centuries to deter vectors from parasitizing their hosts. Kim et al. (2004) used 56 essential oils from plant derivates to assess the acaricidal activity against the poultry red mite *Dermanyssus gallinae*. Of those 56, they found that in a filter paper contact bioassay, 100% mortality at 0.07 mg cm^{-2} was observed in bay, cade, cinnamon, clove bud, coriander, horseradish, lime dis 5F, mustard, pennyroyal, pimento berry, spearmint, thyme red and thyme white oils. Nevertheless, they also reported that the mortality significantly decreased at 0.02 mg cm^{-2}. In another experiment conducted in South Africa the efficacy of a synthetic repellent, diethyl-3-methylbenzamide (DEET), was evaluated at different concentrations for the control of *Amblyomma hebraeum* the principal vector of *Rickettsia africae*, which causes African tick bite fever. The results showed that DEET successfully killed the ticks (>90–77%) but only during the first 2 h after application, so as time went by, the efficacy was decreased significantly and after 3 h post-application the efficacy was below 70% (Jensenius et al., 2005).

The use of traps, and other techniques used, as attractants is another widely used technique to control for insects. Most of the experiments were based on the use of pheromones as attractant for vectors to be used as a trap. A method widely used is the impregnation of a trap with the pheromones and the subsequent adding of insecticides to kill the parasites, this could be used on the host or on the surrounding vegetation (see reviews by Soneshine, 2004, 2006).

Finally, the last example of biological control is the use of predators of ectoparasites, such as other arthropods (ants, beetles, etc.) and bird species. However, the population of predators are not dependant on the size of the vector population (Samish et al., 2004), and therefore the amount of success in controlling the vector is limited. Another major disadvantage concerning the use of this approach is that predators are only occasional feeders, and thus large numbers of predators would be needed to reduce the number of vectors, though, it could also alter the surrounding environment and it will produce major changes in the population of non-target species (Symondson et al., 2002). Nonetheless, the use of oxpeckers (*Buphagus africanus* and *B. erythrorhyncus*) has been noted, since these animals can eat up to a thousand ticks per bird in less than a week; since the population of oxpeckers have declined in recent years due to agricultural practices, some management plans have been put in place and there have been some attempts to reintroduce the birds in cattle production areas in Africa (Samish et al., 2004).

3.3. Genetically-engineered vectors

Genetic techniques rely on two different strategies of control mechanisms, population replacement and population suppression, both ideas are based on the introduction of genetic traits introduced into the wild by mating (Coleman and Alphey, 2004).

Among those genetic techniques identified as population replacement, RNA interference (RNAi) is one example. RNAi is to be considered as an acquired antiviral defence in eukaryotic organisms in which double stranded RNA (dsRNA) controls the process of gene expression (Sanchez-Vargas et al., 2004). In non-mammalian cells long dsRNA are used while in mammalian cells RNAi pathways can be triggered by synthetic short interfering RNA (siRNA) molecules or by DNA-based expression vectors expressing short hairpin (shRNA) molecules (Bernstein et al., 2001). Therefore, the host can control for the replication of viruses. This mechanism could explain why some mosquitoes are better vectors than others. Genetically modified mosquitoes that trigger the RNAi response against specific viruses may therefore resist the viral infection (Sanchez-Vargas et al., 2004). One example of this approach is the work done by Franz et al. (2006) in which they developed transgenic lines of the mosquito *Aedes aegypti* against infections of dengue virus. Bian et al. (2005) using RNAi and reverse genetics, created a transgenic alteration on the same species of mosquito to increase the susceptibility to the entomopathogenic fungus *Beauveria bassiana*.

Readers interested about the different RNAi pathways, their advantages and drawbacks can refer to review papers existing on this topic (e.g. Aljamali et al., 2002, 2003; Aigner, 2006).

Another technology that genetically modifies an organism is the use of transposable elements (TEs) also known as ‘‘jumping genes’’. TEs are large portions of DNA that move around the cell, and the genomes of many organisms contain numerous copies of TEs (Biemont and Vieira, 2006). Genomes are likely to be composed of two main types of TEs, the transposons (or ClassII), which are DNA-based elements and the retro-transposons (or Class I), which are based on RNA (Biemont and Vieira, 2005).

Recent developments in gene transfer technologies in mosquitoes, in which several transposable elements have been used to genetically transform mosquito species to retain genetically desirable characteristics (Tu and Coates, 2004). Kokoza et al. (2001) successfully modified the yellow fever mosquito *Aedes aegypti* by inserting a TE (called piggyBac) into its genome, producing transgenic lines that are able to pass the modification to further generations. This principle has been also used to transmit the capacity to resist infection from undesirable pathogens (Arensburger et al., 2005). Holt et al. (2002) sequenced the complete genome of the mosquito *Anopheles gambiae* and found the presence of TEs. A considerable amount of work has been undertaken to use TEs as gene vectors in Anopheline mosquitoes (Catterucia et al., 2000; Grossman et al., 2001; Perera et al., 2002). As a result, genetically-modified *Anopheles stephensi* showed a lower ability to transmit *Plasmodium berghei* (Ito et al., 2002). Later, Arensburger et al. (2005) also developed an active TE in the *Anopheles gambiae* genome that resists the infection of the *Plasmodium* parasite, thus controlling the spread of malaria. Recently, genetically modified *Anopheles stephensi* were used to block malaria transmission (Marrelli et al., 2007) opening a tremendous potential to block more vector-borne diseases.

3.4. Population suppression—the sterile insect technique

The sterile insect technique consists of the release of mass numbers of vectors that are sterilized by irradiation. The principle is based on releasing a large numbers of sterile males that mate with wild females, thus reducing the reproductive potential, and in this way it reduces the wild population overtime (Coleman and Alphey, 2004). It is a widely used method for biological pest control. It has already been successful in eliminating some agricultural pests such as the Mediterranean fruit fly, the screwworm fly, and the tsetse fly (Coleman and Alphey, 2004). Although it is environmentally friendly, this technology produces a decrease on fitness of the sterile population when compared with their wild counterparts, thus when they have to mate, their reduced fitness plays as a disadvantage and the sterile insects are ineffective at doing so. Drawbacks from irradiation techniques are now slowly replaced by genetic changes of transgenes (Gong et al., 2005) and similar autocidal studies using Sterile Acarine Technique (SAT) in ticks has also produced interesting results (de la Fuente et al., 2006c).

The ideal scenario is that the release of GM vectors into the wild must bear a similar genetic background as their natural counterparts, by releasing individuals with a different genetic background, the possibility of bringing the presence of novel alleles increase in the target population, and as a consequence the genetic makeup will be different and a change of fitness may be present (Zhong et al., 2006).

Could we have transgenic vectors with a reduction in pathogen transmission and a higher mating competitiveness? The use of genetically modified vectors is a promising alternative in the control of vector-borne diseases. Nevertheless, caution must be exercised since the initial optimism produced in the laboratory must be equalled to the one produced under field conditions. Monitoring must be followed after the release of these genetic engineered vectors to ensure the success of them in the wild and if such improved performance is passed onto the next mosquito generation for instance (Riehle and Jacobs-Lorena, 2005).

Following the progress done on RNAi in many living organisms such as *Drosophila*, nematodes or plants we see new publications for ticks as well. For instance, de la Fuente et al. (2006a, 2006c) working on the subolesin gene have reduced or blocked the reproduction performance in *Dermacentor variabilis* ticks and therefore it could also block the transmission of transovarial pathogens between female ticks and their offsprings. Labuda et al. (2006) working on *Rhipicephalus appendiculatus* ticks reduced the viraemia due to tick-borne encephalitis viruses (TBEV) by developing an anti-tick vaccine base on the tick cement protein (64TRP). RNAi experiments on mice and cattle have shown a reduction of the infectivity of both rickettsiae to the vertebrate hosts (de la Fuente et al., 2006b).

Going back to basics about the vector life cycle, behaviour and physiology could also bring more future targets, while the understanding of TEs on genomes could advance our knowledge of their genetic organisation and future manipulations. Currently, the mosquito DNA sequencing program has improved the potential to control for vector-borne diseases. Nevertheless, the general public has expressed some worries on the release of genetically modified organisms into the wild. However, the potential of these techniques to be used as a control of vector-borne diseases is powerful yet laborious.

References

Aigner, A., 2006. Gene silencing through RNA interference (RNAi) in vivo: strategies based on the direct application of siRNAs. J. Biotechnol. 124, 12–25.

Aljamali, M.N., Sauer, J.R., Essenberg, R.C., 2002. RNA interference: applicability in tick research. Exp. Appl. Acarol. 28, 89–96.

Aljamali, M.N., Bior, A.D., Sauer, J.R., Essenberg, R.C., 2003. RNA interference in ticks: a study using histamine binding protein dsRNA in the female tick *Amblyomma variegatum*. Insect Mol. Biol. 12, 299–305.

Arensburger, P., Kim, Y.J., Orsetti, J., Aluvihare, C., O'Brochta, D.A., Atkinson, P.W., 2005. An active transposable element, Herves, from the African malaria mosquito *Anopheles gambiae*. Genetics 169, 697–708.

Beard, C.B., Durvasula, R.V., Richards, F.F., 1998. Bacterial symbiosis in arthropods and the control of disease transmission. Emerg. Infect. Dis. 4, 581–591.

Beard, C.B., Dotson, E.M., Pennington, P.M., Eichler, S., Cordon-Rosales, C., Durvasula, R.V., 2001. Bacterial symbiosis and paratransgenic control of vector-borne Chagas disease. Int. J. Parasitol. 31, 621–627.

Bernstein, E., Denli, A.M., Hannon, G.J., 2001. The rest is silence. RNA 7, 1509–1521.

Bian, G., Shin, S.W., Cheon, H.M., Kokoza, V., Raikhel, A.S., 2005. Transgenic alteration of toll immune pathway in the female mosquito *Aedes aegypti*. Proc. Natl. Acad. Sci. U.S.A. 102, 13568–13573.

Biemont, C., Vieira, C., 2005. What transposable elements tell us about genome organization and evolution: the case of Drosophila. Cytogenet. Genome Res. 110, 25–34.

Biemont, C., Vieira, C., 2006. Junk DNA as an evolutionary force. Nature 443, 521–524.

Catterucia, F., Nolan, T., Loukeris, T.G., Blass, C., Savakis, C., Kafatos, F.C., Crisanti, A., 2000. Stable germline transformation of the malaria mosquito *Anopheles stephensi*. Nature 405, 959–962.

Chauve, C., 1998. The poultry red mite *Dermanyssus gallinae* (De Geer, 1778): current situation and future prospects for control. Vet. Parasitol. 79, 239–245.

Coleman, P.G., Alphey, L., 2004. Genetic control of vector populations: an imminent prospect. Trop. Med. Int. Health 9, 433–437.

de la Fuente, J., Kocan, K.M., 2006. Strategies for development of vaccines for control of ixodid tick species. Parasite Immunol. 28, 275–283.

de la Fuente, J., Rodriguez, M., Garcia-Garcia, J.C., 2000. Immunological control of ticks through vaccination with *Boophilus microplus* gut antigens. Ann. New York Acad. Sci. 916, 617–621.

de la Fuente, J., Almazan, C., Blas-Machado, U., Naranjo, V., Mangold, A.J., Blouin, E.F., Gortazar, C., Kocan, K.M., 2006a. The tick protective antigen, 4D8, is a conserved protein involved in modulation of tick blood digestion and reproduction. Vaccine 24, 4082–4095.

de la Fuente, J., Almazan, C., Blouin, E.F., Naranjo, V., Kocan, K.M., 2006b. Reduction of tick infections with *Anaplasma marginale* and *A. phagocytophilum* by targeting the tick protective antigen subolesin. Parasitol. Res. 100, 85–91.

de la Fuente, J., Almazan, C., Naranjo, V., Blouin, E.F., Meyer, J.M., Kocan, K.M., 2006c. Autocidal control of ticks by silencing of a single gene by RNA interference. Biochem. Biophys. Res. Commun. 344, 332–338.

Durden, L.A., Polur, R.N., Nims, T., Banks, C.W., Oliver, J.H., 2004. Ectoparasites and other epifaunistic arthropods of sympatric cotton mice and golden mice: Comparisons and implications for vector-borne zoonotic diseases. J. Parasitol. 90, 1293–1297.

Eisler, M.C., Torr, S.J., Coleman, P.G., Machila, N., Morton, J.F., 2003. Integrated control of vector-borne diseases of livestock-pyrethroids: panacea or poison? Trends Parasitol. 19, 341–345.

Franz, A.W.E., Sanchez-Vargas, I., Adelman, Z.N., Blair, C.D., Beaty, B.J., James, A.A., Olson, K.E., 2006. Engineering RNA interference-based resistance to dengue virus type 2 in genetically modified *Aedes aegypti*. Proc. Natl. Acad. Sci. U.S.A. 103, 4198–4203.

Gong, P., Epton, M.J., Fu, G., Scaife, S., Hiscox, A., Condon, K.C., Condon, G.C., Morrison, N.I., Kelly, D.W., Dafa'alla, T., Coleman, P.G., Alphey, L., 2005. A dominant lethal genetic system for autocidal control of the Mediterranean fruitfly. Nat. Biotecnol. 23, 453–456.

Grossman, G.L., Rafferty, C.S., Clayton, J.R., Stevens, T.K., Mukabayire, O., Benedict, M.Q., 2001. Germline transformation of the malaria vector, *Anopheles gambiae*, with the piggyBac transposable element. Ins. Mol. Biol. 10, 597–604.

Holt, R.A., Subramanian, G.M., Halpern, A., Sutton, G.G., Charlab, R., Nusskern, D.R., Wincker, P., Clark, A.G., Ribeiro, J.M., Wides, R., Salzberg, S.L., Loftus, B., Yandell, M., Majoros, W.H., Rusch, D.B., Lai, Z., Kraft, C.L., Abril, J.F., Anthouard, V., Arensburger, P., Atkinson, P.W., Baden, H., de Berardinis, V., Baldwin, D., Benes, V., Biedler, J., Blass, C., Bolanos, R., Boscus, D., Barnstead, M., Cai, S., Center, A., Chaturverdi, K., Christophides, G.K., Chrystal, M.A., Clamp, M., Cravchik, A., Curwen, V., Dana, A., Delcher, A., Dew, I., Evans, C.A., Flanigan, M., Grundschober-Freimoser, A., Friedli, L., Gu, Z., Guan, P., Guigo, R., Hillenmeyer, M.E., Hladun, S.L., Hogan, J.R., Hong, Y.S., Hoover, J., Jaillon, O., Ke, Z., Kodira, C., Kokoza, E., Koutsos, A., Letunic, I., Levitsky, A., Liang, Y., Lin, J.J., Lobo, N.F., Lopez, J.R., Malek, J.A., McIntosh, T.C., Meister, S., Miller, J., Mobarry, C., Mongin, E., Murphy, S.D., O'Brochta, D.A., Pfannkoch, C., Qi, R., Regier, M.A., Remington, K., Shao, H., Sharakhova, M.V., Sitter, C.D., Shetty, J., Smith, T.J., Strong, R., Sun, J., Thomasova, D., Ton, L.Q., Topalis, P., Tu, Z., Unger, M.F., Walenz, B., Wang, A., Wang, J., Wang, M., Wang, X., Woodford, K.J., Wortman, J.R., Wu, M., Yao, A., Zdobnov, E.M., Zhang, H., Zhao, Q., Zhao, S., Zhu, S.C., Zhimulev, I., Coluzzi, M., della Torre, A., Roth, C.W., Louis, C., Kalush, F., Mural, R.J., Myers, E.W., Adams, M.D., Smith, H.O., Broder, S., Gardner, M.J., Fraser, C.M., Birney, E., Bork, P., Brey, P.T., Venter, J.C., Weissenbach, J., Kafatos, F.C., Collins, F.H., Hoffman, S.L., 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science 298, 129–149.

HuangFu, T.Q., Lim, L.H., Chua, K.Y., 2006. Efficacy evaluation of Der p 1 DNA vaccine for allergic asthma in an experimental mouse model. Vaccine 24, 4576–7581.

Huntley, J.F., Machell, J., Nisbet, A.J., van den Broek, A., Chua, K.Y., Cheong, N., Hales, B., Thomas, W.R., 2004. Identification of tropomyosin, paramyosin, and apolipophorin/vitellogenin as three major allergens of the sheep scab mite, *Psoroptes ovis*. Parasite Immunol. 26, 335–342.

Ito, J., Ghosh, A., Moreira, L.A., Wimmer, E.A., Jacobs-Lorena, M., 2002. Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. Nature 417, 452–455.

Jensenius, M., Pretorius, A.M., Clarke, F., Myrvang, B., 2005. Repellent efficacy of four commercial DEET lotions against *Amblyomma hebraeum* (Acari: Ixodidae), the principal vector of *Rickettsia africae* in southern Africa. Trans. R. Soc. Trop. Med. Hyg. 99, 708–711.

Keesing, F., Holt, R.D., Ostfeld, R.S., 2006. Effects of species diversity on disease risk. Ecol. Lett. 9, 485–498.

Kim, S.I., Yi, J.H., Tak, J.H., Ahn, Y.J., 2004. Acaricidal activity of plant essential oils against *Dermanyssus gallinae* (Acari: Dermanyssidae). Vet. Parasitol. 120, 297–304.

Kokoza, V., Ahmed, A., Wimmer, E.A., Raikhel, A.S., 2001. Efficient transformation of the yellow fever mosquito *Aedes aegypti* using the piggyBac transposable element vector pBac[3xP3-EGFP afm]. Insect Biochem. Mol. Biol. 31, 1137–1143.

Kwon, S.S., Kim, N., Yoo, T.J., 2003. The effect of vaccination with DNA encoding murine T-cell epitopes on the Der p 1 and 2 induced immunoglobulin E synthesis. Int. Congr. 1240, 1189–1194.

Labuda, M., Trimnell, A.R., Lickova, M., Kazimirova, M., Davies, G.M., Lissina, O., Hails, R.S., Nuttall, P.A., 2006. An antivector vaccine protects against a lethal vector-borne pathogen. PLOS Pathog. 2, 251–259.

Marrelli, M.T., Li, C., Rasgon, J.L., Jacobs-Lorena, M., 2007. Transgenic malaria-resistant mosquitoes have a fitness advantage when feeding on Plasmodium infected blood. PNAS 104, 5580–5583.

Mejia, J.S., Bishop, J.V., Titus, R.G., 2006. Is it possible to develop panarthropod vaccines? Trends Parasitol. 22, 367–370.

Nisbet, A.J., MacKellar, A., McLean, K., Brennan, G.P., Huntley, J.F., 2007. Eukaryotic expression of recombinant Pso o 1, an allergen from *Psoroptes ovis*, and its localization in the mite. Parasitology 134, 83–89.

Ostfeld, R.S., Keesing, F., 2000. Biodiversity and disease risk: the case of Lyme disease. Conserv. Biol. 14, 722–728.

Perera, O.P., Harrel, R., Handler, A.M., 2002. Germ-line transformation of the South American malaria vector, *Anopheles albimanus*, with a piggyBaclEGFP transposon vector is routine and highly efficient. Insect Mol. Biol. 11, 291–297.

Riehle, M.A., Jacobs-Lorena, M., 2005. Using bacteria to express and display anti-parasite molecules in mosquitoes: current and future strategies. Insect Biochem. Mol. Biol. 35, 699–707.

Samish, M., Ginsberg, H., Glazer, I., 2004. Biological control of ticks. Parasitology 129 (Suppl.), S389–S403.

Sanchez-Vargas, I., Travanty, E.A., Keene, K.M., Franz, A.W.E., Beaty, B.J., Blair, C.D., Olson, K.E., 2004. RNA interference, arthropod-borne viruses, and mosquitoes. Virus Res. 102, 65–74.

Scholte, E.J., Ng'habi, K., Kihonda, J., Takken, W., Paaijmans, K., Abdulla, S., Killeen, G.F., Knols, B.G.J., 2005. An entomopathogenic fungus for control of adult African malaria mosquitoes. Science 308, 1641–1642.

Soneshine, D.E., 2004. Pheromones and other semiochemicals of ticks and their use in tick control. Parasitology 129 (Suppl.), S405–S425.

Soneshine, D.E., 2006. Tick pheromones and their use in tick control. Annu. Rev. Entomol. 51, 557–580.

Steenberg, T., Kilpinen, O., Moore, D., 2006. Fungi control of the poultry red mite, *Dermanyssus gallinae*. DIAS report No. 119, 71–73.

Sutherst, R.W., 2004. Global change and human vulnerability to vector-borne diseases. Clin. Microbiol. Rev. 17, 136–173.

Symondson, W.O.C., Sunderland, K., Greenstone, M.H., 2002. Can generalist predators be effective biocontrol agents? Annu. Rev. Entomol. 47, 561–594.

Titus, R.G., Bishop, J.V., Mejia, J.S., 2006. The immunomodulatory factors of arthropod saliva and the potential for these factors to serve as vaccine targets to prevent pathogen transmission. Parasite Immunol. 28, 131–141.

Tu, Z., Coates, C., 2004. Mosquito transposable elements. Insect Biochem. Mol. Biol. 34, 631–644.

Whiteman, N.K., Parker, P., 2004. Effects of host sociality on ectoparasites population biology. J. Parasitol. 90, 939–947.

Willadsen, P., 1999. Immunological control of ectoparasites: past achievements and future research priorities. Genet. Anal. Biomol. Eng. 15, 131–137.

Willadsen, P., 2001. The molecular revolution in the development of vaccines against ectoparasites. Vet. Parasitol. 101, 353–367.

Willadsen, P., 2004. Anti-tick vaccines. Parasitologys 129 (Suppl.), S367–S387.

Zhong, D., Temu, E.A., Guda, T., Gouagna, L., Menge, D., Pai, A., Githure, J., Beier, J.C., Yan, G., 2006. Dynamics of gene introgression in the African malaria vector *Anopheles gambiae*. Genetics 172, 2359–2365.

Available online at www.sciencedirect.com
ScienceDirect
Infection, Genetics and Evolution 8 (2008) S1–S49
Infection, Genetics and Evolution
www.elsevier.com/locate/meegid

Abstracts

Abstract for the 8th International Meeting on Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases Royal River Hotel, Bangkok, Thailand, 30 November–2 December 2006

Available online 20 January 2008

Plenary lectures

The molecular epidemiology of enteric protozoan infections—Emerging issues and paradigm shifts

R.C. Andrew Thompson

World Health Organization Collaborating Centre for the Molecular Epidemiology of Parasitic Infections, School of Veterinary and Biomedical Sciences, Murdoch University, Murdoch, WA 6150, Australia.

In recent years a variety of issues have emerged concerning the epidemiology of zoonotic protozoan infections that result from the ingestion of environmentally resistant infective stages. They have many features in common regarding their transmission, which can be direct, or via water or food, and most exhibit low host specificity. Although they have been the subject of research for many years, recent studies have raised fundamental questions concerning our understanding of the epidemiology of infections with these parasites. *Giardia* and *Blastocystis* both have wide host ranges and are genetically very divergent yet how this variability is reflected in terms of zoonotic potential, clinical significance and virulence is not clear. With *Cryptosporidium*, many taxonomic and epidemiological questions have been resolved but recent studies have not only questioned *Cryptosporidium*'s phylogenetic affinities, but have also revealed new aspects about its life cycle and development. These findings will have a major impact on both surveillance and control. In the case of *Toxoplasma*, recent studies in domestic animals and wildlife have raised questions about how the parasite is maintained in nature. In particular, the role of vertical transmission in wildlife populations may have been underestimated. In the case of *Blastocystis*, *Entamoeba coli*, *Chilomastix* and *Dientamoeba*, they have been largely overlooked in terms of their impact on public health yet their common, and sometimes concurrent occurrence, has raised questions about their clinical and zoonotic potential.

These emerging issues will be discussed with emphasis on how molecular tools and epidemiological studies can help resolve these questions.

Leishmania and sand flies: Parasite–vector co-evolution or opportunism?

Paul Bates

Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK.

Leishmaniasis is an emerging and re-emerging disease in several parts of the world. One of the principle areas of investigation that helps us to understand a change in disease epidemiology are the key biological factors responsible for disease transmission. For vector-borne diseases like leishmaniasis an understanding of vector specificity and the mechanics of transmission are two such key factors. Recent work in several laboratories has shown that *Leishmania* and sand flies provide a range of examples from both ends of the spectrum with regard to vector specificity. For example, *Leishmania major* and *Phlebotomus papatasi* appear to be a very specific parasite–vector combination, and co-evolution has driven the molecular differentiation of a specific ligand on the surface of the parasite that binds to a corresponding galectin on the wall of the sand fly midgut. Thus *P. papatasi* is a representative member of a group that can be called the ''restricted'' vectors of leishmaniasis. At the other extreme lies *Lutzomyia longipalpis*, which transmits *L. infantum* in Central and South America. There is now strong evidence that this parasite has only very recently been introduced into the Americas from Europe in last few hundred years, probably when European colonists brought *L. infantum*-infected dogs to America. *Lutzomyia longipalpis* was already there and was adopted as a vector by the incoming parasites, taking over from *P. perniciosus* and *P. ariasi* found in Southern Europe. All of these vectors belong to a second group, the ''permissive'' vectors of leishmaniasis. Although in nature they usually only transmit one particular species of parasite, this

1567-1348/$ – see front matter
doi:10.1016/j.meegid.2008.01.008

appears to be more due to ecological constraints rather than any intrinsic barrier, as under laboratory conditions they can support the development of many species of *Leishmania*. Thus this parasite–vector combination can be regarded as a case of evolutionary opportunism. Current work is being pursued to investigate the molecular basis of this opportunism. Once the parasite has established an infection in a particular sand fly it must then overcome the challenge of transmission by bite: how can the parasite travel against the flow of an incoming bloodmeal? Recent work has shown that a gel-like material secreted by parasites in the sand fly gut plays a key role in promoting transmission. The so-called promastigote secretory gel (PSG) creates a ''blocked fly'' that cannot feed properly. This material must be egested by regurgitation before bloodfeeding can proceed, thereby egesting the infective parasites at the same time. This mechanism of transmission appears to be common amongst the *Leishmania* parasite–vector combinations examined so far. It may have evolved either before or after the specialisation of individual *Leishmania* species to a particular vector, thus representing either a conserved or convergent evolutionary response. What lies ahead for SE Asia? Leishmaniasis may remain a relatively rare disease, but the emergence of an epidemic of cutaneous leishmaniasis in Sri Lanka in the past 5 years, cases of visceral disease in Thailand, and reports of leishmaniasis in kangaroos in Australia all illustrate that complacency is dangerous. As the examples mentioned above show, both parasites and vectors have shown themselves capable of adapting to new circumstances, either by the spread of a well-established parasite–vector combination due to changes in ecology or the establishment of a novel parasite–vector partnership.

Pharmacogenomics of HIV

Wasun Chantratita

Virology and Molecular Microbiology Unit, Department of Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok 10400, Thailand.

In post-genomics era, we have found the additive effect of individual genetic variations in loci encoding metabolic enzymes, drug transporters, cell surface markers, and cellular growth and differentiation factors may play a significant role in the variability of response and toxicity of a number of drugs. The ''one-size-fits-all'' regimen of antiretroviral treatments results in interpersonal variation in drug concentrations and differences in susceptibility to drug toxicity. Many of the antiretrovirals are metabolized by polymorphically expressed enzymes (cytochrome P450, CYP450; glucuronyl transferase, GT) and/or transported by drug transporters (ABC and SLC families). The nonnucleoside reverse transcriptase inhibitors (NNRTIs), nevirapine and efavirenz, are metabolized primarily by CYP2B6. The associations have been identified between a frequent CYP2B6 variant (G516T) and NNRTI pharmacokinetics. Greater plasma efavirenz exposure was predicted by CYP2B6 G516T and recent data suggest that G516T also predicts nevirapine exposure. Study the effect CYP2B6 polymorphism in mother-to-child HIV transmission of single-dose Nevirapine is currently studied in Thailand. The clearest association between genetic variants and response relates to the hypersensitivity reaction that occurs with abacavir. The identification that the major histocompatibility complex haplotype acts as a strong genetic predisposing factor which can be translated into a pharmacogenetic test. However, much more work needs to be done to define the genetic factors determining response to antiretroviral agents. In Thailand, pharmacogenomics project was established in 2003. Study of allele frequency and linkage disequilibrium of markers in drug related genes loci are relevant to the objective of this project. We genotyped 1536 haplotype tagging SNPs known polymorphic sites in 182 drug related genes in 280 unrelated healthy Thai samples, which comprises 70 samples from each of four geographical Thai populations: North, Northeastern, Central, and South. This data is crucial for pharmacogenomics case–control association studies with clinical records.

Molecular epidemiology of important bacterial pathogens in India: Ancient origins, current diversity and future epidemics

Seyed E. Hasnain[1,2], Niyaz Ahmed[1]

[1]*Pathogen Evolution Group, Centre for DNA Fingerprinting and Diagnostics, Hyderabad, India.*

[2]*University of Hyderabad, Hyderabad, India.*

Mycobacterium tuberculosis, leptospira and *Helicobacter pylori* are some of the bacterial pathogens that trigger diseases with a complex interplay between infection dynamics, pathogen biology and host immune responses. The whole genome sequence determination has greatly facilitated our understanding of these pathogens. Tuberculosis is the disease with a highest morbidity and mortality worldwide. The disease haunts millions of people in India with a huge death rate. The genetic diversity and evolutionary history of the underlying *M. tuberculosis* strains are largely unknown in the context of this country that has earned dubious distinctions for tuberculosis prevalence. Our ongoing, large-scale analysis of hundreds of strains of tubercle bacilli highlighted a clear predominance of ancestral *M. tuberculosis* genotypes in the Indian subcontinent, compared to other regions of the world, and support the opinion that India is a historically ancient endemic focus of tuberculosis. It is hypothesized that such 'ancient' bacilli are relatively 'docile' than some of the highly 'killer' ones such as the highly disseminating Beijing types which harbor inherent propensity to acquire multiple drug resistance (MDR) and are spreading in India through major metropolitan cities. Beijing strains are likely to evade and replace ancestral reservoirs of *M. tuberculosis* in the country. If that happens, India will probably face large, institutional outbreaks involving hospital wards, prisons, schools, etc. This is perhaps a major issue that needs to be addressed in the post-genomic scenario, with the same magnitude of zeal that researchers have shown towards drug discovery and diagnostic or vaccine development. Leptospirosis is another major pestilence, a worldwide zoonosis caused by the spirochetes of the genus Leptospira. The leptospires have been extremely diverse pathogens having more than three hundred different strains or serovars with specific geographic distribu-

tion. But this enormous inventory of serovars, based mainly on an ever-changing surface antigen repertoire, throws an artificial and unreliable scenario of strain diversity. It is therefore difficult to track strains whose molecular identity keeps changing according to the host and the environmental niches they inhabit and cross through. To address this problem, we have developed highly sophisticated genotyping systems based on integrated genome analysis approaches to correctly identify and track leptospiral strains. These approaches are expected to greatly facilitate epidemiology of leptospirosis apart from deciphering the origins and evolution of leptospires in a global sense. The human gastric pathogen *H. pylori* is presumed to be co-evolved with its human host and is again a very highly diverse and robust pathogen. Our 'geographic genomics' study tests the theory that *H. pylori* existed in humans as a benign bacterium for thousands of years until it acquired some virulence factors from the microorganisms abundant in the human societies of the neolithic period, after the domestication of agriculture and livestock. We found traces of East Asian ancestry in the gene pool of Native Peruvian strains (Amerindian?). This finding supports ancient human migration across the Bering-strait (20,000 years BP). We also attempted to support the idea that the major single virulence factor of the bacterium, the *cag* Pathogenicity Island (*cag*PAI) was acquired during different times, at different places in the world and from a 'local' microbial source. We followed this with theoretical approaches to find significant overlap among the *H. pylori* population expansion time and domestication of agriculture in the world. This study provides some new insights into the ancient origins and diversity of *H. pylori* and the significance of such diversity in the development of gastroduodenal pathology. Why has this bacterium survived for this long time in humans? Does this association makes the colonization beneficial or of low biological cost? These are the questions that need to be answered in the near future.

Viral population size is a key element in the risk assessment of the emergence in humans of a pandemic prone H5N1 avian influenza virus

Jean-Claude Manuguerra

Laboratory for Urgent Response to Biological Threats, Institut Pasteur, Paris, France.

The surge of the global avian influenza epizootic caused by genotype Z H5N1 highly pathogenic avian influenza viruses (HPAIVs) has posed numerous questions, in particular to risk managers and policy makers. Scientific knowledge is limited on many aspects of the ecology and environmental properties of HPAIVs, in particular H5N1. In addition to being an animal health issue with strong impact on human nutrition and socio-economics consequences, the current widespread epizootic has spilled over as a human health issue. Indeed, some 250 zoonotic cases of H5N1 infections have been reported worldwide since the end of 2003. Current H5N1 HPAIVs are however poorly transmissible from domestic birds to humans and need specific conditions to achieve this passage. Besides, virus transmission between humans is rare and extremely inefficient. This is due to two probable main reasons: 1/in humans, H5N1 HPAIVs find their preferred receptor structures (terminal sialic acid moieties in α 2,3 bonds) in the lower respiratory tracts (particularly in alveolar cells), 2/their optimum temperature of replication is higher than the temperature of the upper parts of the human respiratory tract. These facts could make it difficult for the virus to reach its proper targets in humans during the contamination process and could confine the virus deep in the lungs without possibility of easy exit, necessary for virus transmission. In the past, new virus subtypes emerged in the human population either by reassortments between human/mammalian and avian influenza viruses, as probably happened around 1957 and again around 1968, or by accumulation of point mutations as probably occurred with the precursor of the Spanish influenza virus. Indeed, some residues have been pointed out as important for the adaptation to new hosts and their accumulation could pave the way to a virus adapted to humans: 1/amino acid (AA) 627 on PB2 is probably involved in temperature dependence, 2/AA 223 in the haemagglutinin is involved in binding to terminal sialic acid moieties, which vary from one host species to another and within a host species from one tissue to another. Other determinants, probably in the NP or NS genes may greatly contribute to viral adaptation to their hosts. Influenza viruses are present in the form of quasi-species, *i.e.* populations of viral genomes bearing point differences between them. Viral diversity increases the probability of a group of minority viral genomes to harbour a set of mutations directly involved in an increased viral capacity for human-to-human transmission. Viral diversity depends both on virus intrinsic variation capabilities and viral population size. Influenza virus polymerase complexes are error prone and generate frequent point mutations. When a virus succeeds in changing host, its mutation rates seems generally higher in the new host from a phylogenetic viewpoint. This is also true among birds when an avian influenza virus (AIV) jumps from a duck species to chickens or turkeys. In the past, the hypothesis has been raised according to which precursor viruses would pre-exist in their current host where they acquire the necessary set of mutations through a hypermutation mechanism. This would be due to a polymerase complex with an error rate higher than that of other viruses, following the acquisition of point mutation (mutator mutations) affecting the enzyme fidelity. *In vitro* studies using avian-like influenza A (H1N1) viruses introduced in the pig population in Germany during the early 1980s suggest that there is no such thing as mutator mutations. Their conclusion was that the increased viral diversity was linked more to the global size of the virus population rather than to an enhanced mutation capacity of the virus. Applying this to the H5N1 current situation, it is probable that the animal host demographic factor, especially in domestic flocks, is a critical factor in viral diversity. For example, the poultry population increased from 1.1 billions in 1980 to 4.9 billions in 2002 in China only, offering a possibility of vast virus populations present at any one time in domestic poultry. Virus global maintenance in nature is a key element to understand its population dynamics. Data from the literature on AIV worldwide and long-term cycles in birds and in the environment are rather limited and there is a lot more to understand. Using a virus population

approach and basing it on sequence variation data, it should be possible to estimate the risk of a set of mutations to occur and thus the risk of viral emergence by accumulation of point mutations.

Genome scan study of clinical malaria in Senegal and Thailand

A. Sakuntabhai

Sakuntabhai A. Laboratoire de la génétique de la prédisposition aux maladies infectieuses, Institut Pasteur, 28 rue du Dr Roux, 75015 Paris, France.

Malaria has exerted considerable selective pressure on the human genome, most notably apparent in the prevalence of hemoglobin mutations in regions endemic for malaria. Epidemiological studies in regions of high malaria endemicity have consistently shown that the severity of disease considerably decreased after the first years of life, whereas parasite prevalence and incidence remain high throughout adolescence and only decrease slowly in adults. Knowledge of the relationship between parasite infection and disease remains one of the major enigmas of malaria epidemiology. No clear picture of the mechanisms underlying naturally acquired immunity to malaria or disease transmission have yet emerged. We carried out a human genetic study of two well-defined cohorts in whom malaria parameters were recorded longitudinally from two continents, Senegal and Thailand. The major difference apart from genetic background between the two cohorts is the presence of *Plasmodium vivax* in Thailand. We first estimated genetic effect for each phenotype as quantitative traits by mean of variance component. We found that number of clinical malaria attacks for the three species (*P. falciparum* (PF), *P. vivax* and *P. ovale*) and trophozoite density of PF are significantly under human genetic influence. In addition human genetic factors showed significant effect on gametocytogenesis of PF, which may influence transmission of the disease. We performed genome screening linkage analysis and tested the effect some known and candidate genes. We confirmed the previous finding of linkage on chromosome 5q31 (*Pfil1*) with parasite infection level. We found a new region on chromosome 5p15, which showed linkage to clinical PF attacks both in Senegal and Thailand. There are genes involved in complement activation, cytokines, etc. We planned to perform systematic screening of this region using information from the public database.

Defining risk to pathogenic infections: Utilizing the HAPMAP database and broad based screens to discover host susceptibility genes to infectious diseases

J. Sheng[1,2], J. Smith[1,3], J. Murray[4], T. Hodge[4], D.H. Rubin[1,2]

[1]*Research Medicine, VA TVHS, United States.*

[2]*Division of Infectious Diseases, United States.*

[3]*Division of Genetics, Department of Medicine, Vanderbilt University, United States.*

[4]*Department of Infectious Diseases, College of Veterinary Medicine, University of Georgia, United States.*

Background: The susceptibility to infection and subsequent manifestations of disease is dependent upon the host and pathogen. Pathogen-encoded proteins have been discovered that regulate specific features of pathogen behavior and affect its survival in the host. Host factors, which affect pathogen survival and can be manipulated to enhance or inhibit infection, have been more difficult to discover.

Methods: We have utilized a process of random insertional mutagenesis coupled with siRNA knockdown of gene expression to discover and validate cellular genes that play roles in various aspects of intracellular pathogen replication. We used the dbSNP database of NCBI to view the single nucleotide polymorphism (SNPs) in these genes, which we then categorized for potential function by virtue of predicted alteration of protein structure or transcript processing. The HAPMAP database provides an assessment of the major haplotypes (ancestral fragments of DNA harboring a specific series of alleles at variant positions) in a given region of the genome (and their frequencies) in a sample of Caucasian, Chinese, Japanese, or Yoruban populations.

Results: The human viruses, reovirus, Ebola and Marburg, were used for selection of mutant cells, which were resistant to lytic infection. HIV was selected as the virus of interest to validate whether the mutant gene had broad based association with virus replication. SNPs were found in validated genes that were predicted to affect the coding sequence or non-coding regions that may affect transcription or translational efficiencies. It was found that for some of the candidate genes, haplotype frequencies were notably different among Caucasians compared to peoples from Asia or Africa.

Conclusions: Mammalian genes were discovered that have roles in infection of Marburg and Ebola viruses, reovirus and HIV. These genes contain potentially functional genetic variation of varying frequency across major populations. Genetic variation of these candidate host genes may be subject to selective pressure by pathogens and may modify susceptibility and disease course following exposure to a potential pathogen. Further analysis will help to develop genetic profiles, which can be used to personalize medicine and target therapeutics to at risk populations.

Today knowledge and future challenges on human fascioliasis in Asia: The who initiative

Santiago Mas-Coma

Departamento de Parasitología, Facultad de Farmacia, Universidad de Valencia, Av. Vicent Andres Estelles s/n, 46100 Burjassot, Valencia, Spain.

Fascioliasis is an important disease caused by two trematode species: *Fasciola hepatica*, present in Europe, Africa, Asia, the Americas and Oceania, and *F. gigantica*, mainly distributed in Africa and Asia. Human fascioliasis was considered a secondary disease until the mid-1990s. This old disease has a great expansion power thanks to the large colonization capacities of its causal agents and vector species, and is at present emerging or re-emerging in many countries, including both prevalence and intensity increases and geographical expansion. WHO (Headquarters Geneva) decided to launch a worldwide initia-

tive against human fascioliasis including two main axes: (A) transmission and epidemiology studies; (B) control activities by mainly treatments with triclabendazole (Egaten®), a single-dose highly effective drug. Results obtained during the last years have furnished numerous hitherto-unknown aspects and new information which have given rise to a complete new general picture of this disease, explaining why human fascioliasis has recently been included within the list of important human parasitic diseases. Fascioliasis is the vector-borne disease presenting the widest latitudinal, longitudinal and altitudinal distribution known. Recent studies have shown it to be an important public health problem. Human cases have been increasing in 51 countries of the five continents. Recent papers estimate human infection up to 17 million people, or even higher depending from the hitherto unknown situations in many countries, mainly of Asia and Africa. Major health problems are known in Andean countries, the Caribbean, northern Africa, and western Europe. In Asia, the area of most concern is the region around the Caspian Sea (Iran and neighbouring countries). Moreover, data from the beginning of this new century indicate that south-east Asian countries may also be seriously affected, with around 500 cases in the 2002–2003 period and up to 2000 cases from the beginning of 2006 to nowadays in Vietnam. When comparing different human endemic areas, a large diversity of situations and environments appear. Fascioliasis in human hypo- to hyperendemic areas appear to present, in the different continents, a very wide spectrum of transmission and epidemiological patterns related to the very wide diversity of environments. This large diversity indicates that, once in a new area where the disease is emerging, studies must be performed from zero and shall comprise a multidisciplinary approach to assess which kind of epidemiological pattern are we dealing with. Within this multidisciplinarity, molecular epidemiology studies become crucial. Molecular markers developed during recent years shall be applied to both liver flukes from humans and animals and to freshwater lymnaeid transmitting snails, in order to establish which combined haplotypes are involved in the disease transmission locally. In Asia, molecular epidemiology studies performed in the area around the Caspian Sea show that the transmission pattern may be very complicated due to the overlapping of both *F. hepatica* and *F. gigantica*, the appearance of intermediate fasciolid forms, and the participation of lymnaeid vector species belonging to different groups as *Radix*, *Galba/Fossaria*, stagnicolines and *Pseudosuccinea*. A similar situation may be expected throughout other Asian regions, as in Vietnam and neighbouring countries. The fluke-snail host species specificity factor plays a fundamental role, although the domestic animal fauna, mainly livestock (mainly sheep, cattle, buffaloes, goats, donkeys and pigs) but sometimes also sylvatic herbivorous mammals as lagomorphs and rodents, are also worth noting. Molecular techniques as DNA target sequencing, single nucleotide polymorphisms (SNPs) and microsatellites are of great help to assess the transmission patterns and origins of human contamination, in the way to establish the appropriate individual prophylaxis and general control measures.

Human immune response gene polymorphism versus HIV-1 and dengue virus diversity in SE Asians

Henry Stephens

Centre for Nephrology, University College London, Royal Free Campus, Rowland Hill Street, London NW3 2PF, UK.

The immune response to pathogens is dependent on the presentation of microbial peptides by human leukocyte antigens (HLA) to T cells and natural killer (NK) cells. The genes encoding HLA and killer immunoglobulin-like receptors (KIR) are highly polymorphic. This diversity has functional implications and is likely to be driven by microbial selection acting on HLA and KIR gene loci. Different ethnic groups vary in their HLA and KIR allele profiles, which in turn can act as highly informative correlates of ethnicity in anthropological studies. The analysis of HLA and KIR gene profiles in ethnic Thais, has revealed that this major ethnic group is highly representative of the overall gene pool within the large populations of mainland SE Asia. Thus, this geographic region is most suitable for large-scale population-based genetic epidemiological studies of emerging infectious diseases such as HIV-1 and dengue, which are of increasing public health concern. HLA and KIR association studies with HIV-1 have been performed in numerous ethnic groups. A variety of effects have been observed particularly with HLA-B57, -B27 and -A11 molecules. There is evidence that the diversity of HIV-1 clades and recombinants infecting different populations is being driven by immune responses controlled by polymorphic HLA molecules. Such an effect may well be responsible for the prevalence of HIV-1 clade E or the CRF01_A/E recombinant in the ethnic Thai, Cambodian Khmer and Vietnamese Kinh populations, while HIV-1 clades B and C and recombinants thereof have seeded predominantly into the more northern Sino–Tibetan–Burman populations of this region. By contrast, all four of the major dengue virus serotypes are known to circulate in mainland SE Asian populations. There is evidence in ethnic Thais that the outcome of exposure to dengue virus in previously exposed and immunologically primed individuals, associates with HLA-A2, -B5 and -B15 molecules, depending on the dengue serotype responsible for secondary infections. Taken together, these studies are of relevance to the design, testing, and implementation of new vaccine control programmes in populations at risk of exposure to HIV-1 and dengue.

Drug targets and drug resistance in malaria

Worachart Sirawaraporn

Department of Biochemistry and Center for Bioinformatics and Applied Genomics, Faculty of Science, Mahidol University, Rama VI Road, Bangkok 10400, Thailand.

Malaria is one of the most life-threatening infectious diseases in tropical and sub-tropical countries. The disease claims approximately 3–5 millions people each year. The emergence of drug-resistance *Plasmodium falciparum* to almost all the currently available antimalarial drugs in many regions of the world has caused treatment of malaria increasing problematic and thus there is an immediate need to search and identify new targets,

develop new and effective antimalarial agents, and understand the molecular basis of drug resistance in malaria parasites. So far, resistance in malaria has been found to be associated with specific single nucleotide polymorphisms (SNPs) in the gene, i.e. *pf*MDR1 (N86Y) and *pf*CRT (K76T) are associated with quinoline resistance, whereas *pf*DHFR (N51I, C59R, S108N and I164L) correlated with resistance of antifolate antimalarial drugs. The DHFR of *P. falciparum* (*pf*DHFR) represents one of the most well-defined drug targets in malaria. Research on *pf*DHFR including gene cloning, expression, generation of mutants resistant to inhibitors and structural studies during the past two decades has contributed tremendously towards the understanding of antifolate binding and molecular mechanism of antifolate resistance in malaria. Studies of malarial DHFR will be discussed with respect to the interactions to malarial thymidylate synthase (TS) domain. The results could provide insights into better understanding of how effective inhibitors could be developed in order to overcome malaria resistance.

Acknowledgements

Works in WS's Laboratory have been supported by TDR grants.

Environmental change, infectious disease emergence, and dengue

Bruce A. Wilcox

Department of Tropical Medicine, Medical Microbiology, and Pharmacology, University of Hawaii Manoa, 651 Ilalo Steet, Biosciences Building 320, Honolulu, HI 96813, USA.

Regional environmental change, driven largely by globalization and population growth, with associated increases in resource consumption and waste generation, plays a primary role in the emergence of infectious disease, especially in tropical developing regions. Associated land use and transformation of resource production (urbanization, agricultural expansion and intensification, and natural habitat alteration), have produced changes in ecological systems, notably in landscapes and, in turn, their natural communities and ultimately in their pathogen, animal host, and human populations. Thus, the altered ''host–pathogen'' dynamics facilitate novelty, including exchange of genetic material among pathogens, resulting in rapid adaptation by the pathogens and more frequent generation of novel pathogen variants. Some will be more virulent, infective, and/or capable of enhanced transmission, contributing to disease reemergence or emergence. Factors related to public health infrastructure and climate variability, and their interactions with regional environmental change, also contribute significantly to disease emergence. Dengue and dengue hemorrhagic fever (DHF) is possibly the clearest case of disease re-emergence and emergence associated with regional environmental change, specifically urbanization. Failure to effectively control of dengue/DHF in many regions argues for new approaches that integrate research on the vector, pathogen, human and environment within a defined ecosystem. Classical ecological concepts are key to understanding population, community, and ecosystem level dynamics influencing disease emergence. But more recent advances and research tools in evolutionary ecology are also fundamental to understanding both vector and pathogen transmission dynamics that underlying emergence. Integrating this research with social ecological concepts represents a promising new, transdisciplinary approach to dengue control.

Symposiums

Symposium "Coevolution host pathogen 1"

(1) Peopling of South America and South Asia insights through *Helicobacter pylori* genomics

Niyaz Ahmed

Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad 500076, India.

The human gastric pathogen *Helicobacter pylori* is co-evolved with its host and therefore, origins and expansion of multiple populations and sub populations of *H. pylori* mirror ancient human migrations. Ancient origins of *H. pylori* in the New World and in India are debatable. It is not clear how different waves of human migrations in these large continents contributed to the evolution of strain diversity of *H. pylori*. We tried to address these issues through mapping genetic origins of *H. pylori* of native Peruvians (of Amerindian ancestry) and Indians and their genomic comparison with hundreds of isolates from different geographic regions. For this purpose, we attempted to dissect genetic identity of strains by fluorescent amplified fragment length polymorphism (FAFLP) analysis, multilocus sequence typing (MLST) of the seven housekeeping genes (*atp*A, *efp*, *ure*I, *ppa*, *mut*Y, *trp*C, *yph*C) and the sequence analyses of the *bab*B adhesin and *oip*A genes. The whole *cag* pathogenicity-island (*cag*PAI) from these strains was analyzed using PCR and the geographic type of *cag*A phosphorylation motif EPIYA was determined by gene sequencing. In case of South American *H. pylori* populations, we observed that while European genotype (hp-Europe) predominates in native Peruvian strains, approximately 20% of these strains represent a sub-population with an Amerindian ancestry *(hsp-Amerind*). All of these strains however, irrespective of their ancestral affiliation harbored a complete, 'western' type *cag*PAI and the motifs surrounding it. This indicates a possible acquisition of *cag*PAI by the *hsp-Amerind* strains from the European strains, during decades of co-colonization. Our observations therefore suggest presence of ancestral *H. pylori* (*hsp-Amerind*) in Peruvian Amerindians which possibly managed to survive and compete against the Spanish strains that arrived to the New World about 500 years ago. We suggest that this might have happened after native Peruvian *H. pylori* strains acquired *cag*PAI sequences, either by new acquisition in *cag*-negative strains or by recombination in *cag* positive Amerindian strains. In case of Indian strains, almost all the isolates analyzed revealed a European ancestry and belonged to MLST genogroup hp-Europe. The *cag*PAI harbored by Indian strains also revealed European features upon PCR based analysis and whole PAI sequencing. These observations therefore suggest that *H. pylori* in India have ancient origins in Europe. Highly similar MLST and *cag*PAI genotypes observed for ethnically and

linguistically diverse Indian people might argue for a European-Central Asian root of population expansion in the Indian subcontinent. Predominance of genogroup hp-Europe in India amidst non-existence of other genogroups such as hp-Africa and hp-East Asia, point to the fact that the strains of former type carried a special fitness advantage in Indian stomachs, possibly conferred by complete and intact 'western' type *cag*PAIs to outcompete endogenous strains, if any. These results also might potentate speculations related to large-scale replacement of the ancient indigenous people of India by outsiders, bringing first Neolithic practices and languages from the Fertile Crescent and Central Asia.

(2) A genomics approach to understanding host response during dengue infection

Martin L. Hibberd

Genome Institute of Singapore, Group Leader, Infectious Disease, Genome Institute of Singapore, #02-01 Genome Building, 60 Biopolis Street, Singapore 138672, Singapore.

Dengue infection results in a wide clinical spectrum, ranging from asymptomatic, through fever (DF), to the life threatening complications hemorrhagic fever (DHF) and shock syndrome (DSS). Although we now understand that factors such as repeat infections and the type or magnitude of the host response are important in determining severity, the mechanisms of these actions remain largely unknown. Understanding this host–pathogen interaction may enable outcome prediction and new therapy options. Developments in biology now allow a ''systems approach'' to be applied to this problem, utilizing whole genomes of both human and virus, *in vitro* and *in vivo* to enable a more complete picture of their interplay to be built up. We have developed a whole genome approach to viral sequencing, to increase efficiency and enable large numbers of genomes to be completed, together with a web-based interpretation tool. We have also applied human genome expression arrays to characterize the types of host response made to the different viruses and also investigate the role of host variation using human whole genome genetic association studies. These technologies have identified novel host pathways involved in viral replication, and also host immune responses, such as the interferon signaling pathway, that are influenced by viral sequence and thus viral evolution.

(3) Taming of host innate response by a potential biothreat agent *Burkholderia pseudomallei*

Stitaya Sirisinha

Department of Microbiology, Faculty of Science Mahidol of University, Bangkok, Thailand.

Burkholderia pseudomallei is a facultive intracellular Gram-negative bacillus classified by CDC as a category B agent. It not only has potential for bioterrorism, but also causes potentially fatal septicemia in humans and animals. We have previously shown that *B. pseudomallei* is able to survive and replicate in mouse macrophage cell lines, escape into the cell cytoplasm and induce cytoplasmic protrusion leading to direct cell-to-cell spreading. The bacteria induces cell fusion, resulting in multinucleated giant cell formation and apoptotic cell death. The macrophages infected with *B. pseudomallei* exhibit reduced and delayed formation of TNF-α and fail to stimulate iNOS and NO production, thus allowing the bacteria to survive intracellularly. One of the mechanisms responsible for the depressed response is most likely associated with a failure to induce IFN-β production required for phosphorylation of STAT1 and induction of IRF1. The latter is one of the transcription factors needed to turn on the iNOS gene. On the other hand, we can favorably modulate host cell response by using immunomodulating agents, e.g., CpG oligodeoxynucleotide, that can boost up its innate immunity by enhancing iNOS production, increasing uptake and intracellular killing capacity of the macrophages. *B. pseudomallei* may also produce negative regulator that in turn turns off a subsequent host cell response to these stimuli. We recently demonstrated that *B. pseudomallei* could readily induce the expression of negative regulators that interfere with host cell response to interferon-γ stimulation, thus allowing the bacteria to escape killing by the activated phagocytes. Furthermore, we now have additional information from DNA microarray study using Affymetrix chips with human lung epithelial cell line infected with *B. pseudomallei.* There was a down regulation of IL-6, IL-8 and the adhesion molecule ICAM-1 when compared with cells infected with its avirulent counterpart. Altogether, the data strongly indicate that *B. pseudomallei* successfully modulate host innate response for its own survival inside the infected host.

(4) The human EMR1 gene is under strong balancing selection, and may have a role in susceptibility to pulmonary tuberculosis

Jenny Tan, Mark Seielstad

Genome Institute of Singapore, Singapore.

One third of humanity is infected by *Mycobacterium tuberculosis* and more than two million people die from the infection each year. And yet, despite this awful toll, only a tenth of the infected billions will ever succumb to or even exhibit symptoms of the disease. This bespeaks a major role for genetic variability in determining the outcome of Tb exposure and infection. Identifying the relevant genetic variation has stymied investigators for some time, with most progress to date arising from studies of severe Mendelian mycobacterial susceptibilities. In a preliminary analysis of two distinct data sets comprised of (1) ∼10,000 SNPs distributed throughout the human genome and genotyped in 42 active Tuberculosis cases matched with 42 household and community controls and (2) ∼110,000 SNPs genotyped in 111 active cases and 116 controls, we have seen statistically significant associations among numerous SNPs. One of the genes that has been strongly implicated in our study is EMR1. This unusual gene is a member of the EGF-TM7 family of receptors that are predominantly expressed by cells of the immune system. In the course of resequencing this gene to search for putatively functional variants that could be involved in the TB disease process, we observed patterns of genetic variation that were strongly suggestive of natural selection. The

gene displays an elevated level of nucleotide diversity that ranks among the very extremes of the empirical distributions for human genes, a skew in the allele frequency spectrum towards intermediate frequency alleles, positive Tajima's *D* values, an elevated nonsynonymous substitution rate within the human lineage, the presence of highly divergent intermediate haplotypes, and a level of population differentiation (Fst) that is lower than the global average. These data suggest that the EMR1 gene does not evolve in a neutral fashion and is more likely to have experienced strong balancing selection. Because proteins containing EGF-like modules are typically involved in protein–protein interactions and the observation that 14 of the 20 nonsynonymous variations reside within the extracelullar portion of the receptor, the target of selection is probably directed against the EGF-like domain, and may involve recognition of pathogen associated molecular patterns (PAMPs). Biological studies that seek to identify the interacting target(s) of the EMR1 receptor can shed further insight into the nature of the evolution of the gene, and its involvement (if any) in susceptibility to infections.

Symposium "Medical entomology 1"

(5) *Aedes aegypti*: Experimental data supports a genetic background for shape variation

N. Jirakanjanakit[1], S. Leemingsawat[2], J.P. Dujardin[3]

[1]*Center for Vaccine Development, Institute of Science and Technology for Research and Development, Mahidol University, Salaya, Puttamonthon, Nakhonpathom, Thailand.*

[2]*Department of Medical Entomology, Faculty of Tropical Medicine, Mahidol Unversity, Bangkok, Thailand.*

[3]*IRD Montpellier, GEMI, UMR CNRS-IRD 2724, Montpellier, France.*

Shape variation of the wing of *Aedes aegypti* was examined in three isofemale lines A, B and C under controlled laboratory conditions during ten generations. The landmark coordinate data were collected from cross veins and junctions for morphometric analysis. To quantify shape variation, we used the ''metric disparity'' index, known to be independent of sample sizes. Statistical comparisons were performed by non-parametric tests (bootstraps). It was assumed that isofemale lines had been founded by parents having different genotypes, and that no new genotypes appeared during the 10 generations of follow-up. Metric disparity was scored across lines within a given generation and across different generations within a given line. It was shown that the metric disparity index behaved as expected for an indicator of genetic diversity: increasing when mixing different lines, not increasing when adding individuals of the same line. In addition, a simple classification tree of the total sample showed that even after 10 generations, the wings were clustered into three groups according to the initial founders. This study suggests a genetic basis for wing geometry of *A. aegypti*. The epidemiological interest of wing shape behaving as a genetic character would be to help in detecting natural patterns of population structuring at a low cost. In the same way three experimentally isolated lines were recognized by individual wing traits, it is expected that any isolated field population could also be detected. Similar conclusions were obtained previously on various old laboratory lines of *A. aegypti*.

Keywords: *Aedes aegypti*; Isofemale line; Shape variation; Metric disparity

(6) Population structure of main malaria vectors in Asia, members of *Anopheles* species complexes: Implications on vector control

Manguin Sylvie[1], Garros Claire[1,3], Dusfour Isabelle[1,4] Coosemans Marc[2]

[1]*Institute of Research for Development (IRD), Centre of Biology and Management of Populations, Montpellier, France.*

[2]*Institute of Tropical Medicine (ITM), Antwerp, Belgium.*

[3]*University of California, Irvine, College of Health Sciences, Program in Public Health, 3501 Hewitt Hall, Irvine, CA 92697-4050, USA.*

[4]*Uniformed Services University of the Health Sciences (USUHS), Department of PMB, 4301 Jones Bridge Road, Bethesda, MD 20814, USA.*

In Asia, the anopheline biodiversity is very rich and the main malaria vectors belong to complexes in which species are morphologically indistinguishable. Recent advances in systematics and molecular identifications have allowed a clarification of phylogenetic relationships and a simplification of species identification among and within the sibling species or *Anopheles* groups. This is of primordial importance for applying appropriate vector control programs. The presentation of the latest data on the main malaria vectors in southern Asia will highlight the importance of precisely identifying the species, assessing relationships among members of complex, and testing phylogenetic hypotheses involving closely related *Anopheles* species to conduct adequate and efficient vector control strategies. The Minimus Complex is widely distributed on the Asian Continent and is composed of two species considered as malaria vectors in SE Asia. This complex belongs to the Funestus Group which comprises 27 closely related species distributed in Africa and in Asia. Based on molecular and morphological characters and a complete phylogenetic work, a new systematic scheme was recently presented which reflects the evolutionary relationships within species of this group. The Sundaicus Complex is distributed along the coast of Asia and is known as one of the main malaria vector in southern Asia. Recent molecular works on this complex have allowed the recognition of at least three species for which phylogeographic evolutionary scenario will be presented along with the malaria risk linked to specific human activities. Other *Anopheles* complexes with major malaria vectors will be mentioned such as the Dirus, Fluviatilis, and Culicifacies in relation to systematics and malaria transmission.

(7) From population structure to genetically engineered vectors: New ways to control vector-borne diseases?

Olivier Sparagano

School of Agriculture, Food and Rural Development, Newcastle University, Agriculture Building, Newcastle-upon-Tyne NE1 7RU, UK.

Epidemiological studies on ticks and the pathogens they can carry (such as Lyme Disease) are showing some correlations between infection rates and biodiversity highlighting the ''dilution'' effects on potential vectors while other studies comparing sympatric small rodent species demonstrated that rodent species transmitting more pathogens are parasitized by more ectoparasite species. Further studies on host dispersion also showed some impacts on genetic diversity in the ticks with some other comparisons between tick sex, location and genetic flows within these ectoparasite populations. However, other studies highlighted no evidence in gene flows in *Ornithodoros coriaceus* and a far more complex situation with *Ixodes uriae*. The ongoing sequencing of *Ixodes scapularis* (vector of the Lyme Disease spirochaete *Borrelia burgdorferi*, the zoonotic *Babesia microti* and the HGE agent now part of the *Anaplasma phagocytophilum* complex). Furthermore, complementing results in genetic improvement in mosquitoes (genetic markers, sexing, genetic sterilization and fail-safe systems) will also increase performance as it has already been shown in field applications in developing countries. Recent results have greatly improved the fitness of genetically modified insects compared to wild type populations with new approaches such as the post-integration elimination of transposon sequences, stabilising any insertion in genetically modified insects. Encouraging results using the Sterile Insect Technique (SIT) highlighted some metabolism manipulation to avoid the viability of offspring from released parent insect in the wild, if necessary. Recent studies on vector symbionts would also bring a new angle in vector control capabilities. These new potential approaches will improve the levels of control or even in some cases would eradicate vector species and consequently the vector-borne diseases they can transmit. This paper will review the work on different genetic approaches to understand host/pathogen interface in vectors and new genetically modified techniques used to control them.

Keywords: Ticks; Mosquitoes; Biodiversity; Genetic flow; Genetic manipulations

(8) A comparative study of genetic lineages of dengue vectors, *Aedes aegypti* and *Aedes albopictus*, from Thailand and France

Walairut Tuntaprasart[1], Philippe Barbazan[2], Jean Paul Gonzalez[2], Sylvie Manguin[3]

[1]*Faculty of Public Health, Mahidol University, 420/1 Rajvithee, Ratchathewee, Bangkok 10400, Thailand.*

[2]*IRD-Mahidol University, CVVD-UR 178, Project, Emerging Viral Diseases and Information System Research, Faculty of Sciences, Mahidol University, Thailand, 25/25 Phutthamonthon 4 Road, Nakonpathom 73170, Thailand.*

[3]*Institute of Research for Development, Center of Biology and Management of Populations, Campus Baillarguet CS 30016, Montferrier/Lez 34988, France.*

In order to compare the genetic diversity among populations of two main vectors of Dengue in Thailand, *Aedes aegypti* and *A. albopictus*, were collected from several sites including urban, rural and forested areas of Bangkok, Nakhonpathom, KhonKaen, ChiangMai and Kanchanaburi Province. Various methods were used such as ovitrap, landing catch and/or aspirator. In addition, *A. albopictus* from France was also collected and manipulated compared with the Thai populations. All adult mosquitoes were species identified and kept at $-20\,^{\circ}C$ until processed. DNA extraction was carried out using a classical extraction buffer as previously described (Collins, 1987). Molecular characterization and genetic lineages identification were done, among all collected sample populations, by using three genetic markers including the mitochondrial NADH dehydrogenase subunit 5 (ND5) fragment, the nuclear ribosomal DNA second internal transcribed spacer region (ITS2) and, the mitochondrial cytochrome oxidase I (COI). Aligned sequences of *A. aegypti* and *A. albopictus* genes from six localities were compared pairwise. The preliminary result showed marqued differences in nucleotide composition among *Aedes* mosquito populations of Nakhonpathom Province as compared to the others. This study reveals information on divergence of dengue vector from endemic areas and will help to understand vector competence and efficiency in transmitting the virus. Furthermore, it will serve as an informative knowledge on the species dispersal modalities and mean for implementing control strategies.

Keywords: Genetics; *Aedes aegypti*; *Aedes albopictus*; Thailand; France

"Student symposium"

(9) Molecular characterization of Thai *Ehrlichia canis* and *Anaplasma platys* strains isolated from dogs

Danai Pinyoowong[1], Sathaporn Jittapalapong[2], Fanan Suksawat[3], Roger W. Stich[4], Arinthip Thamchaipenet[1]

[1]*Department of Genetics, Faculty of Science, Kasetsart University, Thailand.*

[2]*Department of Parasitology, Faculty of Veterinary Medicine, Kasetsart University, Thailand.*

[3]*Department of Medicine, Faculty of Veterinary Medicine, Khon Kaen University, Thailand.*

[4]*Department of Veterinary Pathobiology, University of Missouri, Columbia, MO, USA.*

Canine monocytic ehrlichiosis caused by *Ehrlichia canis* is veterinary importance worldwide. In Thailand, there has been little information available on *E. canis* and its phylogeny. The objective of this study was to characterize and establish molecular structure and phylogeny of Thai *Ehrlichia* and *Anaplasma* strains. *Ehrlichia*-positive blood samples of dogs were extracted for genomic DNAs. 1.5 kb PCR products of 16S rRNA gene were obtained using designed genus-specific primers for *Ehrlichia* and *Anaplasma*. Nearly complete sequences of the 16S rRNA gene were compared with other sequences available in the Genbank database. Percentage of similarity as well as secondary structure analysis of 16S rRNA sequences indicated that they are new *E.* canis and *A. platys* strains. Phylogenetic analysis revealed that two strains of Thai *E. canis* were closely related and formed a single cluster within the cluster amongst previously published *E. canis* from different countries. *A. platys* found in this study showed close relation-

ship with earlier report of *A. platys* in Thailand. This report represents the first molecular characterization of *E. canis* in dogs from Thailand.

Keywords: *Ehrlichia canis*; *Anaplasma platys*; PCR; Thailand; Dogs

(10) *Leishmania braziliensis*: Population structure and reproductive modes

V. Rougeron[1], T. De Meeûs[1], M. Hide[1], E. Waleckx[1], J. Arevalo[2], H. Bermudez[3], A. Llanos-Cuentas[2], J.-C. Dujardin[5], A.-L. Bañuls[1]

[1]*Laboratoire GEMI UMR CNRS-IRD 2724, Centre IRD, Montpellier, France.*

[2]*Alexander Von Humboldt Institute of Tropical Medicine, Universidad Peruana Cayetano Heredia, Lima, Peru.*

[3]*Faculty of Medicine, Universidad Mayor San Simon, P.O. Box 4866, Cochabamba, Bolivia.*

[4]*Department Parasitology, Institute of Tropical Medicine, Antwerp, Belgium.*

Leishmaniases are severe diseases affecting humans and animals caused by protozoan parasites belonging to the *Leishmania* genus and transmitted by female sandflies' bites. These parasitoses are widespread over all continents, except Antarctica. Nowadays, leishmaniases still pose considerable public health problems. At present, it is suggested and even admitted that *Leishmania* species present a basically clonal population structure associated to rare sexual recombination events. However, such a statement mostly relies on population genetics studies that may be criticised. The markers used were little adapted (lack of resolution or dominant markers) and clonality was inferred from the analysis of linkage disequilibria across loci that are far from ideal in that respect. *Leishmania braziliensis* is an important leishmaniasis agent in South America. The principal objective of our work was to study the population structure and reproductive mode of this species in Peru and Bolivia and, for the first time, using microsatellite markers. On the whole, 124 human isolates (68 from Peru and 56 from Bolivia) were genotyped on 12 microsatellite loci. Various population genetics tests were applied. The results obtained appear in contradiction with a simple clonal propagation. Indeed, strong homozygosities found at each locus, associated to strong linkage disequilibria across loci, advocated for an inbred reproductive strategy. Further analyses suggest that a significant part of the high heterozygote deficits observed in our samples is likely the consequence of a Wahlund effect, i.e. the coexistence of strongly differentiated genetic entities within each sample. This work brings key information concerning the biology of these organisms and opens new prospects on the study of this species and other members of the genus.

(11) Clonal strains of *Pseudomonas aeruginosa* isolated from patients with cystic fibrosis

Pholawat Tingpej[1], Mark Elkins[2], Veronica Yozghatlian[1], Barbara Rose[1], Jim Manos[1], Peter Bye[2,3], Colin Harbour[1]

[1]*Department of Infectious Diseases and Immunology, Blackburn Building DO6, University of Sydney, 2006 NSW, Australia.*

[2]*Department of Respiratory Medicine, Royal Prince Alfred Hospital, Sydney, Australia.*

[3]*Medicine, University of Sydney, Sydney, Australia.*

Cystic fibrosis (CF) is the most common genetic condition among Caucasians. Eighty percent of people with CF are infected with *Pseudomonas aeruginosa* by adulthood and most die from complications arising from chronic lung infections. *P. aeruginosa* is widespread in various environments including hospitals. It was once generally accepted that individual CF patients acquired *P. aeruginosa* from their environments and thus each patient carried their own unrelated (or unique) strain. Recently however, clonal (or epidemic) strains have been reported in Europe and Australia. In this study, DNA restriction fragment length polymorphisms (RFLP) of *P. aeruginosa* isolates from 112 patients attending an adult CF clinic at Royal Prince Alfred Hospital in Sydney, Australia were analysed using pulsed-field gel electrophoresis (PFGE). Cluster analyses were performed using computer-aided software (GelCompar II). Isolates sharing a similarity greater than 85% defined by the Dice coefficient with a position tolerance of 1.0% and an optimisation of 1.0% were considered as a clone. The DNA fingerprint of each isolate from the same clone had up to a three-band difference, which confirmed their close relatedness. Two major clones, AES-1 and AES-2, were isolated from 38% and 5% of 112 patients, respectively. There is a 66% similarity between AES-1 and AES-2 clones. The AES-1 strain had an identical DNA banding pattern with a previously reported Melbourne epidemic strain. The AES-2 strain was also identical to an epidemic strain reported from Brisbane known as a Pulsotype II. AES-1 isolates were significantly more resistant to gentamycin, amikacin and Timentin® than non-clonal isolates reflecting treatment difficulties. These strains have not been detected from the environment suggesting that person-to-person transmission may play a role in such cases. These results have led us to implement a segregation policy in our clinic as well as emphasising the important role of molecular typing in infection control.

(12) Influenza: An idea model bridging epidemiological and evolutionary dynamics

Zhenggang Wang, Chung-Chau Hon, Tsan-Yuk Lam, Fanya Zeng, Frederick Chi-Ching Leung

Annual outbreaks of influenza cause substantial morbidity and mortality, and also cause heavy economic losses. In recent years the threat of a human influenza pandemic has increased considerably as humans have become susceptible to infection by the avian influenza virus H5N1. However, our current understanding of influenza and the ability to evaluate the threat are limited. Several important issues, including the influence of climate variability on influenza epidemic patterns and intraspecific and interspecific interactions between various circulating influenza types, subtypes and strains, have not yet been sufficiently studied. In this study, we explored the immunological dynamics and epidemiological dynamics of influenza using our host immune unit-virus-susceptible (HVS) model. By matching model output to epidemiological patterns identi-

fied in surveillance data collected from United States, we found that three types of interspecific competitions (between influenza A and B, influenza A subtypes H1 and H3, and new and circulating strains) are essential to depict phylogenetic patterns of influenza. The study therefore, illustrates the population dynamics of the emergence, circulation and elimination of new influenza variants (subtypes or strains).

Keywords: Influenza; Co-circulation; Predator–prey model; Epidemiological dynamics; Evolutionary dynamics

Symposium "Phylogeny of pathogens"

(13) The evolution of gene overlap in RNA viruses

Robert Belshaw[1], Oliver G. Pybus[1], Andrew Rambaut[2]

[1]*Department of Zoology, University of Oxford, United Kingdom.*

[2]*Institute of Evolutionary Biology, University of Edinburgh, United Kingdom.*

Two of the most striking attributes of RNA viruses are their small genome size (the average RNA virus is only 9 kb long) and their high mutation rate. It has been argued that these attributes are linked, with genome size being limited by the accumulation of deleterious mutations in a long stretch of RNA. Genome size is also linked to mutational robustness, which can be thought of as the local gradient of the adaptive landscape around its peak. Sharper peaks represent less robust genomes where mutations have a proportionately greater negative effect on fitness. Viruses with larger genomes should evolve to be more mutationally robust because on average they can expect to experience more mutations per round of replication. We investigate the relative mutational robustness of RNA viruses by measuring the amount of gene overlap. Most RNA viruses have some nucleotides that code for more than one protein by being in two overlapping reading frames. In such viruses, some mutations will therefore affect more than one gene and hence will have an increased negative effect on viral fitness. We analysed the sequences of 700 viral species, correcting for phylogenetic non-independence, and found that – as predicted – gene overlap is strongly negatively correlated to genome length. Furthermore, in the relative abundance of different frameshifts, we find evidence for two evolutionary processes having been at work: new genes being created in other frames within older genes, and creeping overlap between originally contiguous genes that happen to be in different frames. We propose two simple evolutionary models to explain these processes.

(14) Phylogenetic and antigenic analysis of *Orientia tsutsugamushi* isolated from scrub typhus patients in Thailand

Stuart D. Blacksell[1,2], Rungnapa Luksameetanasan[1,3], Thareerat Kalambaheti[3], Vanaporn Wutheikanun[1], Wirongrong Cheirakul[1], Daniel H. Paris[1], Sunee Chueasuwanchai[1], Rose McGready[1], Sharon J. Peacock[1,2], Nicholas P.J. Day[1,2]

[1]*Wellcome Trust, Mahidol University, Oxford Tropical Medicine Programme, Faculty of Tropical Medicine, Mahidol University, Bangkok 10400, Thailand.*

[2]*Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Oxford, United Kingdom.*

[3]*Department of Microbiology and Immunology, Faculty of Tropical Medicine, Mahidol University, Bangkok 10400, Thailand.*

Orientia tsutsugamushi is the causative agent of scrub typhus infection and is a major cause of human disease in rural areas of South East Asia. Seventeen in vitro isolates of *O. tsutsugamushi* from patients with scrub typhus disease in north-eastern and western Thailand between 2003 and 2005 were examined to determine phylo-temporal and phylo-geographic relationships between the samples. Implications for diagnosis were also investigated. Phylogenetic analysis of the entire 56 kDa-type-specific antigen gene (~1.6 kb) found that the majority (11/17; 64.7%) of the isolates clustered within Karp genotype, and 4 of 17 (23.5%) isolates within the Gilliam/Kawasaki cluster. Two isolates grouped with each of the historical Thai isolates TA763 (Karp-like) and TA716 (Kato-like). Two-dimensional cross-binding of patient antibody reactivity against *O. tsutsugamushi* isolate antigens demonstrated relationships similar to 56 kDa gene nucleotide sequence results with distinct differences between the binding of Gilliam/Kawasaki antibodies and Karp antigens. Results from 56 kDa genetic analysis demonstrates a Karp type strain dominance similar that reported in studies from 1960s and 1970s. There were no clear geographical associations from this study however more isolates are required to confirm this observation. The majority of scrub typhus vaccine candidates are based on the 56 kDa protein of Karp type strain and the results presented here demonstrate that Karp type strain should be a major component of a future vaccine however it is unclear what is the efficacy of such a vaccine with other type strains?

(15) Sequence analysis of the C-terminal region of merozoite surface protein-1 of *Plasmodium falciparum* (PfMSP-119) and *P. vivax* (PvMSP-119) as vaccine candidate antigens among Iranian clinical isolates

A.A. Mehrizi[1,2], N.D. Djadid[1], A.H. Salmanian[2], S. Zakeri[1]

[1]*Malaria Research Group (MRG), Biotechnology Department, Pasteur Institute of Iran, Pasteur Avenue, P.O.Box 1316943551, Tehran, Iran.*

[2]*National Research Center for Genetic Engineering and Biotechnology, Tehran, Iran.*

In this study we analyzed the sequence variation of the C-terminal region of merozoite surface protein-1 of *Plasmodium falciparum* (Pfmsp-119) and *P. vivax* (Pvmsp-119) genes as the most promising blood stage vaccine target antigens, in 70 *P. vivax* and 92 *P. falciparum* infected blood samples collected from areas with different malaria endemicity in Iran. The presence of polymorphism in this region may compromise its use as a vaccine candidate. All *P. vivax* samples have shown 100% conserved sequences among northern and southern isolates, however the MAD20 allele was found significantly among *P. falciparum* clinical isolates in south. Furthermore, MAD20 allelic type showed four different allelic forms, while the K1 allelic type

showed no polymorphism. These results are discussed with regards to evaluation of these vaccine antigens in both malaria species, and in compare with the studies that were conducted in other areas in Southeast Asia, and Africa. Such study would complement this information and would allow comparing the Iranian *P. falciparum* and *P. vivax* populations with those found in distinct and contrasting epidemiological settings.

(16) Sequence variation in the T-cell epitopes of the *Plasmodium falciparum* circumsporozoite protein in Iranian clinical isolates

S. Zakeri[1], M. Avazalipour[1], A.A. Mehrizi[1], G. Snounou[2], N.D. Djadid[1]

[1]*Malaria Research Group (MRG), Biotechnology Department, Pasteur Institute of Iran, Iran.*

[2]*Parasitologie Comparee et Modeles Experimentaux, Museum National d'Histoire Naturelle, Paris, France.*

The first subunit malaria vaccine tested in humans was based on the *Plasmodium falciparum* circumsporozoite protein (PfCSP). The T-cell helper epitopes Th2R and Th3R found in the carboxyl-terminus region included in some vaccine formulations, showed sequence diversity which could be a potential problem for vaccine efficacy. The aim of this study was to define the nature and extend of *Pfcsp* genetic polymorphism in isolates collected from patients in Iran. The data would complement information obtained in other endemic settings. A total of 21 complete and 69 partial *Pfcsp* sequences were derived from isolates collected in the south-eastern hypoendemic area of Iran. Although nine different allelic forms were observed in the 21 complete sequences, they were mainly due to variation in the repeat units number and arrangement, whereas only two haplotypes were noted for the combined Th2R /Th3R epitopes, for each of which only two allelic variants were noted. Comparison of the 3′-end region of all 90 sequences revealed only one more Th2R variant, and a total of five combined haplotypes of which three were dominant, and two only found in a minority of samples collected from non-Iranian patients who acquired the infection abroad. Thus, the *Pfcsp* gene of the parasites circulating in Iran displays a very low level of diversity. These results contrast with observations made in Africa, but are akin to those observed in other regions (Papua New Guinea, Thailand and Brazil).

Symposium "Coevolution host pathogen 2"

(17) The analysis of candidate genes and their influence on tuberculosis susceptibility in a Canadian Aboriginal population

Linda Larcombe[1], Peter Nickerson[2], Iga Dembinski[3], Andrew Lodge[2], Erin Larcombe[4], Leisel Milligan[4], Bruce Martin, Pam Orr[1,2]

[1]*Department of Medical Microbiology, University of Manitoba (Presenter), Room 512, Basic Medical Sciences Building, 730 William Avenue Winnipeg, MB, Canada R3E 0W3.*

[2]*Faculty of Medicine, University of Manitoba, Canada.*

[3]*Canadian Blood Services, 777 William Avenue Winnipeg, Manitoba, Canada.*

[4]*Faculty of Science, University of Manitoba, Canada.*

Despite the ability to treat tuberculosis and the continued decline of mortality rates, tuberculosis continues to have a detrimental impact on the quality of life for many Aboriginal people in northern Canada where morbidity rates remain high. In a small isolated community, the rate of tuberculosis is 60 times higher than that of the average Canadian Aboriginal population. Previous research demonstrated that significant differences exist in the frequency of cytokine gene polymorphisms maintained by distinct Aboriginal and Caucasian populations in Manitoba. The Dené are a discrete Aboriginal cultural group and recent analysis has shown that this group maintains a high frequency of cytokine gene polymorphisms (TNFa, IL-6, IFNg, IL-10, TGFb) related to an effective Th2 immune response but a less effective Th1 response to infectious diseases. In addition, the Dené have a high frequency of gene polymorphisms in the Vitamin D Receptor gene which may in part, contribute to their susceptibility to tuberculosis. This presentation will describe the analysis of a panel of purported tuberculosis-susceptibility genes (Vitamin D receptor and cytokine SNPs) from a northern Canadian Dené cohort. The Dené have a unique history and prehistory in relation to other northern Canadian Aboriginal populations and as a result they have preserved their cultural identity and along with that, their distinct immunogenetic profile that is well adapted to a specific pathogen environment.

(18) Malaria *Plasmodium* agent induces alteration in the head proteome of their *Anopheles* mosquito host

Thierry Lefèvre[1], Frédéric Thomas[1], Alex Schwartz[2], Elena Levashina[2], Stéphanie Blandin[2], Jean-Paul Brizard[3], Laure Le Bourligu[4], Edith Demettre[4], François Renaud[1], David G. Biron[1]

[1]*GEMI, UMR CNRS-IRD 2724, IRD, 911 Av. Agropolis BP 64501, 34394 Montpellier Cedex 5, France.*

[2]*UPR 9022 CNRS, Avenir Group INSERM, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, 67084 Strasbourg Cedex, France.*

[3]*UMR 5096 (UP-IRD-CNRS), IRD, 911 Av. Agropolis BP 64501, 34032 Montpellier Cedex 5, France.*

[4]*LGF 2580 CNRS, 141 rue de la Cardonille, 34094 Montpellier Cedex 5, France.*

Host behavioural changes induced by parasites that increase the likelihood of parasite transmission has long captured the interest of parasitologists and behavioural ecologists. For instance, in pathogens—insect vector systems, several studies support the idea that pathogen manipulates the behaviour of their vectors, such as feeding behaviour, in a way that increase the contact with the vertebrate host and hence favour the parasites' transmission. Despite increasing evidence of such behavioural changes, the underlying mechanisms causing infected vectors to act in ways that benefit pathogen transmission remain enigmatic in most cases. Here, 2D difference gel electrophoresis coupled with mass spectrometry were employed to analyse and compare the head proteome between malaria (*Plasmodium berghei*) infected mosquitoes and uninfected mosquitoes (*Anopheles gambiae*). This proteomics

approach detected 12 protein spots in two cohorts of mosquitoes with altered levels in the head of sporozoite infected individuals. These proteins were subsequently identified using mass spectrometry and functionally classified as metabolic, synaptic, molecular chaperone, signalling, and cytoskeletal proteins. Our results indicate an altered energy metabolism in the head of sporozoite infected mosquitoes. Some of the up/down regulated proteins identified such as synapse associated protein, 14-3-3 protein, and calmodulin have previously been shown to play critical roles in the central nervous system of invertebrates and vertebrates. Furthermore, a heat shock response (HSP 20) and a variation of cytoarchitecture (tropomyosins) have been evidenced. These proteins shed light on potential molecular mechanisms underlying behavioural modifications and offer new insights into the study of intimate interactions between *Plasmodium* and its *Anopheles* vector.

Keywords: Malaria; Mosquitoes; Host–parasite systems; 2D difference gel electrophoresis; Mass spectrometry

(19) The ORF2 glycoprotein of hepatitis E virus retrotranslocate from the endoplasmic reticulum to the cytoplasm and down-regulates NF-κB activity in human hepatoma cells

Milan Surjit, Sunil K. Lal

Virology Group, International Centre for Genetic Engineering & Biotechnology, New Delhi 110067, India.

NF-κB is a key transcription factor that has been implicated to play a crucial role in host survival during infection by pathogens. Therefore, it has been a priority of many pathogens to manipulate the cellular NF-κB activity in order to create a favorable environment for their survival inside the host. In this report, we provide evidence for a novel mechanism of inhibition of NF-κB activity, which is mediated by the major capsid (ORF2) protein of the hepatitis E virus. Heterologous expression of the ORF2 protein in human hepatoma cells was found to inhibit IκBα ubiquitination by interfering with the assembly of the $SCF^{\beta TRCP}$ complex, thus resulting in stabilization of the cellular IκBα pool, with a concomitant reduction in the activity of NF-κB and its downstream targets. NF-κB inhibitory activity exhibited by the ORF2 protein was found to depend on its ability to retrotranslocate from the endoplasmic reticulum (ER) to the cytoplasm, where it was observed to be stably present. Further, retrotranslocation of the ORF2 protein was dependent upon the glycosylation status of the protein, mediated in a p97 dependent pathway and independent of ubiquitination of the former. The ORF2 protein, therefore, exploits the ER associated degradation pathway to gain access to the cytoplasm, where it interferes with the IκBα ubiquitination machinery, leading to inhibition of host cell NF-κB activity.

Keywords: Retro-translocation; NF-κB; ER stress; ERAD pathway; IκBα ubiquitination; ORF2 protein of hepatitis E virus

(20) Development of a novel immunome-based *Candida* vaccine

Songsak Tongchusak[1], Chanvit Leelayuwat[2], Vladimir Brusic[3], Sansanee C. Chaiyaroj[1]

[1]*Department of Microbiology, Faculty of Science, Mahidol University, Bangkok, Thailand.*

[2]*Department of Clinical Immunology, Faculty of Associated Medical Sciences, Khon Kaen University, Khon Kean, Thailand.*

[3]*Cancer Vaccine Center, Dana-Farber Cancer Institute, Harvard Medical School, Boston, USA.*

Candida albicans is one of the most important opportunistic dimorphic fungi responsible for hospital acquired fungal infection in humans. Generally, candidiasis rarely occurs in healthy individuals but it is frequently associated with patients who receive immunosuppressive drug therapy or long-term catheterization and patients who suffer from AIDS (Navarro-Garcia et al., 2001). To date, there are neither effective vaccines nor therapeutic protocols to eradicate these fungal infections. We therefore utilized immunomics approach to assist in the identification of the fungal immunoprotective epitopes. First, a molecular database of *C. albicans* virulence factors called CandiVF was developed (URL http://antigen.i2r.a-star.edu.sg/Templar/DB/-CandiVF/). The database contained 153 virulence proteins of *C. albicans.* It also provided a T-cell epitope predictive algorithm called *Hotspot Hunter* which was previously developed by our group. *Hotspot Hunter* facilitated the prediction of promiscuous peptides that can bind specifically to eight supertypes of HLA-DRB1 (DRB1*0101, *0301, *0401, *0701, *0801, *1101, *1301, *1501). In order to verify T-cell epitope prediction by *Hotspot Hunter*, secretory aspartyl proteinase 2 (Sap2) was selected as a study model. Sap2 was a *C. albicans* common antigen during infection and capable of inducing IgE-mediated allergic reaction in atopic individuals (Suenobu et al., 2002). Primary sequence analysis revealed that Sap2 contained two different sequences, groups 1 and 2. A total 40 conventional overlapping peptides of Sap2 (20-mer overlapped 10-mer in length) were then synthesized. All peptides were used to stimulate peripheral blood mononuclear cells form HLA-DRB1 specific blood donors to determine the proliferative response. Eleven of eighteen peptides within the prediction areas were able to induce PBMCs proliferation. However, when anti-IL-2 ELISpot assay was used to confirm the cell proliferation result, only two of eleven peptides stimulated significant T-cell activation. Outside the predictive areas, peptide 11 could induce proliferation of IL-2 producing clone in one donor of HLA-DRB1*04/04. The use of immunomics can assist the identification of immunoprotective eptiope and the development of a potential peptide-based vaccine.

Keywords: Immunomome; Immunomics; T-cell epitope; Database and *Candida albicans*

References

Navarro-Garcia, F., Sanchez, M., Nombela, C., Pla, J., 2001. Virulence genes in the pathogenic yeast *Candida albicans*. FEMS Microbiol. Rev. 25 (2), 245–268.

Suenobu, N., Kweon, M.N., Kiyono, H., 2002. Nasal vaccination induces the ability to eliminate *Candida* colonization without influencing the pre-existing antigen-specific IgE Abs: a possibility for the control of *Candida*-related atopic dermatitis. Vaccine 20, 2972–2980.

Symposium "A revolutionary way to detect pathogens: Epidemiological and clinical Importance of ultra-sensitive detection of endemic and emerging diseases"

(21) New revolutionary ultrasensitive technique to detect pathogens based on their capture and concentration by ApoH coated nano-magnetic beads

Ilias Stefas[1], Dorothée Missé[2], Johanna Gajardo[2], Alvaro Martinez[2], Priscille Bosc[1], Francisco Veas[2]

[1]*Apo-H Technologies, Montpellier, France.*

[2]*Laboratoire Immunologie Virale et Moléculaire, Viral Emerging Diseases Unit UR 178, Institut de Recherches pour le Développement, Montpellier, France.*

Despite remarkable advances in medical research and treatments during the 20th century, infectious diseases remain among the leading causes of death worldwide. The main causes for this phenomenon are: (1) emergence of new infectious pathogens; (2) re-emergence of old infectious diseases have reappeared after a significant decline in incidence; and (3) persistence of intractable infectious pathogens. Indeed, the previous decades have been marked by several striking episodes of emerging and re-emerging pathogens such as HIV, Marburg virus, Hepatitis C virus, Hantavirus, Ebola virus, West-Nile virus, Dengue virus, Yelow Fever virus and more recently SARS coronavirus and avian flu. New infectious diseases continue to evolve and emerge. According to the Center for Disease Control and Prevention in Atlanta 70% of emerging infectious diseases in humans are zoonotic pathogens. In order ta anticipate the measures to be taken to prevent or control future epidemics, considerable attention has recently been directed to emerging and re-emerging infections at national and international level (http://euro.who.int/surveillance).

To adopt the appropriate containment measures towards emergent pathogens, fast, sensitive and reliable diagnostics are key element. Nucleic acid amplification is widely used for the detection and identification of pathogens. One of the main problems for pathogens detection in clinical but also in environmental samples is that they generate false negative results. This problem is mainly due to three reasons: presence of inhibitors, absence of a universal extraction method, lack of a rapid and reliable pathogen concentration methodology. The above-mentioned disadvantages would be compensated by the use of a very sensitive method consisting of a matrix-bound APOH which has a particular property to fix a broad panel of pathogens. Interestingly, APOH strongly interacts with various viruses such as, HBV, HCV, orthopoxviruses, Dengue virus, Hantavirus, H5N1, West Nile which are either endemic or emerging diseases in South-East Asia.

(22) Capture and concentration of orthopox virus using ApoH coated nanomagnetic beads for ultrasensitive detection

Nancy Ernst[1,2], Jana Rödig[1], Sonja Linke[1], Elias Stefas[2], Francisco Veas[3], Heinz Ellerbrok[1]

[1]*Robert Koch-Institut, Centre for Biological Safety, Berlin, Germany.*

[2]*Apo-H Technologies, Montpellier, France.*

[3]*Laboratoire Immunologie Virale et Moléculaire, Viral Emerging Diseases Unit UR 178, Institut de Recherches pour le Développement, Montpellier, France.*

Major obstacles for the detection of pathogens in clinical or environmental samples are false negative results. This is mainly due to the lack of a rapid and reliable pathogen concentration methodology that allows detection of highly diluted samples, and the inability of most of the currently used technologies to eliminate or neutralize interfering ''natural inhibitors'' that could be present in biological samples. In order to improve virus diagnostics we wanted to exploit the ''non-self'' recognition and binding properties of human apolipoprotein H (ApoH. ApoH binds and captures pathogens enabling their concentration from different kinds of biological samples. We have used magnetic beads coated with ApoH recombinant protein as a pre-treatment step for orthopox viruses to improve the detection threshold and to increase the sensitivity for diagnosis. With this approach virus was concentrated, DNA was extracted and subsequently detected and quantified by real-time PCR. After ApoH-treatment, Vaccinia Virus was detected from highly diluted samples where diagnosis had been negative with a standard DNA preparation protocol. At present the concentration and improved detection of other viruses with an ApoH-enhanced protocol is under investigation.

(23) ApOH-capture technology enhances Andes hantavirus detection allowing virus concentration from plasma and urine samples of patients with acute hantavirus cardiopulmonary syndrome

Paula Godoy[1], Elias Stefas[2], Pablo Ferrer[1], Valeska Vollrath[1], Pablo Vial[3], Marcela Ferrés[1], Francisco Veas[4], Marcelo López-Lastra[1]

[1]*Laboratorio de Infectología y Virología Molecular, Centro de Investigaciones Médicas, Pontificia Universidad Católica de Chile.*

[2]*ApoH Technologies, Montpellier, France.*

[3]*Facultad de Ciencias de la Salud, Universidad del Desarrollo, Santiago, Chile.*

[4]*Viral and Molecular Immunology Laboratory, U178/IRD, Montpellier, France.*

Background: Hantavirus cardiopulmonary syndrome (HCPS) is an emerging disease caused by new world hantaviruses. Hantaviruses (HV) are segmented RNA viruses belonging to the genus *Hantavirus* in the family *Bunyaviridae*. Hantavirus (HV) infections are mainly transmitted to humans by inhalation of virus-contaminated aerosols of rodent excreta and secretions, however, sporadic person-to-person transmissions of the Andes hantavirus (ANDV) have been reported. Based on the knowledge that human apolipoprotein H (ApoH), a constituent of human plasma, interacts with viral proteins, we wished to asses a possible interaction between ApoH and ANDV, the major etiological agent HCPS in South America.

Materials and methods: Blood and urine samples from acute-HCPS patients were selected on the basis of their availability. Samples collected as part of the research initiative NIH/NIAID #AI 45452 were kindly supplied for this study. Donor patients met the clinical criteria for HCPS and harbored IgM antibodies reactive with hantavirus antigens. HV genomic RNA was confirmed in plasma by an in-house developed RT-PCR/hemi-nested PCR, using primers designed to partially encompass the S segment ORF of the Andes virus, strain CHI-7913. Samples used as negative control were collected among the laboratory staff. ApoH-coated magnetic beads and ApoH-coated ELISA plates used in this study were supplied by ApoH Technologies S.A. and used following their instructions.

Results: We report that ANDV interacts with ApoH, and that ApoH-coated magnetic beads or ApoH-coated ELISA plates can be used to capture and concentrate virus from serum and urine samples, allowing virus detection by both immunological and molecular approaches. We then developed an ANDV-high throughput screen assay and assessed ANDV in urine samples, from 50 patients with acute ANDV-HCPS, collected during 5 days following hospitalization. 45 patients showed detectable amounts of ANDV in urine in at least one tested sample.

Conclusions: ApoH capture assay increases the sensitivity of virus detection by both molecular and immunological methods. This apparent enhancement in sensitivity most probably stems from the fact that virus is being concentrated from a larger sample volume. Additionally, we demonstrate that ANDV can be shed in the urine of infected individuals. Although, our data do not necessarily predict the presence of infectious virus in urine, the fact that ANDV is readily detected in urine samples of acute-HCPS patients not only lends support to the possibility that urine is a route for person-to-person transmission of HCPS but also raises the intriguing prospect that virus might be present in other biological secretions.

Acknowledgments

This study was supported by the Jeunes Equipes Associées à l'IRD program to MLL, and by US PHS grant number AI 45452 to MF. MLL is member of the Núcleo Milenio en Inmunología e Inmunoterapia.

(24) Elsevier reviewer workshop

Michel Tibayrenc, Bas Straub

Symposium "Epidemiology and genomics of HIV"

(25) Host genetic polymorphisms, HIV variability resistance to infection and disease progression

Nicole Ngo Giang-Huong[1,2,3], Gonzague Jourdain[1,2,3], Tim Cressey[1,2,3], Marc Lallemant[1,2,3]

[1]*U174 "Epidémiologie Clinique, Santé Maternelle et Infantile, and HIV in South East Asia", Institut de Recherche pour le Développement, France.*

[2]*Department of Immunology and Infectious Disease, Harvard School of Public Health, USA.*

[3]*Faculty of Associated Medical Sciences, Chiang Mai University, Thailand.*

Research on host and viral factors associated with susceptibility to HIV-1 infection, disease progression, and response to antiretroviral therapy has proved critical for designing efficient interventions. HIV entry into cells results from complex interactions between env gp120 with CD4 and co-receptors. CCR5, the major receptor for cytokines, is also the main co-receptor for macrophage-tropic R5 strains. The CCR5-delta32 homozygous deletion and the CCR5-m303 variant, have been associated with resistance to HIV infection, and CCR5-delta32 heterozygosity with delayed disease progression. Also, the minor co-receptor CCR2 64I variant has been associated with delayed progression. The effect on progression of SDF1 polymorphisms, a ligand for the chemokine receptor CXCR4, remains controversial. With regards to innate or specific host immune response to HIV, while there are conflicting results on the role of neutralizing antibodies, studies have shown an association between HLA class I alleles and natural resistance to infection or disease progression. Discordance between maternal and infant HLA genotypes may have a protective effect. Response to antiretroviral drugs involves both host genetics and HIV variation. While the ability of HIV to mutate and escape drug pressure varies with each drug's specific mechanism of action, human polymorphisms have been associated with increased toxicity of some antiretroviral drugs. Until today, the challenge has been to demonstrate the clinical significance of identified polymorphism following a pathogenesis hypothesis driven approach (gene/pathway candidates). Following completion of the human genome sequencing, numerous SNPs have been identified whose biological significance remains unknown, and the challenge is to discover associations between such SNPs and characterized phenotypes. Understanding a polymorphism functional significance becomes the next step. Discordant couples, transmitting mothers/partners, rapid progressors under therapy or long term untreated non-progressors are phenotypes which have been widely used. In the rapidly evolving field of HIV medicine, defining in large/diverse populations stable/unambiguous phenotypes is increasingly difficult.

(26) Synonymous substitution rates predict HIV disease progression as a result of underlying replication dynamics

Philippe Lemey[1], Sergei L. Kosakovsky Pond[2], Alexei J. Drummond[3], Oliver G. Pybus[1], Beth Shapiro[1], Helena Barroso[4], Nuno Taveira[4], Andrew Rambaut[1]

[1]*Department of Zoology, University of Oxford, Oxford, UK.*

[2]*Department of Pathology, University of California San Diego, La Jolla, USA.*

[3]*Department of Computer Science, University of Auckland, Auckland, New Zealand.*

[4]*Centro de Patogénese Molecular, Faculdade de Farmácia de Lisboa, and Instituto Superior de Ciências da Saúde Egas Moniz, Lisbon, Portugal.*

To investigate HIV evolution in relationship to disease progression, we developed a new computational technique that can estimate changes in the absolute rates of synonymous and nonsynonymous divergence through time from molecular sequences. This allows separating changes in generation time

and mutation rate from changes in selective pressure and effective population size. Using this technique, we have identified a previously unknown association between the 'silent' evolutionary rate of HIV and the rate of disease progression in infected individuals. This finding demonstrates that cellular immune processes, which are already known to determine HIV pathogenesis, also determine viral replication rates and therefore impose important constraints on HIV evolution. Humoral immune responses, on the other hand, are the major determinant of nonsynonymous rate changes through time in the envelope gene, and our relaxed clock estimates support a decrease in selective pressure as a consequence of immune system collapse.

(27) The presence of anti-R7V antibodies in HIV-1 infected patients: A novel efficient marker for the non-progression to AIDS

Chonlaphat Sukasem[1,2], Chutatip Srichunrusami[2], Jiraporn Keatkla[2], Montip Tiensuwan[3], Wasun Chantratita[2]

[1]*Department of Pathobiology, Faculty of Science, Mahidol University, Bangkok, Thailand.*

[2]*Department of Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Bangkok, Thailand.*

[3]*Department of Mathematics, Faculty of Science, Mahidol University, Bangkok, Thailand.*

Background: Concerning HIV life cycles, it is well known that the HIV acquires a cellular antigen when the virus is released from the host infected cell. The epitope is defined as a peptide, seven amino acids (RTPKIQV; R7V). The previously study reported that the presence of antibodies directed to the peptide R7V in HIV infection correlated with the non-progression to AIDS.

Objectives: T5o determine the correlation between the anti R7V antibodies in Thai treated-naive HIV-1 infected patients and the clinical status.

Methods: A retrospective study was carried out in 124 normal population and 128 Thai treated-naive patients, infected with HIV-1 for more than 5 years. Presence of anti R7V antibody was detected by the anti R7V ELASA. OD was read at 450 nM.

Results: We found that treated-naive HIV-infected patients presented anti-R7V antibodies at higher level than in normal population (32.8% and 0.8%, respectively; $P < 0.001$). Relative to clinical status of patients, the frequency of positive anti R7V antibodies level had significantly higher in non-progressors and moderate progressors (100% and 49%, respectively; $P < 0.001$). We did not find anti R7V antibodies in rapid progressors. These results demonstrated a strong correlation between the presence of anti-R7V antibodies and a good prognostic status of Thai HIV-1 infected patients.

Conclusion: This study provides the strongest evidence to date for the presence of the anti R7V antibodies in non-progression of HIV infection.

Keywords: Progression markers; Anti-R7V antibody; Non-progression; HIV-1 infected patients; ELASA; AIDS

(28) Origin of HIV-1

Fran Van Heuverswyn

UMR145, Institute for Research and Development (IRD) and University of Montpellier, Montpellier, France.

West central African chimpanzees (*Pan troglodytes troglodytes*) are now recognized as a natural reservoir of simian immunodeficiency virus (SIVcpz) and the immediate source of at least two human cross-species infections: (i) HIV-1 group M, the pandemic form of human immunodeficiency virus type 1, and (ii) HIV-1 group N, thus far identified only in a few AIDS patients from Cameroon. A third lineage, HIV-1 group O also falls within the SIVcpz*Ptt* radiation, but the ape reservoir of this virus has not yet been identified. First, we report here, the geographic distribution and the genetic diversity of groups M- and N-like viruses in wild chimpanzee communities in southern Cameroon and secondly the detection and molecular characterization of SIVs closely related to HIV-1 group O in wild-living gorillas (*Gorilla gorilla*) in the same country. More than 1300 ape fecal samples were collected at 18 remote forest sites in Cameroon. Overall, 62% were from chimpanzees and 21% were from gorillas. The remainder were found to be degraded or from other primate/mammal species following mitochondrial DNA analysis. All were tested for HIV crossreactive antibodies using a commercial HIV-1/2 confirmatory assay. Thirty tree different *Pan troglodytes troglodytes* apes were found to be SIV-infected and six samples, corresponding to three different gorillas (as determined by microsatellite analysis) contained antibodies reactive with the HIV-1 envelope glycoprotein (gp41). Fecal RNA was isolated and partial pol and/or gp41 sequences were amplified by RT-PCR. Phylogenetic analysis of these SIV sequences showed that the 33 newly identified SIVcpz and the 3 SIVgor strains fall within the HIV-1/SIVcpz*Ptt* radiation. The identified SIVcpz strains were characterised by a high genetic diversity and a phylogeographic clustering. The latter allowed us to trace the origins of HIV-1 group M and group N to distinct chimpanzee communities in southern Cameroon. Phylogenetic analysis of the 3 SIVgor strains revealed a monophyletic lineage within the SIVcpz*Ptt* radiation which was most closely related to HIV-1 group O. We also confirmed further absence of SIVcpz infection in 78 samples from *Pan troglodytes vellerosus.*

These findings showed that chimpanzees likely served as the primary reservoir of SIVs now found in chimpanzees, gorillas and humans. HIV-1 groups M and N clearly arose by transfer of viruses from chimpanzees to men, while the origin of HIV-1 group O is less clear. Chimpanzees could have transmitted group O-like viruses to gorillas and to humans independently, or they could have transmitted the virus first to gorillas, which in turn transmitted it to humans.

Symposium "Epidemiology and evolution of malaria"

(29) Molecular detection of malaria parasite *Plasmodium falciparum* in a member of *Anopheles hyrcanus* group from northern Iran

Saber Gholizadeh[1], Sedigheh Zakeri[1], Ahmad Raeisi[2], Mozaffar Mosslem[3], Navid Dinparast Djadid[1]

[1]*Malaria Research Group (MRG), Biotechnology Department, Pasteur Institute of Iran, P.O. Box 1316943551, Pasteur Avenue, Tehran, Iran.*

[2]*Center for Diseases Management and Control (CDMC), Tehran, Iran.*

[3]*Guilan University of Medical Sciences, Guilan province, Iran.*

Nested polymerase chain reaction which targets the conserved 18s small-subunit RNA genes of parasite, not only permits a malaria infection in *Anopheles* mosquitoes to be detected, but also allow each *plasmodium* species present to be detected. Mosquitoes were collected from Guilan province in northern Iran. After morphological identification and dissection, head and thorax of 197 pools (985 individual mosquitoes) of *Anopheles maculipennis* complex, *Anopheles sacharovi* and *Anopheles hyrcanus* group were used for DNA extraction. PCR amplified a 205 bp fragment of *Plasmodium falciparum* in one pool of *An. hyrcanus* group specimens from Fooman district. The PCR method shows greater sensitivity and specificity and confirmed the existence of *P. falciparum* in *An. hyrcanus* group that so far has been considered as non-vector in Iran. This unexpected presence of *P. falciparum* in *Anopheles hyrcanus* population urges prompts investigation and standardization of control methods in Islamic Republic of Iran, Azerbaijan and Armenia.

Keywords: *An. hyrcanus* group; *P. falciparum*; Malaria; Guilan; Iran

(30) Molecular epidemiology of *Plasmodium falciparum* resistance to antimalarial drugs in Iran

S. Noorizadeh[1], M. Afshar Pad[1,2], N.D. Djadid[1], A. Raeisi[3], P. Gil[4], A. Bjorkman[4], S. Zakeri[1]

[1]*Malaria Research Group (MRG), Biotechnology Department, Pasteur Institute of Iran, P.O. Box 13164, Tehran, Iran.*

[2]*Research & Science Campus, Azad Islamic University, Tehran, Iran.*

[3]*Center for Diseases Management and Control, Tehran, Iran.*

[4]*Malaria Research Laboratory, Division of Infectious Diseases, Karolinska Institute, Stockholm, Sweden.*

As chloroquine (CQ) resistance spreads across Iran, the sulfadoxine/pyrimethamine (SP) combination therapy with CQ are being used as alternative first-line drugs for the treatment and prevention of *Plasmodium falciparum* malaria. We determined the genotype of three different genes, *P. falciparum* chloroquine resistance transporter (*pfcrt*), dihydrofolate reductase (*dhfr*), and dihydropteroate synthase (*dhps*) in 208 clinical *P. falciparum* isolates (pre-treatment and post-treatment) from Iran, which were collected during 2003–2005. DNA was isolated and analyzed by using polymerase chain reaction and restriction fragment length polymorphism (PCR/RFLP) to detect polymorphisms previously shown to be associated with resistance. The results showed that 77% of field isolates carried parasites with double mutant alleles of *pfdhfr* (C59R + S108N), while retaining a wild type mutation at position 51 and all *pfdhps* (S436F/A, A437G, K540E and A581E). 18.7% of field isolates carried parasites with double mutant alleles of *pfdhfr* (C59R + S108N) and single mutant alleles of *pfdhps* (A437G). Limited and stable polymorphism over the time particularly in *pfdhps* revealed that the SP is still effective as antimalaria drug in Iran. Furthermore the putative key codons of novel candidate gene for chloroquine resistance, *Pfcrt* was determined and the high levels of CQ pressure have let to strong selection of the *pfcrt*76T, 220S and 326S polymorphism among *P. falciparum* isolates in Iran. These results may have important implications for the future surveillance of both CQ and SP resistance by the use of molecular markers in Iran.

(31) Epitope mapping and analysis of sequence variation in VAR2CSA DBL3X involved in *P. falciparum* placental sequestration

Thomas S. Rask[2], Madeleine Dahlbäck[1], Pernille H. Andersen[2], Morten A. Nielsen[1], Nicaise T. Ndam[1,3], Mafalda Resende[1], Louise Turner[1], Philippe Deloron[3], Lars Hviid[1], Ole Lund[2], Anders Gorm Pedersen[2], Thor G. Theander[1], Ali Salanti[1]

[1]*Centre for Medical Parasitology at University of Copenhagen and Copenhagen University Hospital (Rigshospitalet), Denmark.*

[2]*Center for Biological Sequence Analysis, BioCentrum-DTU, Denmark.*

[3]*Institut de Recherche pour le Développement (IRD), Faculté de Pharmacie, France.*

Pregnancy-associated malaria (PAM) is a major health problem, which mainly affects primigravidae living in malaria endemic areas. The syndrome is precipitated by accumulation of infected erythrocytes in placental tissue through an interaction between chondroitin sulphate A (CSA) on syncytiotrophoblasts and a parasite encoded protein on the surface of infected erythrocytes, believed to be VAR2CSA. Women in endemic areas develop immunity to PAM and anti-VAR2CSA antibody levels correlate with protection. VAR2CSA is a polymorphic protein of approximately 3000 amino acids forming six Duffy-binding-like (DBL) domains. For vaccine development it is important to define the antigenic targets for protective antibodies and to characterize the consequences of sequence variation. In this study, we show that the VAR2CSA DBL3X domain mediates binding to CSA which makes it a leading vaccine candidate. We characterize sequence variation in the DBL3X domain, comprising single nucleotide polymorphisms, deletions and variable number of tandem repeats, using bayesian inference of selection pressure and recombination rate analysis. Combination of these results with structural modelling, shows that sequence variation mainly occurs in regions under strong diversifying selection, predicted to form flexible surface loops. From peptide array data we show that these regions are main targets of naturally acquired IgG and accessible for antibodies reacting with native VAR2CSA on infected erythrocytes. Interestingly, surface reactive anti-VAR2CSA antibodies also target a conserved DBL3X region predicted

to form an α-helix. Finally, we identify DBL3X sequence motifs that are more likely to occur in parasites isolated from primi- and multigravidae, respectively. These findings strengthen the vaccine candidacy of VAR2CSA and will be important for choosing epitopes and variants of DBL3X to be included in a vaccine protecting women against PAM.

Keywords: Malaria; VAR2CSA; Pregnancy; Vaccine; PfEMP1; DBL; Selection; Recombination; pepScan; Epitope

(32) Polymorphisms in the genes of interleukin 12 and its receptors in association with resistance to severe malarial anemia in children residing in western Kenya

Lyna Zhang[1], Donald Prather[1], Jodi Vanden Eng[1], Sara Crawford[1], Simon Kariuki[2], Feiko ter Kuile[3], Bernard Nahlen[1], Altaf A. Lal[1], Venkatachalam Udhayakumar[1], Ya Ping Shi[1]

[1]*Malaria Branch, Division of Parasitic Diseases, National Center for Zoonotic, Vector-Borne & Enteric Diseases, Coordinating Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30341, United States.*

[2]*Kenya Medical Research Institute, Kisumu, Kenya.*

[3]*Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK.*

Of the more than 1 million Africans who die from *Plasmodium* infection each year, most are children under five years of age. The majority of these deaths are from severe malarial anemia. P*lasmodium falciparum* has been shown to drive selection of human genetic variants for conferring protection against severe forms of malaria, such as severe malarial anemia (SMA). Malarial anemia is characterized by the destruction of malaria infected red blood cells and suppression of erythropoiesis. Recent studies in murine models of malarial anemia have demonstrated that interleukin 12 (IL12) significantly boosts erythropoietic responses. Furthermore, several immunological studies conducted in Africa have shown that IL12 production was suppressed in children with SMA compared to asymptomatic children. For these reasons the genes encoding the two IL12 subunits, IL12A and IL12B, and its receptors, IL12RB1 and IL12RB2, are attractive candidate genes for studying SMA. In this study, a total of 75 tagging single nucleotide polymorphisms (tagSNPs) covering these four genes were examined. Genotyping was performed with the iPlex MassARRAY technology (Sequenom) in a cohort of 940 children from the Asembo Bay region of western Kenya, an area with intense malaria transmission. Individuals possessing two copies of IL12A common allele (rs2243140) at 3′UTR showed increased susceptibility to SMA (Hb < 6 g/dl and the presence of *P. falciparum* $> 10{,}000/\mu L$) ($p = 0.009$, RR 1.85, 95%CI: 1.10–3.10). Individuals possessing two copies of a rare variant in IL12RB1 (rs429774) appeared to be strongly protected against SMA ($p = 0.0002$, RR 0.24, 95%CI: 0.08–0.72) and overall severe anemia ($p = 0.002$, RR 0.49, 95% CI: 0.29–0.82). Identification of genetic polymorphisms that influence human host susceptibility to malaria infection and severe disease outcomes may help us to better understand the immune response to malaria and design novel treatments against severe malarial anemia.

Symposium "Coevolution host pathogen 3"

(33) The development of functional apical—basolaterally differentiated midgut cell cultures for the study of arbovirus and *Plasmodium*- host interactions

Robert Butcher[1], Youngyuth Yuthavong[2], Buntika Areekul-Butcher[3]

[1]*National Nanotechnology Centre (NANOTEC), National Science and Technology Development Agency (NSTDA), 111 Thailand Science Park, Paholyothin Road, Klong 1 Klong Luang, Phathumthani 12120, Thailand.*

[2]*National Centre for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA), 113 Thailand Science Park, Paholyothin Road, Klong 1 Klong Luang, Phathumthani 12120, Thailand.*

[3]*Department of Biology, Faculty of Science, Chulalongkorn University, Phayathai Road, Paphumwan, Bangkok 10330, Thailand.*

The intractability of studying host–pathogen interactions within the arthropod mid-gut remains a problem in arthropod vector–pathogen research. For example, neither currently available cell lines nor primary mid-gut cell cultures support *Plasmodium* ookinete invasion, penetration and differentiation to oocytes whilst the absence of apical-basal differentiation and thus the control (temporal and cellular route) of mature infective arbovirus virion export is not correctly mimicked and indeed may be biochemically somewhat artifactual and affect cellular defective interfering virion function and accumulation.

Using a biphasic culture system and nanofiber based lamina hydrogels, we have established apical—basal differentiated primary mid gut cell cultures from Anopheline and Culicine mosquitoes. Although heterogeneous in cell morphology, they are principally comprised of columnar epithelial cells including vATPase^{+ve} cells with mid-gut specific gene expression, as determined by RT-PCR. These cultures support *Plasmodium* ookinete invasion, penetration and basal lamina differentiation to oocytes (ca. 20% of cells) and so potentially provide the first amenable cell culture system for the detailed proteomic and genomic analysis of *Plasmodium gallinaceum*, *P. falciparium* and *P. vivax* mid-gut invasion and subsequent differentiation. They may also prove to be a tractable model system for the screening of both potential resistance genes and allelic variances between parasite and host isolates, and the subsequent screening of the efficacy of transgenes prior to the labour intensive genetic modification of mosquitoes.

Moreover, following apical infection (ca. 90% of cells), Dengue and Japanese encephalitis viruses are replicated and baso-laterally secreted; suggesting these cell cultures may be amenable to arbovirus–host/symbiont interaction studies as well as virus–virus interaction within the same host tissue or cell. To illustrate this point we present preliminary results of inducible RNAi and defective interfering particle encoding phagemids upon dengue replication and baso-lateral secretion,

as well as interactions with natural symbionts, in these cell cultures.

(34) Susceptibility to hantavirus infection and MHC class II gene diversity in rodents

J.F. Cosson[1], J. Bryja[2], J. Deter[1], M. Galan[1], Y. Chaval[1], H. Henttonen[3], J.P. Hugot[4], J.R. Michaux[1], S. Morand[1], N. Charbonnel[1]

[1]*Centre de Biologie et de Gestion des Populations, CBGP, INRA, Campus International Baillarguet, CS 30016, 34988 Montferrier sur Lez, France.*

[2]*Institute of Vertebrate Biology, Academy of Sciences of the Czech Republic, Studenec, Czech Republic.*

[3]*Finnish Forest Research Institute, Vantaa, Finland.*

[4]*Museum National d'Histoire Naturelle, Origine et Evolution de la Biodiversité, 55 rue Buffon, 75231 Paris, France.*

Immunogenetics may provide key insight into epidemiology and transmission ecology. It may contribute to the understanding of the distribution of area of endemism and of the risk of emerging diseases in natural populations. One of the leading goals of immunogenetics has been to understand the associations of genetics to immune related diseases. In this context, the Major Histocompatibility Complex (MHC) has been extensively studied. It is a central component of the vertebrate immune system. There are multiple lines of evidence supporting the idea that this polymorphism is maintained by some form of balancing selection mediated by pathogens and parasites through frequency-dependent selection. Among MHC, Class II genes are known for their role in Puumala Hantavirus infections severity in humans. During this talk, we are going to analyse the genetic diversity of two class II MHC genes (DQA and DRB) at different evolutionary scales in rodents in relation with Hantaviruses distribution and phylogeny. First, we analysed the phylogenetic organisation of MHC allelic forms within vole and mice species in Europe and Southeast Asia (two area of endemism of human disease caused by Hantaviruses). This revealed the existence of trans-species polymorphism (TSP) among voles. Additional data on mice is needed to test whether TSP occur among mice and between vole and mice. TSP indicates that a balancing selection acts on these genes, probably through the mediated-selection exerted by shared pathogens. Second, we searched for associations between MHC haplotypes and the presence of Hantavirus in the bank vole (*Clethrionomys glareolus*), the reservoir of the Puumala Hantavirus responsible of Hemorrhagic Fever with Renal Syndrome in humans. Voles were serologically checked for antibodies. Associations between genetic parameters (haplotypes or heterozygosity) and infection status were explored using multivariate analyses. We detected significant associations between one MHC-haplotype and the susceptibility to Hantavirus infection. Similar studies are under progress among other rodent species in Europe and Southeast Asia. Our preliminary results highlights the potential importance of immunogenetics in understanding the emerging of rodent-born diseases like hantavirus infection.

(35) Rodents biodiversity and associated infections in Southeast Asia

J.P. Hugot[1], Y. Suputtamongkol[2], S. Jittapalapong[3], S. Morand[4], J. Michaux[5], J.F. Cosson[5], G. Dobigny[5], V. Herbreteau[6]

[1]*OSEB, UMR 5202 CNRS, MNHN, Paris, France.*

[2]*Siriraj Hospital, Bangkok, Thailand.*

[3]*Faculty of Veterinary Medicine, Kasetsart University, Thailand.*

[4]*ISEM, CC064, Université Montpellier-2, France.*

[5]*CBGP, CS30016, Montferrier-sur-Lez, France.*

[6]*IRD, Paris, France.*

We are currently involved in a Franco-Thai program devolved to a multidisciplinary investigation of Muridae rodents (''mice and rats''), their parasites and the pathogens that they may carry and/or transmit to human, with a better understanding of diseases emergence as an ultimate goal. In such a context, pathogens circulation in the wild is a complex but pivotal phenomenon which requires a continuum of scientific approaches to be accurately apprehended. In particular, a rigorously comparative study is mandatory in order to take into account the interactions between parasites, wild and domestic hosts (uncluding human) and their environment. This is the reason why we rely on both concepts and techniqueq from a wide range of disciplines including taxonomy, cytogenetics, phylogenetics, phylogeography, population genetics, ecology, geography as well as modeling. With the following objectives.

Objective 1: To precisely identify and characterize the rodent species acting as reservoirs, and to document their ecology, geographic distribution as well as the genetic structure of their populations. To assess parasite and pathogen diversity in relation to their associated rodent hosts. To provide co-phylogenies and co-phylogeographies in order to enlight the evolutionary relationships of hosts and parasites.

Objective 2: To map and correlate the observed rodents distributions with their species-specific environmental landscapes in order to extrapolate their potential ''real'' range and to anticipate their future distributions in relation to landscape modifications.

Objective 3: Computing epidemiological databases. Cross-checking field data and GIS data. Definition and standardization of a risk-scale. Finalization of maps of risks and distributions. Atlas of Thai Muridae. Reference collection of Thai Muridae. Education and training of students.

Following the listing of the objectives, we expose the first results of our studies and we sketch future projects.

(36) Human macrophage variability of *in vitro* infection by *Leishmania donovani*: A new approach to dissect human susceptibility to visceral leishmaniasis

Frédéric Vezilier[1], Rachel Bras-Goncalves[2], Cristina Riera[3], Salima Fatha[2], Roser Fisa[3], Teresa Serra[4], Maria Udina[5,6], Jaume Carrio[3], Enric Girona[5,6], Teresa Jimenez[5,6], Alejandro Forteza[5,6], Matil de Sedeiro[5,6], Mallorie Hide[1], Anne-Laure Banuls[1], and Bruno Bucheton[1]

[1]*UMR CNRS/IRD 2724 Genetics and Evolution of Infectious Diseases, Montpellier, France*.
[2]*IRD UR008 Pathogénie des Trypanosomatidae, Montpellier, France*.
[3]*Laboratorio de Parasitologia, Facultad de Farmacia, Universtitat de Barcelona, Spain*.
[4]*Servei de Microbiologia, Hospital Universitari San Dureta, Palma, Spain*.
[5]*Fundacio Banc de Sang i Teixit de les Illes Balears, Palma, Spain*.
[6]*Institut Universitari de Investigacio en Ciencies de la Salut (IUNICS), Palma, Spain*.

In its human host, *Leishmania donovani* is an intracellular parasite of macrophages and parasitism display a wide spectrum of clinical manifestations ranging from asymptomatic infections to fatal visceral leishmaniasis (VL). A bulk of evidence from mice and man now indicate that host genetic factors play an important role in determining susceptibility to VL. Until now studies aimed at unravelling human genetic susceptibility to VL have focussed on the analysis of clinical phenotypes. These clinical phenotypes probably result from the interaction of both host and parasite genetic factors but also environmental factors, making these phenotypes complex for genetic dissection. We propose here a novel alternative approach based on the study of a cellular phenotype established *in vitro*. In this study we have established the *in vitro* macrophage infection status of 20 healthy blood donors and their response to IFNγ. We demonstrate the existence of an important inter-individual variability of the macrophage permissive phenotype (parasitic index ranging from 20 to 100 amastigotes/100 macrophages) as well as important differences in the individual response to IFNγ (parasitic index reduced in only half of the subjects upon IFNγ activation). In a second step we attempted to correlate the observed cellular phenotypes with the expression of various macrophage genes implicated in *Leishmania* recognition (CR1, MSF1R), macrophage activation/deactivation (IL1, IL6, IL10, IL12, RANTES, MIP1β, TNFα) and macrophage microbicide activity (NOS2, ARG1, NADPH p40phox). Significant correlations were found between the ratio of NOS2/ARG1 and the p40 phox unit of the NADPH oxidase but none of the measured variables were explicative of the capacity to respond to IFNγ. These preliminary results indicate that this approach provide a valuable tool to analyse host-parasite interactions at the host cell level. Indeed studies in mouse derived macrophages or human monocytic cell line have shown that infection by *Leishmania* has a profound influence on the macrophage transcriptome. However studies are lacking that address specifically the question of the variability of macrophage responses to *Leishmania* infection in human and its relation with resistance/susceptibility to VL.

Symposium "Molecular biology and epidemiology of pathogens and vectors"

(37) Detection of glutathione *S*-transferase e2 (*gste*2) gene in Iranian and Pakistani populations of main malaria vector *Anopheles stephensi*

Nooshin Bayat[1,2], Saber Gholizadeh[1], Sedigheh Zakeri[1], Navid Dinparast Djadid[1]
[1]*Malaria Research Group (MRG), Biotechnology Department, Pasteur Institute of Iran, P.O. Box 1316943551, Pasteur Avenue, Tehran, Iran*.
[2]*Biology Department, Science and Research Branch, Islamic Azad University, Iran*.

One of the most important genes involved with metabolic insecticide resistance (specially DDT resistance) in *Anopheles gambiae* is the epsilon class of the glutathione *S*-transferase super family. In current study, PCR analysis of gste2 region have shown nucleotide variations within populations of *Anopheles stephensi,* the most important malaria vector in Iran and middle east. Specimens were collected from three different zones including; Chabahar, Sarbaz, Nikshahr, Iranshahr, Saravan and Khash districts in Sistan and Baluchistan province, Iran, which are under insecticide application for a long time; areas that has not been treated with insecticides for a long time (Kazeron), Fars province, Iran; and *An. stephens*i population from Pakistan. The result revealed that Iranian strains collected from Sistan and Baluchistan province were 100% identical in GSTe2 DNA sequences except Saravan strain which has showed 100% identity with Pakistani strain, and 99% identity with others. Kazeroon strain was 99% identical with both Pakistani and Iranian strains with a C → G transversion and A → C transition in 105th and 174th nucleotides, respectively. Pakistani and Saravan strains showed A → G transition in position 243 and C → T transition in nucleotide number 351. The follow up study on further specimens from those areas has detected two types of nucleotide variation in Sarbaz samples; one type is identical to Saravnian and Pakistani samples and the other type is similar to other Sistan and Baluchisatn samples. However, in amino acid level, all the sequences were 100% identical proving that the nucleotide variation which was observed does not involve with the insecticide resistance. We will further discuss the cloning results related to *gste2* in Sarbaz populations of *An. stephensi*.

Keywords: *Anopheles stephensi*; gste2; Glutathione *S*-transferase; DDT resistance; Malaria vector

(38) Detection of *Wolbachia* spp. in Iranian *Culicidae* species based on 16srDNA and wsp genes

Naimatallah Daneshinia[1,2], Saber Gholizadeh[1], Sedigheh Zakeri[1], Navid Dinparast Djadid[1]
[1]*Malaria Research Group (MRG), Biotechnology Department, Pasteur Institute of Iran, P.O. Box 1316943551, Tehran, Iran*.
[2]*Biology Department, Science and Research Branch, Islamic Azad University.*

Wolbachia are a group of cytoplasmically inherited bacteria that cause reproduction alterations, including parthenogenesis, cytoplasmic incompatibility, feminization of genetic males and male killing. Based on 16srDNA and protein-coding gene (wsp) sequences wolbachia can be typed into six different super groups or strains (A, B, C, D, E and F). Three A, B and E

types and member of F type are found in arthropods. In this study, we aimed to detect the *wolbachia* in Iranian *Culicidae* (*Anopheles*, *Aedes*, and *Culex* spp) specimens by molecular technique. Mosquitoes were collected from Guilan, Kerman, Sistan and Baluchistan, Khorasan, Azarbijan, and Hormozgan provinces in Iran. DNA extracted from 204 specimens was amplified by using specific diagnostic PCR of *wsp* and *16srDNA* genes that can differentiate between A and B groups of *wolbachia*. The length of amplified fragments by 16srDNA B and wsp B primers was 261 and 442 nucleotide in sequenced specimens, respectively. Within Iranian *Culicidae* only B type of *wolbachia* was detected, while there was no discrepancy in the amplified products of *16srDNA B* and *wsp* B. Therefore, we postulate that the similarity of 100% in 16srDNA and *wsp* genes could be a reason for close evolutionary relation in *Wolbachia pipientis* populations circulating in Iranian *Culicidae* species.

Keywords: *Culicidae*; *Wolbachia pipientis*; Group B; 16srDNA; *wsp* genes

(39) Genetic variability of tick-borne bacterial pathogens in Central Europe, and the development of DNA chip.

M. Derdáková[1,2], D. Blaškovič[3], R. Ivanová[1], I. Cíglerová[2], D. Mullerová[3], I. Barák[3], L. Grubhoffer[4], B. Peťko[1]

[1]*Parasitological Institute, Slovak Academy of Sciences, Košice, Slovakia.*

[2]*Institute of Zoology, Slovak Academy of Sciences, Dúbravská cesta 9, Bratislava, Slovakia.*

[3]*Institute of Molecular Biology, Slovak Academy of Sciences, Dubravská cesta 23, Slovakia.*

[4]*Institute of Parasitology, Academy of Sciences of the Czech Republic, České Budějovice Czech Republic.*

In the recent years due to the global warming, agricultural changes as well as other human impact on the environment, several tick-borne pathogenic organisms have (re)-emerged in Europe. A total of 991 ticks from two different geographical areas in Slovakia (Carpathians Mountain and Pannonian Plain) and one in the Czech Republic (Czech Massif) were tested by PCR analysis for the presence of *Borrelia burgdorferi* s.l. and the members of the family Anaplasmataceae. The overall borrelial prevalence varied between 17–36%. *B. afzelii, B. garinii, B. valaisiana, B. burgdorferi* s.s. were present in all three sites. Moreover, *B. lusitaniae* was present at the locality in North-Central sub-mountain area of Slovakia where 80% of all positive ticks belonged to this species. *B.* garinii (43%) was consistently the predominant species in Carpathian basin regions of Eastern Slovakia. The most common species in the Czech Republic was *B. afzelii* (52%). The sequencing of 5S–23S rDNA revealed high intraspecific variability within the detected species. The DNA of *I. ricinus* ticks was also tested for the presence of *Anaplasma/Ehrlichia* spp. by PCR-SSCP analysis and sequencing of a variable 247 bp fragment of 16S rDNA. *Anaplasma phagocytophilum* was detected in 4.3% of ticks. Moreover 2.1% ticks were infected with *Neoehrlichia mikurensis,* and 0.6% ticks carried an *Anaplasma*-like microorganism, recently detected in *I. ricinus* ticks from Northern Africa. Furthermore, we have developed a new detection system—microarray-based assay of bacterial tick borne pathogens present in Central-Europe. Our oligo-chip is a sensitive and reliable for detection of following genera: *Borreliae*, *Rickettsiae*, *Anaplasmae* and *Ehrlichiae*. This method is based on detection of fluorescent signal after hybridization reaction by using laser scanner, and has a potential of use in the large-scale epidemiological studies as well as fast diagnostic method, since all relevant pathogens can be detected in one step.

Keywords: *Ixodes ricinus*; *Borrelia burgdorferi* sensu lato; *Anaplasma*; *Ehrlichia*; Oligo-chip; Genetic variability; Tick-borne pathogens

(40) Molecular detection of feline hemoplasma in stray cats in Bangkok, Thailand

T. Jiyipong[1], P. Amavisit[2], S. Tasker[3], I. Peters[3], R.W. Stich[4], S. Jittapalapong[5]

[1]*Department of Genetic Engineering, Kasetsart University, Thailand.*

[2]*Department of Microbiology and Immunology, Faculty of Veterinary Medicine, Kasetsart University, Bangkok, Thailand.*

[3]*School of Clinical Veterinary Science, University of Bristol, Langford, Bristol, UK.*

[4]*Department of Veterinary Pathobiology, University of Missouri, Columbia, MO, USA.*

[5]*Department of Parasitology, Faculty of Veterinary Medicine, Kasetsart University, Bangkok, Thailand.*

The stray cat population in Thailand is a considerable problem because they are usually poorly cared for and thought to carry infectious agents. Such agents include the feline 'haemoplasma' species, *Mycoplasma haemofelis* and '*Candidatus M. haemominutum*', together previously known as *Haemobartonella felis*. The larger species, *M. haemofelis*, is an etiologic agent of infectious hemolytic anemia in cats, while the less virulent *M. haemominutum* has only been associated with disease in immunosuppressed cats. These pleomorphic bacteria adhere to erythrocytes and can spread among cats by blood transfer. Although not yet confirmed, these apparently obligate prokaryotic parasites are thought to be naturally transmitted by arthropod vectors that are abundant in tropical climates including Thailand. Conventional detection of *M. haemofelis* and *M. haemominutum* involves light microscopy of stained blood smears, and these organisms can be difficult to detect and differentiate from each other and stain artifacts. The objective of this study was to investigate the distribution of these pathogens in stray cats from Bangkok. To accomplish this we utilized applied molecular methods to detect and differentiate *M. haemofelis* and *M. haemominutum.* Five hundred blood specimens randomly derived from stray cats, were tested with a PCR-RFLP assay. This technique was able to detect and distinguish these two feline haemoplasma species based on their 16S rRNA gene (16S rDNA) sequences. Digestion of amplicons with *Hind* III yielded 76 and 117 bp fragments for *M. haemominutum*, and a single 170 bp fragment for *M. haemofelis.* Of the total specimens tested, 7.8% (39/500) were positive for *M. haemominutum* and 5.4% (27/500) were positive for *M. haemofelis*. Sequences from 170 and 193 bp amplicons

were 98 and 99% identical to corresponding 16S rDNA reported for *M. haemofelis* and *M. haemominutum*, respectively. To our knowledge, this is the first report of both feline haemoplasma species in cats from Bangkok.

Keywords: *Mycoplasma haemofelis*; *Candidatus Mycoplasma haemominutum*; PCR; Bangkok; Stray cats

Symposium "Modeling infectious diseases dynamics and diversity"

(42) Mechanical transmission of pathogens by tabanids: Development of a mathematic model; consequences in epidemiology and population genetic

Marc Desquesnes

Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), UR-Trypanosomes, TA30G Campus de Baillarguet, 34398 Montpellier, France.

It is known for long that a number of pathogens (trypanosomes, rickettsia, bacteria, viruses) circulating in the blood of their mammalian hosts can be mechanically transmitted from a host to another by biting insects (Krinsky, 1976; Foil, 1989). However, parameters of mechanical transmission are not well known, as well as medical, epidemiological, economic and genetic impacts. A series of experiments carried out in cattle allowed to demonstrate mechanical transmission of trypanosomes (*T. vivax* and *T. congolense*) by tabanids (*Atylotus agrestis* and *A. fuscipes*) (Desquesnes and Dia, 2003a,b, 2004). High transmission power of tabanids was demonstrated as well as pathogenicity of trypanosomes mechanically transmitted. *T. vivax* is more easily mechanically transmitted than *T. congolense* due to its high parasitaemia. Data collected daily during the experiments allowed to develop a mathematic model of the transmission which dynamic has important epidemiological applications. The epidemiology of cattle trypanosomosis due to *T. vivax* in the absence of tsetse (Latin America) is very different from that observed in the presence of cyclical vectors (Africa) (Desquesnes, 2004). Medical and economical impacts are also different. In mixed transmission areas, the relative impact of mechanical transmission is very difficult to establish. In the presence of tsetse, cattle trypanosomosis is most often highly endemic with prevalence regularly above 70%. In such conditions, addition of mechanical transmission to cyclical one has little epidemiological impact; however, it contributes to the predominance of *T. vivax versus T. congolense*. Coming to population genetic, mechanical transmission increases the circulation of *T. vivax* amongst cattle, which tends to homogenize the parasitic genetic material present in hosts. Transferring a very little quantity of blood, mechanical transmission tends to clone parasites; not only, it tends to select the most prolific parasitic sub-populations. Consequences on genetics of mechanically transferred pathogens should be studied.

Keywords: *Trypanosoma* sp.; Tabanids; Mechanical transmission; Mathematic model; Epidemiology; Population genetic

(42) Ecological and immunological factors in tuberculosis transmission and control

Gabriela Gomes

Instituto Gulbenkian de Ciência Apartado 14, 2781-901 Oeiras, Portugal.

Humans are exposed to populations of environmental mycobacteria (EM) whose composition varies between regions. Human populations are subject to variable potentials for Mycobacterium tuberculosis (Mtb) transmission due to differences in demographic and socioeconomic backgrounds. The bacille Calmette-Guérin (BCG) vaccine is widely used worldwide but its efficacy has revealed great variability against pulmonary tuberculosis. We develop mathematical models that describe the transmission of Mtb under constraints that are imposed by host immunity elicited by previous exposures to EM, Mtb and BCG. We describe how levels of tuberculosis and vaccine efficacy depend on the hypothesized interactions among the three mycobacterial populations. We determine a threshold in Mtb transmission – the reinfection threshold – above which tuberculosis endemicity is high and insensitive to both EM and BCG. By contrast, variability rules below the reinfection threshold.

(43) Of mice and men—Asymmetric interactions between *Bordetella* species

Olivier Restif[1], Daniel N. Wolfe[2], Elizabeth M. Goebel[2], Ottar Bjornstad[2], Eric T. Harvill[2]

[1]*University of Cambridge, UK*.

[2]*Pennsylvania State University, USA*.

Bordetella pertussis and *B. parapertussis* are two closely related human pathogens causing whooping cough. *B. parapertussis* has received limited interest because its symptoms are typically milder, and its incidence allegedly lower, than those of *B. pertussis*. However, some epidemiological studies suggest that the prevalence of the two pathogens may actually be similar, because of under-reporting of *B. parapertussis* cases. Sustained coexistence of these two competing species is surprising. Recently, experiments in mice have demonstrated that cross-immunity between *B. pertussis* and *B. parapertussis* is not symmetric: immunity induced by *B. parapertussis* infection efficiently protects against subsequent infections by either species, while immunity induced by *B. pertussis* infection does not efficiently protect against *B. parapertussis* infections. Using mathematical models, we explored the possible consequences of this asymmetry on the coexistence of the two pathogen species at the population level. In particular, we investigate the effects of anti-pertussis vaccination and fitness variation on both the short-term dynamics and the longer-term equilibrium of the system.

(44) Reciprocity between modelling and experiment to meet the challenge of controlling antigenically diverse pathogens

Lisa White

Epidemiology and Ecology Group, Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK.

Previous models for antigenically diverse pathogens have predicted vaccine induced strain replacement under certain con-

ditions. To assess the risk of such an event, predictive models for the transmission dynamics of these pathogens must be developed. When modelling the transmission dynamics, there is a conflict between the level of detail required to capture the interactions between strains and the sparseness of the data available to parameterise the models. It is necessary therefore to make simplifying assumptions about individual immune responses to develop viable models. This presentation will demonstrate how such assumptions impact on the population level behaviour of antigenically diverse pathogens. I will highlight the need for further experimental investigation to develop and parameterise models to a point where they have predictive value. As an example I will present a case study of human respiratory syncytial virus in rural Kenya where models were involved in the design of a cohort study. This example will demonstrate the challenges and potential for the fusion of modelling and experiment to address this problem.''

Symposium "Molecular epidemiology of tuberculosis and leprosy"

(45) A bioinformatic analysis of tuberculosis cases in New York

Jeffrey Driscoll[1], Inna Vitol[2], Kristin Bennett[2,3], Jamie Sommer[4], Barry Kreiswirth[5]

[1]*Wadsworth Center, New York State Department of Health, P.O. Box 22002, New Scotland Avenue, Albany, NY 12201-2002, USA*.

[2]*Computer Science Department, Rensselaer Polytechnic Institute, Troy, NY, USA*.

[3]*Mathematical Sciences Department, Rensselaer Polytechnic Institute, Troy, NY, USA*.

[4]*Bureau of Tuberculosis Control, New York State Department of Health, Albany, NY, USA*.

[5]*Public Health Research Institute, Newark, NJ, USA*.

The systematic DNA fingerprinting of human pathogens has led to valuable insights into the dynamics of disease transmission, evolution of strain families, and global movement of pathogens. In the case of tuberculosis (TB), multiple laboratory methods of varying resolutions have been developed to differentiate strains of the *Mycobacterium tuberculosis* complex (MTC). The three most commonly employed methods today are spacer oligonucleotide typing (spoligotyping), mycobacterial interspersed repetitive unit (MIRU) analysis, and insertion sequence IS*6110*-based RFLP typing. Currently in New York an MTC isolate of every newly identified case of TB undergoes spoligotype, MIRU, and IS*6110*-based RFLP analyses. This system of universal TB genotyping has produced a large collection of diverse data. Bioinformatic tools are being developed which permit analysis of TB genotype and epidemiologic databases *in toto*. In this study we utilized two web-based software tool collections to analyze New York spoligotype data: SPOTCLUST (Vitol et al., 2006) and SIT VIT (Brudey et al., 2006). SIT VIT was also used for examining MIRU data. By combining basic patient demographic data and MTC global family assignments we examine such questions such as the degree of recent TB infection in non-US-born persons, differences in strain families versus age at TB diagnosis, and global MTC family evolution. In New York City epidemiologists identified a group of *M. bovis* infections in US-born children (Winters, 2005). This same group is observed by graphing age at diagnosis versus TB strain global family. Graphs plotting patient MTC strain family versus country of birth and time in the US suggests compartmentalization of some strain families from particular areas of the globe. US-born patients in New York are only rarely observed to be infected with East Asia/India (EAI) associated strains. Bioinformatics analyses of TB genotyping and epidemiologic data could play an important role developing TB control and prevention programs.

Keywords: Tuberculosis; Bioinformatics; Databases; Molecular epidemiology; Genotyping

References

Brudey, K., Driscoll, J., Rigouts, L., Prodinger, W., Gori, A., Al-Hajoj, S., Allix, C., Aristiumuno, L., Baumanis, V., Bouza, E., Cataldi, A., Cheong, S., Ellermeier, C., Evan, J., Fauville, M., Ferdinand, S., de Viedma, D., Garzelli, C., Gazzola, L., Gomes, H., Guttierez, M., Hawkey, P., van Helden, P., Kadival, G., Kreiswirth, B., Kremer, K., Kubin, M., Kulkarni, S., Lillebaek, T., Ly, H., Martin, C., Martin, C., Mokrousov, I., Narvskaia, O., Ngeow, Y., Naumann, L., Niemann, S., Parwati, I., Rahim, M., Rasolofo-Razanamparany, V., Rasolonavalona, T., Rossetti, M., Rusch-Gerdes, S., Sajduda, A., Semper, S., Seth, P., Shemyakin, I., Singh, U., Somoskovi, A., Skuce, R., van Soolingen, D., Streicher, E., Suffys, P., Tortoli, E., Tracevska, T., Vincent, V., Victor, T., Yap, S., Portaels, F., Rastogi, N., Sola, C., 2006. An appraisal of the geographic prevalence of major genotyping families of *Mycobacterium tuberculosi*s complex through the updated SpolDB4 database. BMC Micro. 6, 23 pp.

Winters, A., Driver, C., Macaraig, M., Clark, C., Munsiff, S., Pichardo, C., Driscoll, J., Salfinger, M., Kreiswirth, B., Jereb, J., LoBue, P., Lynch, M., 2005. Human tuberculosis caused by *Mycobacterium bovis*—New York City, 2001–2004. MMWR 54, 605–608.

Vitol, I., Driscoll, J., Kreiswirth, B., Kurepina, N., Bennett, K., 2006. Identifying *Mycobacterium tuberculosis* complex families using spoligotypes. Infect. Genet. Evol. 6 (6), 491–504.

(46) Mycobacterium tuberculosis isolates from South India belongs to an ancient lineage

Sujatha Narayanan[1], Kathy Deriemer[2], Sebastien Gagneux[2], Lalitha Hari[1], Anthony Tsolaki[2], Suganthi Rajasekhar[1], P.R. Narayanan[1], Peter Small[2]

[1]*Department of Immunology, Tuberculosis Research Center, Chennai 600031, India*.

[2]*Division of Infectious Diseases and Geographic Medicine, Stanford University Medical Center, Stanford, CA 94305, United States*.

The magnitude of the tuberculosis problem worldwide and the global traffic requires the application of effective approaches to decipher the portable genotype patterns of a particular region in comparison with the global patterns of

disease transmission. The genetic information resulting from high prevalence areas such as ours would prove useful for defining phylogenetic links that exists with TB genomes and for constructing models of genome evolution. This will also help in the elucidatation of the evolutionary history of ancient tubercle bacilli. We have used a number of genotypic tools like IS6110 RFLP, Spoligotyping, Deletion Microarray, and Deletion PCR to understand the lineage of *M. tuberculosis* in southern India. We used a standardized international protocol for Is6110 RFLP and spoligotyping was done according to the protocol supplied by (Isogen Bioscience BV, Manarssen) and instructions of the manufacturer. All of the spoligotype patterns were coded using the octal code system. We then referred to a standardized international database of spoligotype patterns, SpolDB3 available at http://www.pasteur-guadeloupe.fr/tb/spol3 www.pasteur-guadeloupe.fr/tb/spol3) to determine whether each pattern had been previously reported. We selected 25 isolates for comparative whole-genome hybridization using an Affymetrix DNA chip (Santa Clara, California, USA). We identified putative deletions in the experimental strains relative to the sequenced reference strain H37Rv using DelScan software (AbaSci, San Pablo, California, USA), and confirmed the putative deletions by direct sequencing. For the detection of LSPs by multiplex real-time PCR, we designed a series of assays based on different TaqMan primer/probe combinations. We also used previously published studies of genomic deletions to identify phylogenetically important genomic deletions and screen for them using PCR. Our results showed that the major clade present in Tiruvallur, S. India belong to the group which is designated as the most ancient lineage of *M. tuberculosis*.

(47) The use of proteomics and bioinformatics to identify novel antigens of *Mycobacterium leprae* towards the development of a rapid diagnostic test for the early detection of leprosy

John S. Spencer[1], Hee Jin Kim[1], Rocio Lopez Roa[1], Maria A.M. Marques[1], Renato da Silva Filho[1], Nathan A. Groathouse[1], Diana L.Williams[2], Thomas H.M. Ottenhoff[3], Annemieke Geluk[3], Sang-Nae Cho[4], Robert Gelber[5], Hazel M. Dockrell[6], Stewart T. Cole[7], Warwick J. Britton[8], Malcolm S. Duthie[9], Greg Ireton[9], Maria C.V. Pessolani[10], Richard A. Slayden[1], Patrick J. Brennan[1]

[1]*Department of Microbiology, Immunology and Pathology, Colorado State University, Fort Collins, CO, United States.*

[2]*Laboratory Research Branch, Division of the National Hansen's Disease Programs, Louisiana State University, Baton Rouge, LA, United States.*

[3]*Department of Immunohematology & Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands.*

[4]*Department of Microbiology, Yonsei University College of Medicine, Seoul, Republic of Korea.*

[5]*Leonard Wood Memorial Leprosy Research Center, Cebu, Philippines.*

[6]*Department of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK.*

[7]*Unité de Génétique Moléculaire Bactérienne, Institut Pasteur, Paris, France.*

[8]*Mycobacterial Research Group, Centenary Institute of Cancer Medicine and Cell Biology, Sydney, Australia.*

[9]*Infectious Disease Research Institute, Seattle, WA, United States.*

[10]*Department of Mycobacterial Diseases, Oswaldo Cruz Institute, Rio de Janeiro, Brazil.*

Comparative genomic analysis of the *M. leprae* genome has identified 1614 open reading frames, and up to 165 genes with no homologues in *M. tuberculosis*. Diagnosis of leprosy is a major obstacle to disease control, and has been compromised in the past by the lack of specific reagents. Understanding which antigens might be useful in either cell mediated or antibody response assays has been hampered by the fact that *M. leprae* is an uncultivatable microorganism, capable of growth only in humans, armadillos, and immunocompromised nude mice. Nevertheless, using proteomic analysis, we have identified 256 proteins from the native subcellular fractions of *M. leprae*, which represents approximately 16% of the potential open reading frames thought to be coded by the genome. In addition, using an oligonucleotide chip, we have further identified up to 702 genes that are thought to be transcriptionally active. Of these, there were 41 genes that coded for proteins that are unique to *M. leprae*, and 205 genes that coded for proteins that are conserved hypotheticals, proteins which are limited to mycobacteria. Using the information from these studies, a picture of those antigens that might be both specific to *M. leprae* and important in evoking either cell mediated or antibody responses in leprosy patients is unfolding. In previous work, we found that a number of recombinant *M. leprae* proteins or peptides behaved reproducibly as T cell antigens, inducing stronger IFN-γ responses from tuberculoid leprosy patients than tuberculosis patients or endemic controls. In addition, we are examining the potential of protein/peptide arrays to define the humoral immune response to identify disease-state-specific antigen profiles. Those antigens that provide specific responses in leprosy patients could be developed into a rapid diagnostic test for the early detection of leprosy and epidemiological surveys of the incidence of leprosy, of which little is known.

(48) Differentiation of *Mycobacterium* species that cause tuberculosis by genome analysis

S.E. Vasconcellos[1], R. Huard[2], S. Niemann[3], A.R. Santos[1], J.L. Ho[4], P.N. Suffys[1]

[1]*Laboratory of Molecular Biology Applied to Mycobacteria. FIOCRUZ, Rio de Janeiro, Brazil.*

[2]*Clinical Microbiology Service, Columbia University Medical Center, The New York Presbyterian Hospital, University Hospital Columbia and Cornell, New York, United States.*

[3]*National Reference Center for Mycobacteria, D-23845 Borstel, Germany.*

[4]*Division of International Medicine and infectious Diseases, Department of Medicine, Joan and Sanford I Weill Medical College, Cornell University, New York, United States.*

Mycobacterium tuberculosis is the principal etiologic agent of human tuberculosis (TB) in most regions of the world while *M. africanum* is frequent in certain regions in the Africa. *Mycobacterium bovis* is the principal cause of TB in cattle and causes TB in humans upon ingestion of unpasteurized milk and products or after prolonged contact with infectious cattle. The other species of the *M. tuberculosis* complex (MTC) *M. microti, M. pinnipeddi and M. caprae* are less frequently observed and the classification within the MTBC of rare variants such as the dassie and oryx bacilli remains to be defined. The genetic differences between MTBC strains, besides being the cause of host preference, can be useful for diagnosis and better define epidemiology and appropriated treatment, help identifying potential virulence markers and fine-tune MTBC phylogeny. Genetic characteristics such as deletions, insertions, inversions, duplications and Single Nucleotide Polymorphism (SNPs) are nowadays used to help identifying MTBC isolates to the species level. In the present study, 47 strains isolated from different TB patients from Ghana were characterized using either PCR-based protocols for detection of genomic deletion or PCR-RFLP for detection of earlier or presently described SNPs. A total of 13 deletions (IS1561', *cpf32,* RD701, RD702, RD713 RD711, TbD1, TbD1,RD7, RD8, RD9 and RD10) and 16 SNPs ($hsp65^{540}$, $rpoB^{1049/116}$, $katG^{203/463}$, $gyrB^{1450}$, $gyrA^{95}$, $Rv\ 1510^{1129}$,$aroA^{285}$, $PRO\text{-}narGHJI^{215}$, $3'cfp32^{311}$,$RD13^{174,}$ $PPE55^{2154/2148}$, nat^{751} and $RD711^{390}$) were analyzed. Using this approach, it was possible to differentiate all *M. tuberculosis* ($n = 12$) from *M. africanum* ($n = 29$) strains, and to recognize *M. africanum* subtype Ia ($n = 9$) and *M. africanum* Ib ($n = 20$). In addition, the presence of the hereby described SNPs (nat^{751} *and* $RD711^{390}$) were analyzed in a well-characterized collection of MTBC isolates (Huard, 2006); nat^{751} was only pesent in *M. africanum* Ia, while $RD711^{390}$ was only pesent in *M. africanum* Ib. This suggests that some species or sub-species within the MTBC can be differentiated looking at particular SNPs and that newly discovered SNPs could help fine tuning the MTBC to the species level.

Symposium "Molecular Epidemiology of Parasites"

(49) Identification of trypanosomes using molecular methods from wild caught tsetse in Tanzania

Emily Adams[1], Patrick Hamilton[2], Imna Malele[3] Wendy Gibson[1]

[1]*School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK.*

[2]*School of Biological Sciences, University of Exeter, Exeter EX4 4PS, UK.*

[3]*Tsetse and Trypanosomiasis Research Institute, PO Box 1026, Tanga, Tanzania.*

Tsetse flies transmit many species of trypanosomes in Africa, some of which are human and livestock pathogens of major medical and socio-economic impact. Identification of trypanosomes is essential to assess the disease risk posed by particular tsetse populations.

We have developed two tests to replace the multiple species-specific PCR tests used previously to identify the trypanosome species carried by individual tsetse flies.

1. Generic PCR test: interspecies size variation in the PCR product of the ITS-1 region of the ribosomal RNA repeat region enables species identification.
2. Fluorescent primers: size variation within the 18S rDNA region is used to identify species, products are sized with greater accuracy by the use of an automated DNA sequencer. This enables identification of species with increased precision and sensitivity.

Using the generic systems, we have been able to identify a new species of trypanosome. The 18s region of this trypanosome has been sequenced and we can confirm the new trypanosome is most closely related to *Trypanosoma godfreyi*, a suid trypanosome. The easy identification of this new trypanosome by PCR will facilitate studies of its epidemiology.

Both methods facilitate the identification of samples quickly and accurately, and have been used for large-scale field studies. The results of these studies and the respective advantages and disadvantages of the two methods will be discussed.

(50) The human African trypanosomiasis: Interactions between the tsetse fly, its secondary symbiont *Sodalis glossinidius*, and the parasite

Anne Geiger[1], Sophie Ravel[1], Thierry Mateille[2], Jérome Janelle[1], Delphine Patrel[1], Roger Frutos[1], Gérard Cuny[1]

[1]*UMR 17, IRD-CIRAD, CIRAD TA 207/G, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France.*

[2]*IRD, UMR Biologie et Gestion des Populations, CBGP, Campus International de Baillarguet, CS 30016, 34988 Montferrier-sur-Lez Cedex, France.*

Tsetse flies transmit African trypanosomes, the causative agents of sleeping sickness in human and Nagana in animals. This disease affects many people with considerable impact on public health and economy in sub-Saharan Africa, while trypanosomes resistance to drugs is rising. *Sodalis glossinidius,* a symbiont of tsetse flies, is considered to be involved in vector competence. In a former study no direct correlation was found between the presence of *S. glossinidius* and the ability of the insect to acquire *Trypanosoma congolense*. However, *Glossina palpalis gambiensis* and *Glossina morsitans morsitans* were shown to harbour genetically distinct populations of *S. glossinidius,* suggesting that vector competence for a given trypanosome species could be linked to the presence of specific genotypes of the symbiont rather than a mere presence /absence. In order to assess this hypothesis, *Glossina palpalis* individuals were fed on blood infected with either *Trypanosoma brucei gambiense* (T.b.g.) or *Trypanosoma brucei brucei* (T.b.b.) species, and the genetic diversity of *S. glossinidius* strains isolated from dissected flies was investigated using AFLP markers. Correspondence between

occurrence of these markers and parasite establishment was analysed using multivariate analysis. We demonstrated that the distribution of *S. glossinidius* strains from T.b.g.-infected flies differed strongly from those from T.b.b-infected individuals. Some AFLP markers were shown to be significantly linked to the ability of T.b.g. or T.b.b. to establish in the *Glossina* midgut. This suggests the differential presence of *Sodalis* genotypes to influence parasite establishment and could explain variations in *Glossina* vector competence in the wild. The markers identified may be useful to assess prevalence of *Sodalis* genotype facilitating establishment of parasite species, and to develop novel risk-management strategies.

(51) Microsatellite EmsB, a relevant tool to explore the genetic diversity of *Echinococcus multilocularis* at different geographical scales

Jenny Knapp, Jean-Mathieu Bart, Marie-Hélène Guislain, Akira Ito, Bruno Gottstein, Patrick Giraudoux, Renaud Piarroux

The cestode *Echinococcus multilocularis* is widely known in the Northern hemisphere to infect mainly foxes and rodents and accidentally human. However human seems not to be equal facing the alveolar echinococcosis (AE) disease. Many factors such as environmental conditions, immunology or genetic susceptibility can explain the heterogeneity of infection in human. Few data existed concerning the influence of the *E. multilocularis* genetic diversity. Up to now, the geno-tracking of subspecies was impossible due to a lack of fast evolutionary markers. Recently, a tandem repeated multi-locus microsatellite, named EmsB, was discovered in the *E. multilocularis* genome. Our objective was to assess the limit of EmsB discriminatory power at three geographical scales: (i) at a world scale, by studying Alaskan, European and Asian endemic areas, with parasitic lesions from definitive and intermediate hosts, (ii) at a continental scale by studying adult worms circulating in so-called old and new endemic areas in eight European countries and (iii) at a local scale ($n \times 100\ km^2$) in the French region of the Ardennes. The hierarchical clustering analysis of the results helped us to explore relevant questions as the spatial dynamic, the circulation of the parasite between different hosts, the emergence or the re-emergence of AE in human, to increase our knowledge on the parasite transmission and to enhance the public health management.

Keywords: *Echinococcus multilocularis*; Microsatellite; Different geographical scales; Transmission dynamic

(52) Identification of the iron superoxide dismutase gene repertoire in *Trypanosoma brucei gambiense*

Gustave Simo[1,2], Gérard Guny[3], Tazoacha Asonganyi[2], Flobert Njiokou[4], Stéphane Herder[3]

[1]*Medical Research Center, Institute of Medical Research and Study of Medicinal Plants (IMPM), PO Box 6163, Yaoundé, Cameroon.*

[2]*Faculty of Medicine and Biomedical Sciences, University of Yaoundé 1, Yaoundé, Cameroon.*

[3]*Laboratoire de Recherche et de Coordination sur les Trypanosomoses IRD/CIRAD, UR 177 (IRD), TA 207G, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France.*

[4]*Faculty of Science, University of Yaoundé 1, Yaoundé, Cameroon.*

Superoxide dismutase (SOD) forms part of the defense mechanism that helps to protect organisms from superoxide anions. This enzyme is one of the isoenzyme systems commonly used to differentiate T.b. gambiense from T.b. brucei and T.b. rhodesiense. To understand the genetic basis of the differences observed between SOD electrophoretic profiles of T. brucei sub-species, we undertook the identification and the characterization of SOD gene repertoire in T.b. gambiense. This study was performed on seven stocks (4 T.b. gambiense group 1 and 3 group 2) showing different SOD profiles. Four SOD genes (soda, sodb1, sodb2 and sodc) were identified in T.b. gambiense genome. These genes were cloned and their predicted amino acid sequences were deduced. Few differences were observed between nucleotide sequences of the four SOD genes of T.b. gambiense group 1 and 2 stocks. Even with T.b. brucei, few differences were observed. Several amino acids specific to FeSOD were found in the four SODs sequences of T.b. gambiense. Aligning the four T.b. gambiense protein sequences with those of others organisms, important differences were found with MnSOD and Cu/ZnSOD, but high similarity with FeSOD; indicating that the SODs of T.b. gambiense are FeSOD. High similarity exists between the proteins sequences of T.b. gambiense and T.b. brucei. Despite the differences observed in SOD electrophoretic profiles, there is a genetic stability of the SODs genes in *T. brucei* sub-species.

Symposium "Emerging viral diseases: Avian flu and SARS"

(53) Can the evolution of avian influenza virus (AIV) be predicted?

Prasert Auewarakul

Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand.

Because the previous influenza pandemics were caused by either viruses of avian origin or reassortant viruses with some of their genomes derived from avian viruses, it is believed that the H5N1 AIV, which already causes limited outbreaks in human, may become the source of the next pandemic. Continuously evolving, the virus may adapt to the point that it can transmit efficiently in human population. An insight in the viral evolution may help us to estimate the risk, understand the cause, and prevent the emergence of a pandemic strain. Trying to understand the viral evolution, we need to look at the viral dynamics and selection pressures. While the evolution of human influenza viruses is driven mainly by immunological selection pressure, it is not clear what drives the evolution of AIV. The current H5N1 AIV has diverged into multiple sublineages. Because these sublineages or clades are antigenically distinct, it is likely that immunological

pressure has played a role in the divergence of H5N1 AIV. It is also likely that the immunological pressure was the result of massive immunization in poultry, especially in China. The recent emergence of the new Fijian-like strains further supports this notion. In general, genomes of most AIV are in evolutionary stasis because of they are in equilibrium with their natural host. Once transmitted to a new host species, the virus starts rapid evolution to adapt itself to the new host. Virus–host interaction is, therefore, considered an important factor in the evolution of H5N1 AIV. Genome analysis has revealed positive selection pressure on some parts of the viral genome. Some of these sites may eventually fix and become host-specific residues when the virus reaches the optimum in the new hosts. Although the avian-human inter-species barrier may involve several genetic determinants, the most important one may be the receptor binding preference. A few point mutations in the receptor-binding pocket of hemagglutinin have been shown to increase human-type receptor binding of H5. We have found such mutation in viral quasispecies from a human respiratory specimen. This suggests that there is a selection pressure driving the virus toward human-type receptor specificity in human infection. Viral genomes may be shaped not only by selection pressures but also mutational bias. Recent observation has shown that avian and human influenza viral genomes are different in their GC content. It was suggested that this indicates a difference in mutational bias between the two hosts. However, I would like to point it out that this may not be entirely correct. Genome composition may also be influenced by codon composition, which could be affected by selection pressure. We analyzed codon volatility, which represents the propensity to change non-synonymously, and found some differences in the genome volatility between the two virus groups. Higher codon volatility in human viruses is likely the result of higher degree of evolution and changes, and may explain at least part of the difference in genome composition. We also found that hemagglutinin and neuraminidase genes of H5N1 AIV have high codon volatility suggesting that the genes are rapidly evolving. Such analyses may help us to identify viral genome regions with high potential to change, and may help us to predict the pattern of viral evolution.

(54) Avian influenza H5N1 in Russia

T.N. Ilyicheva, V.A. Evseenko, V.A. Ternovoy, A.G. Dyrimanov, S.I. Zolotikh, A.M. Shestopalov, S.V. Netesov, I.G. Drozdov

State Research Center of Virology and Biotechnology "Vector", Koltsovo, Novosibirsk Region 630559, Russia.

In July of 2005 in Suzdalka village (Novosibirsk Region, Russia) the mass mortality of poultry and wild birds was registered. Highly pathogenic influenza viruses of H5N1 subtype were isolated from different organs of three dead birds. The A/Turkey/Suzdalka/Nov-1/05 strain was isolated from turkeys' spleen, the A/Chicken/Suzdalka/Nov-11/05 and A/Chicken/Suzdalka/Nov-12/05 stains—from chickens' kidney. This was the first case of isolation of high pathogenic influenza viruses of H5N1 subtype in Russian Federation.

The phylogenetic analysis of hemagglutinin gene sequences showed that these strains form a cluster (support index is equal to 99) with H5N1 strains isolated from birds during the spring outbreak of the avian influenza virus in Qinghai Lake in China in 2005. The analysis of birds' migration ways indicates that virus was imported to Siberia during the spring migration in 2005. In summer and autumn of 2005 high pathogenic influenza viruses of H5N1 subtype spread apart all South part of Siberia, reached Caspian Sea, countries of Caucasian region and were registered in various countries of Black Sea basin. In winter of 2005–2006 H5N1 virus was isolated repeatedly in different European countries, in Turkey and in the Southern part of Russian Federation: in Dagestan and in Krasnodar Territory. In 2006 during the spring migration, H5N1 virus was isolated again in South of Western part of Siberia.

The phylogenetic analysis of sequences of genes that code hemagglutinin, neuraminidase, nonstructural protein (NS) and matrix protein showed that all the viruses isolated in period from July of 2005 till May of 2006 in territory of Russian Federation were genetically close to H5N1 virus isolated in 2005 in Qinghai lake (Northern China) and were the products of microevolution of this virus. The analysis of birds' migration ways and spreading of high pathogenic influenza viruses of H5N1 subtype in territory of Eurasia during 2005–2006 showed the important role of swamps and lakes of South of Western part of Siberia in expansion of H5N1 virus to European countries, countries of Near and Middle East and Northern Africa.

Acknowledgements

The authors show the sincere gratitude to everybody who helped authors to carry out this work, especially Dr. Robert G. Webster, Dr. Alexander Lipatov and Dr. Michael Callahan.

This work at the FRSI "Vector" was carried out under the auspices of the Government of the Russian Federation and financial support Russian Ministry of Public Health and BioIndustry Initiative, USA (grants BII/RUX2-20411-NO-04 and BII/RUB2-20429-NO-05).

(55) Genetic diversity of bat SARS-like coronavirus and its interaction with ACE2

Wuze Ren[1], Wendong Li[1,3], Xiuxia Qu[2], Meng Yu[3], Hongkui Deng[2], Shuyi Zhang[4], Linfa Wang[3], Zhengli Shi[1]

[1]*State Key Laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences (CAS), Wuhan, Hubei 430071, China.*

[2]*Department of Cell Biology and Genetics, College of Life Sciences, Peking University, Beijing, China.*

[3]*CSIRO Livestock Industries, Australian Animal Health Laboratory, Geelong, Victoria 3220, Australia.*

[4]*Institute of Zoology, CAS, Beijing 100080, China.*

Bats were recently identified as natural reservoirs of SARS-like coronavirus (SL-CoV) or SARS coronavirus-like virus. These viruses, together with SARS coronaviruses (SARS-CoV) isolated from human and palm civet, form a distinctive cluster within the group 2 coronaviruses of the genus *Coronavirus*, tentatively named group 2b (G2b-CoV). The bat G2b-CoV isolates have an identical genome organization and share an overall genome sequence identity of 88 to 92% among

themselves, and between them and the human/civet isolates. The most variable regions are located in the genes coding for nsp3, ORF3a, spike protein and ORF8 when bat and human/civet G2b-CoV isolates are compared. Genetic analysis demonstrated that a diverse G2b-CoV population exist in bat habitat. The spike protein (S protein) of coronavirus is known responsible for receptor binding to the host cells. The receptor binding domain (RBD) of human/civet G2b-CoV and its receptor angiotensin-converting enzyme-2 (ACE2) were well characterized. However, two deletion sites (5 and 12 aa, respectively) are located in S protein of bat G2b-CoV compared with that of human G2b-CoV. Thus the interaction between bat G2b-CoV S protein and the ACE of human and bat was investigated.

(56) Detection of avian influenza virus in lung tissues of naturally infected chickens in Thailand by *in situ* hybridization

Kriangkrai Witoonsatian[1,2], Theera Rukkwamsuk[2], Thaweesuk Songserm[2], and Theerapol Sirinarumitr[2]

[1]*Center for Agricultural Biotechnology, Kasetsart University, Kamphaengsaen, Nakorn-Pathom 73140, Thailand.*

[2]*Faculty of Veterinary Medicine, Kasetsart University, Kamphaengsaen Campus, Nakorn-Pathom 73140, Thailand.*

In situ hybridization (ISH) was developed to detect of avian influenza virus (AIV) infection in lung tissues of chickens in Thailand. Each of 30 samples of AIV-infected chicken and non-AIV-infected chicken lung tissues were used in this study. The nonradioactive digoxigenin (DIG)-labeled 604 base pairs (bp), 544 bp, and 274 bp cDNA probes for viral RNA encoding the matrix protein, hemagglutinin, and neuraminidase, respectively, of AIV type A H5N1 strain were generated by the reverse transcription polymerase chain reaction (RT-PCR). The hybrid formation was detected with anti–DIG conjugated alkaline phosphatase and nitrobule tetrazolium/5-bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) substrate. The strong positive signal typically was exhibited a dark purple reaction product in the nucleus and cytoplasm of the bronchial and bronchiolar epithelial cells with red background of the nuclear fat red-staining whereas the less intense signal was detected in the interstitial and alveolar macrophages. In contrast, all of non-AIV-infected chicken lung tissues were negative by ISH. Therefore, ISH developed in this study was useful for detection of AIV in tissues taken from naturally infected chickens.

Keywords: Avian influenza virus; Digoxigenin-labeled DNA probe; *In situ* hybridization; Chickens

Symposium "Evolution of animal and human trypanosomes"

(57) Molecular detection of *Trypanosoma lewisi*-like infections in rodents of Thailand

S. Jittapalapong[1], T. Inpankaew[1], N. Sarataphan[2], V. Herbreteau[3], J.P. Hugot[4], S. Morand[5], R.W. Stich[6]

[1]*Faculty of Veterinary Medicine, Kasetsart University, Bangkok, Thailand.*

[2]*National Institute of Animal Health, Pathumthani, Thailand.*

[3]*IRD, Paris, France.*

[4]*OSEB, UMR 5202 CNRS, MNHN, Paris, France.*

[5]*CBGP Campus International de Baillarguet FRANCE*

[6]*University of Missouri, Columbia, MO, USA.*

Trypanosoma lewisi is the type species for *Herpetosoma*, a homogenous subgenus of several dozen named species often described as morphologically indistinguishable, "*T. lewisi*-like" parasites. These parasites normally infect rodents and utilize fleas as vectors. Although this trypanosome subgenus is considered non-pathogenic to normal hosts, some of them are on rare occasion reported in association with human disease. Recently, a *T. lewisi*-like infection was detected in a sick Thai infant, thus the objective of this study was to investigate the prevalence of *T. lewisi* infections among different rodents indigenous to Thailand. Blood was collected from a total of 129 rodents trapped from urban and rural areas of three Thai provinces (Loei, Kalasin and Phrae) between 2005 and 2006. These samples were processed for DNA isolation and tested with a PCR assay universal for the genus *Trypanosoma*, followed by internal transcribed spacer 1 (ITS1) sequence analysis to identify infections in positive samples. Amplicons of approximately 623 bp, the size consistent with the stercorarian trypanosome group, were generated from 21.7% of all rodents tested. Further analysis suggested that ITS1 sequences from these amplicons were 98% identical to that reported for *T. lewisi* from an experimentally infected rat. Only two of six rodent genera tested PCR-positive for these parasites. The highest prevalence of *Herpetosoma* infections was found among rodents from Phrae province (26%). These results suggest that, in Thailand, humans exposed to certain rodents or their ectoparasites could be at risk of infection with *T. lewisi*-like parasites. More work is warranted to identify vectors of these trypanosomes and to compare the prevalence of these infections among rodent and human populations in Thailand.

Keywords: *Trypanosoma lewisi*; Rodents; Internal transcribed spacer 1 (ITS1); Thailand

(58) Impact of mixed infections on Benznidazole treatment efficacy in BALB/c mice infected with*Trypanosoma cruzi* major genotypes

H.R. Martins[1], R. Moreira Silva[2], H.M.S. Valadares[3], M.J.O. Toledo[4], G.L.L. Machado-Coelho[5,1], M.T. Bahia[1,6], A.O. Martins-Filho[7], A.M. Macedo[6], M. Lana[1,4]

[1]*Núcleo de Pesquisas em Ciências Biológicas (NUPEB), Instituto de Ciências Exatas e Biológicas (ICEB), UFOP, Brazil.*

[2]*Departamento de Análises Clínicas, Universidade Estadual de Maringá, Preto, Brazil.*

[3]*Departamento de Ciências Biológicas, ICEB, UFOP, Brazil.*

[4]*Departamento de Análises Clínicas, Escola de Farmácia, Universidade Federal de Ouro Preto, Brazil.*

[5]*Departamento de Farmácia, Escola de Farmácia, UFOP, Brazil.*

[6]*Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, UFMG, Brazil.*

[7]*Centro de Pesquisa René Rachou Fiocruz, MG, Brazil.*

In the present study the impact of *T. cruzi* dual-clone infections on benznidazol treatment efficacy compared with the respective monoclonal infections was investigated. For this, eight clonal stocks, two of each major genotypes 19 and 20 (*T. cruzi* I), hybrid genotype 39 and 32 (*T. cruzi* II) were combined into 24 different mixtures. BALB/c mice were inoculated by intraperitoneal route, with 5000 blood trypomastigotes of each clone and treated with oral doses of 100 mg benznidazol/kg/20 days. The cure control was performed by fresh blood examination, hemoculture, PCR, ELISA and detection of anti-live trypomastigotes antibody. The identification of each clone from not cured mice was performed by microsatellites assay. Cure in dual-clonal infection was detected in 28.4% of treated animals. Considering the cure rates of mice for all *T. cruzi* I (35,4%), T. *cruzi* II (60%) groups and their associations were not observed difference in relation to the expected benznidazole (BZ) susceptibility for combinations *T. cruzi* I + I (0%), I + II (22,1%), except II + II (60,0% susceptible). For major genotypes, combinations 19 × 32 (26.7%) and 19 × 39 (25.6%) shiped their phenotypes to resistant profile and 39 × 32 (60.8%) to susceptible profile. Genotype 20 was 100% resistant to BZ in monoclonal infections, but the cure rates of their combinations ranged from 0 to 24,5%. Nine out of 24 dual infections changed their profile of BZ susceptibility: 20 + 39 ($n = 2$), 20 + 32 ($n = 1$), 19 + 39 ($n = 3$), 19 + 32 ($n = 2$) and 39 + 32 ($n = 1$). Although molecular characterization had identified few mixed infections in isolates from not cured mice, very interesting results were observed. In some mixtures (sensitive + resistant), the selected clone identified after BZ treatment was that previously identified as sensitive to BZ in monoclonal infections. These results suggest that mixed infections, so current in nature, may have important impact on chemotherapy efficacy. Further studies to elucidate the mechanisms involved in this process are essential for advances in the knowledgment of Chagas disease chemotherapy.

(59) The epidemiology of surra (*Trypanosoma evansi* infection) in SE Asia: Can molecular tools provide the answers we need?

Simon Reid

School of Veterinary and Biomedical Sciences, Murdoch University, Western Australia, Australia.

Trypanosomiasis caused by *Trypanosoma evansi* (Surra) is endemic in SE Asia where it is a significant, but often underestimated cause of mortality in livestock. Recent estimates suggest that 22% of Indonesian buffalo infected with *T. evansi* either die or are sold for salvage slaughter. In the last 10–15 years a series of severe outbreaks of surra have occurred in the Philippines. What is most puzzling is that the epidemiology observed in these outbreaks differs from other parts of SE Asia such as Indonesia. These differences include the observation of fatal disease in small ruminants and cattle and the observation of different presenting signs. In addition, there have been reports of isolated cases of human trypanosomiasis in India. A variety of epidemiological tools are required in order to gain a better understanding of the biological basis of the epidemiological conundrum that we are faced with. In particular molecular tools can provide us with unique methods of providing us with insight but they must be used with caution to ensure that the benefit from the research reaches the key stakeholders—the farmer. This talk will describe a multidisciplinary approach to unraveling the epidemiology of surra in the Philippines.

(60) Human infection by *Trypanosoma evansi* in India: Diagnosis, treatment, genetic and epidemiological investigations

Philippe Truc[1], Benoit Vanhollebeke[2], Wendy Gibson[3], Stephane Herder[4], Philippe Poelvoorde[2], Annette Pays[2], Prashant P. Joshi[5], Rahul Katti[6], Vijay R. Shegokar[7], Rajaram M. Powar[7], Jean G. Jannin[8], Etienne Pays[2]

[1]*Institut de Recherche pour le Développement, UR177, Instituto de Combate e Controlo das Tripanossomiases (ICCT), CP 2657, Luanda, Angola.*

[2]*Laboratory of Molecular Parasitology, IBMM, Université Libre de Bruxelles, 12 rue des Professeurs Jeener et Brachet, B-6041 Gosselies, Belgium.*

[3]*School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK.*

[4]*Institut de Recherche pour le Développement, UR177, LRCT, TA 207/G, Campus International de Baillarguet, 34398 Montpellier Cedex 5, France.*

[5]*Department of Medicine, Government Medical College, Nagpur, India.*

[6]*Department of Microbiology, Government Medical College, Nagpur, India.*

[7]*Directorate of Health Services, Maharashtra State, Mumbay, India.*

[8]*World Health Organization, Innovative and Intensified Disease Management, Neglected Tropical Diseases Control, Communicable Diseases, Geneva, Switzerland.*

The first case of human trypanosomiasis in Asia was evidenced in 2004 in Seoni, India. A farmer presented a fluctuating trypanosome parasitaemia associated with febrile episodes for five months. Clinical and biological examinations, including parasitological, serological and molecular biological tests confirmed the infecting species as *T. evansi* (Joshi et al., 2005). Suramin was efficient to cure the patient (Joshi et al., 2006). The *SRA* gene (Xong et al., 1998) was not detected by PCR in the *T. evansi* DNA, which had homogeneous kDNA minicircles of Type A (Truc et al., 2007). The parasite appeared to be a typical *T. evansi*, suggesting that the explanation for this unusual infection may lie with the patient. Normally humans exhibit innate immunity against *T. evansi* and *T. brucei*, and in the latter case this immunity is known to involve apolipoprotein L-I (apoL-I, Vanhamme et al., 2003). The serum of the patient

was found to be devoid of trypanolytic activity, and this was linked to the absence of apoL-I due to frameshift mutations in both *apoL-I* alleles (unpublished data). Therefore, the lack of apoL-I was sufficient to explain the human infection by *T. evansi*. Because of a mechanical transmission by insects was suspected, a serological investigation was conducted in the patient village in 2005. Out of 1806 individuals tested using the Card Agglutination Test for Trypanosomiasis/*Trypanosoma evansi* (Pathak et al., 1997), no trypanosomes were detected in the blood of 60 people who were positive at a significant serum dilution (1:4). The results indicate a frequent exposure of the human population to *T. evansi* in the study area, suggesting frequent vector transmission of parasites to humans (Shegokar et al., 2006). Further investigations are required to evaluate the importance of this phenomenon and the potential emergence of a new zoonotic disease (Brun, 2005).

Keywords: *Trypanosoma evansi*; Human trypanosomiasis; India; SRA gene; CATT/*Trypanosoma evansi*; Genetic markers; apoL1

References

Brun, R., 2005. Human Asian trypanosomiasis, a new threat to human health? Am. J. Trop. Med. Hyg. 73, 484.

Joshi, P.P., Chaudhari, A., Shegokar, V.R., Powar, R.M., Dani, V.S., Somalwar, A.M., Jannin, J., Truc, P., 2006. Treatment and follow-up of the first case of human trypanosomiasis caused by *Trypanosoma evansi* in India. Trans. R. Soc. Trop. Med. Hyg. 100 (10), 989–991.

Joshi, P.P., Shegokar, V.R., Powar, R.M., Herder, S., Katti, R., Salkar, H.R., Dani, V.S., Bhargava, A., Jannin, J., Truc, P., 2005. Human Trypanosomiasis caused by *Trypanosoma evansi* in India: the first case report. Am. J. Trop. Med. Hyg. 73, 491–495.

Pathak, K.M.L., Singh, Y., Van Meirvenne, N., Kapoor, M., 1997. Evaluation of various diagnostic techniques for *Trypanosoma evansi* infections in naturally infected camels. Vet. Parasitol. 69, 49–54.

Shegokar, V.R., Powar, R.M., Joshi, P.P., Bhargava, A., Dani, V.S., Katti, R., Zare, V.R., Khanande, V.D., Jannin, J., Truc, P., 2006. Human trypanosomiasis caused by *Trypanosoma evansi* in India: a preliminary serological survey of the local population in the patient's village. Am. J. Trop. Med. Hyg. 75, 869–870

Truc, P., Gibson, W.C., Herder, S., 2007. Genetic characterization of *Trypanosoma evansi* isolated from a patient in India. Infection, Genetics and Evolution 7 (2), 305–307.

Vanhamme, L., Paturiaux-Hanocq, F., Poelvoorde, P., Nolan, D.P., Lins, L., Van Den Abbeele, J., Pays, A., Tebabi, P., Van Xong, H., Jacquet, A., Moguilevsky, N., Dieu, M., Kane, J.P., De Baetselier, P., Brasseur, R., Pays, E., 2003. Apolipoprotein L-I is the trypanosome lytic factor of human serum. Nature 422, 83–87.

Xong, H.V., Vanhamme, L., Chamekh, M., Chimfwembe, C.E., Van Den Abbeele, J., Pays, A., Van Meirvenne, N., Hamers, R., De Baetselier, P., Pays, E., 1998. A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense*. Cell 95 (6), 839–1346.

Symposium "Medical entomology 2"

(61) Structuration of tsetse (Diptera: Glossinidae) metapopulations according to landscape fragmentation in Burkina Faso

Jeremy Bouyer[1,2], Sophie Ravel[3], Laurence Vial[1], Sophie Thévenon[1], Jean-Pierre Dujardin[4], Laure Guerrini[1,2], Issa Sidibé[2], Philippe Solano[2,3]

[1]*Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Département d'élevage et de Médecine Vétérinaire (CIRAD-EMVT), BP 5035, 34032 Montpellier, France.*

[2]*Centre International de Recherche-Développement sur l'élevage en Zone Subhumide (CIRDES), BP 454, Bobo-Dioulasso, Burkina Faso, France.*

[3]*IRD UR 177, Laboratoire de Recherche et de Coordination sur les Trypanosomoses IRD/CIRAD, Campus de Baillarguet, 34398 Montpellier Cedex 5, France.*

[4]*GEMI, UMR CNRS IRD 2724, Centre IRD Montpellier, 911 avenue Agropolis, BP 64501, 34394 Montpellier Cedex 5, France.*

The impact of landscape fragmentation due to human and climatic pressures on the structure of a metapopulation of *Glossina palpalis gambiensis* (Diptera: Glossinidae) was analysed in the Mouhoun river basin, Burkina Faso. Allele frequencies at eight microsatellite loci, and morphometric features based on 11 wing landmarks, were compared among four populations. The populations originated from the Mouhoun River and one of its tributaries. The among-populations distances were 74, 61 and 81 km upstream to downstream, totalizing 216 km between the first and the fourth. Both microsatellites and wing geometry demonstrated a structuration between the populations, but no isolation. There was no clear relation between gene flow and geographic distance. Nevertheless, the type of gallery forest and particularly their disturbance level assessed using phytosociological censes, seemed to be of tricking importance. The impact of the fragmentation of peri-riverine landscapes on tsetse metapopulations structure and its potential implications for control campaigns is discussed.

(62) The potential use of temperate chimeric phages in arthropod vectored disease control

Robert Butcher[1], Timothy L. Karr[2], Stephen F. Hubbard[3], Buntika Areekul-Butcher[4]

[1]*National Nanotechnology Centre (NANOTEC), National Science and Technology Development Agency (NSTDA), 111 Thailand Science Park, Paholyothin Road, Klong 1 Klong Luang, Phathumthani 12120, Thailand.*

[2]*Sex and Symbiosis Laboratory, Department of Biochemistry, Bath University, Claverton Down, Bath BA2 7PY, United Kingdom.*

[3]*Division of Environmental and Applied Biology, College of Life Sciences, University of Dundee, Dundee DD1 4HN, Tayside, Scotland, United Kingdom.*

[4]*Department of Biology, Faculty of Science, Chulalongkorn University, Phayathai Road, Paphumwan, Bangkok 10330, Thailand.*

Transgenic arthropod technology potentially suffers from recombination mediated loss or alteration of the transgene, especially if virus derived, and typically provides a single shot gene drive system. Paratransformation systems may theoretically overcome recombination but are both hard to stably manipulate and to get the correct transgene expression profile and secretion out of the bacteria. Likewise cytoplasmic incompatibility inducing isolates of the arthropod symbionts *Wolbachia*, *Cardinium* and *Rickettsia* as potential transgene drive candidates may provide multiple drive opportunities but are hard to transfect. However, at least some of the temperate phage isolates of *Wolbachia* reveal a chimeric nature at both the DNA sequence level and in the ability to form virions and tails. The recent development of host-specific replication or infection incompetent phage strains, cell free assembly systems, and evaluation of the control of lysogenic insertion and lytic switch mechanisms, has enabled considerable targeted manipulation of this system. Thus cytoplasmic incompatibility inducing *Wolbachia* isolates act not only as ''factories'' for the *in situ* production of chimeric phage within the target arthropod vector cells, but provide the gene drive system leading to rapid near fixation of the infected phenotype in the arthropod population which can be replaced by other CI types, providing the ability to replace or remove transgenic phages from the population as required. We outline (i) the development of chimeric phages that encode for defective interfering particles to reduce vector competency of arthropod vectored arboviruses or arthropod pathogens and illustrate current progress using dengue and Japanese encephalitis viruses within mosquito cell lines and mid-gut tissue culture; (ii) the adaptation of this system to control ssRNA viral pathogens of commercial shrimps and honey bees and (iii) the engineering of key genes from ichnoviruses (AgIcVak4/5) into *CI-Wolbachia* for paratransformation of insect parasitoids and their lepidopteran pests to enhance integrated pest management schemes through altered host immunity to parasitoids.

(63) *Ornithodoros savignyi* as a vector of AHFV—ecological, molecular and evolutionary implications

Shamsudeen Fagbo

Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria, Private Bag X04, Onderstepoort 0110, Pretoria.

We recently isolated Alkhurma Hemorrhagic Fever virus (genus *Flavivirus*, AHFV), a variant of Kyasanur forest disease virus (KFDV), from *Ornithodoros savignyi* (Charrel, Fagbo, Sarah and de Lamballerie; accepted manuscript, Emerging Infectious Diseases) Manuscript) collected in Saudi Arabia; this was followed by molecular characterization and phylogenetic comparison. AHFV is the first human viral pathogen to be isolated from *O. savignyi*, a cryptic and multiple-host seeking Argasid tick endemic is the Middle East and Africa. AHFV and KFDV are the only tick borne haemorrhagic fever viruses known to be associated with *Ornithodoros* spp vectors. AHFV was the first tick-borne hemorrhagic fever-inducing flavivirus for which the complete genome sequence was determined (Charrel, Zaki, Attoui, Fakeeh, Billoir, Yousef, de Chesse, De Micco, Gould, de Lamballerie, Biochem. Biophys. Res. Commun., 2001). Together with our tick isolate, the accumulated molecular data provides confirmation that AHFV is tick-borne and updates the nascent AHFV literature. It clearly questions, in the absence of verifiable data, recent assertions that it is a mosquito borne flavivirus. It also provides a rethinking of previously help notions on the common ancestor of both viruses. Our findings may be integrated with previous work done on the evolutionary and adaptive behaviour of O. *savignyi* in Africa vis-a-vis AHFV vectoring. Here, we discuss the ensuing and interrelated ecological, molecular and evolutionary implications in the Arabian Gulf and beyond.

(64) Molecular cloning and sequencing analysis of Bm91 (angiotensin converting enzymes) cDNA from salivary glands of Thai Cattle ticks, *Boophilus microplus*

P. Kaewhom[1], T. Sirinarumitr[2], R.W. Stich[3], S. Jittapalapong[4]

[1]*Center for Agricultural Biotechnology, CAB, Kasetsart University, Thailand.*

[2]*Department of Pathology, Faculty of Veterinary Medicines, Kasetsart University, Thailand.*

[3]*University of Missouri, Columbia, MO, USA.*

[4]*Department of Parasitology, Faculty of Veterinary Medicines, Kasetsart University, Thailand.*

Cattle ticks, *Boophilus microplus*, are the most important ectoparasites of livestock in Thailand, and are responsible for severe economic losses through direct effects of feeding and as vectors of pathogens. The feasibility of an anti-tick vaccine against *B. microplus* has recently been successful. Excluding Bm86 and Bm95, other potential candidate proteins were searched and developed. A protein, Bm91, has regions of amino acid sequence similarity to angiotensin converting enzymes. Bm91 has also been found to be effective as the candidate antigen on its own for the anti-tick infestation. The objective of this study was to clone and sequence cDNA encoding Bm91 from *B. microplus* indigenous to Thailand. mRNA was isolated from salivary glands, and cDNA encoding Bm91 was amplified with PCR, cloned into the pPICZα vector and transformed into *Escherichia coli* DH5-α competent cells. Purified plasmid DNA was sequenced with dye terminator cycle sequencing reactions, and Bm91 nucleotide and deduced amino acid were analyzed. Nucleotide sequence analysis showed open reading frames of 1893 bps encoding proteins of 631 amino acids. By using the NCBI-Blast conserved domain search for similar domain. The result showed similarities to peptidase_M2, the angiotensin-converting enzyme. The predicted N-glycosylation sites by NetNGlyc 1.0 server was found at amino acid

positions 51, 58, 88, 113, 207, 306, 313, 475, and 628. Comparison of both nucleotide sequence and deduced amino acids showed high identity to an angiotensin-converting enzyme-like protein precursor of *B. microplus* registered in GenBank (AC no. U62809), resulting in 96% identity. To our knowledge this work represents the first report of Bm91 sequence analysis from an Asian strain of *B. microplus*. Bm91 divergence among Thai and other *B. microplus* strains suggests that further work is warranted to determine if a geographic strain-specific vaccine would be more effective in Thailand.

Keywords: Bm91; Angiotensin converting enzymes; Salivary glands; *Boophilus microplus*

Abstracts poster sessions (by alphabetical order of first author)

Proteome analysis by mass spectrometry

Abedian Zeinab[1], Abedian Farshideh[2]

[1]*Pasteur Institute of Iran, 69 Av. Pasteur, 13164 Tehran, Iran.*

[2]*Immunology Department, Mazandaran University, Iran.*

Proteomics is an experimental approach to understand the message present in genomic sequences by analysis of proteins expressed in cell or tissue.

Objective: The first requirement for proteome analysis is the separation and analysis of proteins from organism or cell by electrophoresis. In two-dimensional gel electrophoresis (2D-E). Isoelecteric focussing (IEF), separates proteins according to their isoelecteric point (p*I*), the second step, SDS-PAGE separates proteins according to their molecular weights. Mass spectrometry has become an important and powerful tool for large-scale protein and polypeptide analysis and as the niche methodology in the emerging field of proteomics.

Methods: In this investigation, lysates were prepared from a B lymphoma cell line (A20) and the proteins were resolved across a pI range of 3–10 using immobiline DryStrip gel. This was followed by the second dimension, on a 12% polyacrylamide gel. Gel protein digestion was carried out by trypsin and sample became ready for electrospray (ESI) mass spectrometric analysis. ESI coupled on-line with high-pressure liquid chromatography (HPLC), shortened the time to solve the primary structure of proteins and peptides (LC–MS). ESI produces multiply charged species extending its capability to analyze a mass of over 100,000 Da. The ionization process took place in atmosphere and the charged species were transfered into mass spectrometer with high efficiency for analysis.

Results: Mass spectrum was formed by charged species for B cell proteins. Charges and masses were indicated on the top of the respective peaks. They were Max.75.9 counts. The highest molecular weight was 198.5977 kDa with pI 4.98 and the lowest was 16.4186 kDa with pI 9.59.

Conclusion: Mass spectrometry is an indensable technique for proteomics. This facilitates the study of all protein complexes and organelles that can be purified.

Monitoring of influenza A viruses in synanthropic birds in South of Western Siberia in H5N1 epizootic and postepizootic period

A.Yu. Alekseev, K.A. Sharshov, S.I. Zolotykh, I.G. Drozdov, A.M. Shestopalov

FSRI State Research Center of Virology and Biotechnology "Vector", Koltsovo, Novosibirsk region, Russia.

Wild birds are a natural reservoir of influenza viruses from which influenza can penetrate in population of domestic birds. Migrations of wild birds promote geographical spreading of various strains of influenza A viruses. It is necessary to monitor the circulation of influenza viruses among wild birds in regions located on birds' migration ways. The special importance should be given to monitoring of influenza viruses among synanthropic birds because of their close contacts with man and domestic animals on the one hand, and with wild biocenoses on the other. The susceptibility of synanthropic birds to high pathogenic avian influenza (HPAI) virus was shown in many works. Since September 2005 till March 2006 we have collected samples from synanthropic birds in Omsk and Novosibirsk region (South of Western Siberia) to investigate the influenza virus infection carrier state. We have investigated 458 samples from 10 bird species. As a result of screening research in chicken embryos we have isolated five viruses that showed gemmaglutination activity. Serological typing of isolates was carried out with the help of serums kindly given by Dr B. Webster (Memphis, USA). Isolate of H4 serotype was isolated only from one bird (from *Pica pica*) during the epizootic period, that is equal to 0.2%. All the other four isolates belonged to a New Castle virus. Thus, we have not registered high pathogenic avian influenza viruses in synanthropic birds during epizootic and postepizootic periods. Probably, it accounts for weak susceptibility of synanthropic birds or small amount of samples.

The problems of property preservation of museum and collection strains of mycobacteria of the tuberculosis complex at long-term cultivation

A.Yu. Alekseev, A.G. Durymanov, A.M. Shestopalov

FSRI State Research Center of Virology and Biotechnology "Vector", Koltsovo Novosibirsk region, Russia.

Cultivating microorganisms on an artificial medium, we recover it from the process of interaction with macroorganism and start the process of changes in genotype stability. The recent data are indicates on property changes of many bacterial pathogens after two to three passages while cultivation on artificial media.

We have developed a new method of cultivation of pathogenic mycobacteria on a basis of continuous cell culture. The investigated strains of pathogenic mycobacteria are passaged with the help of two methods (on Levenstein–Yensen medium and on DM Eagle's MEM growth medium with *Vero* cell culture). Sensitive laboratory animals (guinea pigs) were used for pathogenicity studies. The guinea pigs

were euthanasied on day 80 post-infection. Estimation of pathological changes in internal organs of guinea pigs showed that pathologic changes started eliminating with third passage while cultivating on Levenstein–Yensen medium. Evaluation of internal organs of guinea pigs infected with *M. tuberculosis* grown on DM Eagle's MEM Vero medium showed that pathological changes are observed at least up to passage 13. The examination of internal organs of guinea pigs infected with M. tuberculosis grown on DM Eagle's MEM-Vero medium showed that pathological changes were high and constant.

The following conclusions can be drawn from the obtained results: the virulence of mycobacterium of the tuberculosis complex decreases while cultivation on Levenstein–Yensen medium. Pathological pictures differ in animals infected with mycobacteria cultivated using a continuous cell culture and in those infected with mycobacteria grown on Levenstein–Yensen medium. We have developed a new method of cultivation that allows to remain the most important biological properties unchanged at long-term cultivation on a continuous cell culture.

Identification of genetic diversity within *Brugia* species in feline based on internal transcribed spacer regions

Supatra Areekit, Arda Pakpitchareon, Sunsanee Phasook, Sintawee Khoochareontaworn, Phaisarn Khawsak, Kosum Chansiri

Department of Biochemistry, Faculty of Medicine, Srinakharinwirot University, Bangkok 10230, Thailand.

It has been reported that *Brugia malayi* have infected not only human but also animals such as cats, monkeys, and dogs, whereas *B. pahangi* causes morbidity only in cat reserviors. Due to their similarity in morphologies and others, the identification of both brugian species in such carriers based on traditional detection tools can be difficult and mostly lead to misdiagnosis. Hence, the data based on genes is the alternative useful information for not only parasite identification but also their genetic diversities. The internal transcribed spacer (ITS) regions were used to determine the genetic diversity of *Brugia* spp. within domestic cat reservoirs from different geographical areas in Thailand. Microfilaria was separated and their DNA was extracted prior to PCR amplification. The specific primers of ITS1 and ITS2 regions were used to yield the PCR products of 580 bp and 660 bp in size, respectively. The fragments were cloned, sequenced, and aligned in comparison to the reported data of *B. malayi* and *W. bancrofti*. It was found that ITS1 and ITS2 phylogenetic trees demonstrated the genetic variation among *Brugia* spp. Phylogenetic trees based on Neighbor Joining (NJ) and DNA Parsimony (DNA PARS) revealed both single infection of either *B. malayi* (cats 1, 3 and 4) or *B. pahangi* (cats 6 and 7) and mix infection of both *Brugia* spp. (cats 2 and 5). It can be proposed that ITS regions could be used for studying genetic diversity of *Brugia* spp., especially, in cat reservoirs which will be beneficial for epidemiological survey.

Spatial approach of the production of *Aedes aegypti* pupae using GIS and remote sensing

Philippe Barbazan[1], Wasana Bonyuan[1], Florent Demoraes[1], Marc Souris[1], Walairut Tuntaprasart[2], Jean-Paul Gonzalez[1]

[1]*IRD-Mahidol University, CVVD-UR 178, Project, Emerging Viral Diseases and Information System Research, Faculty of Sciences, Mahidol University, Thailand.*

[2]*Faculty of Public Health, Mahidol University, Thailand.*

DHF is a permanent challenge for Public Health authorities in Thailand, as epidemics in 1997–1998 and 2001, spread over most of the country. Wide variations of level of incidence over areas mean that to be efficient the control strategy needs the delineation of risk areas. Classical entomological indices are used by public health authorities to launch local vector control activities but their reliability to identify areas with higher incidence and to reduce it, is limited. In the frame of a WHO-TDR program to develop new entomological indices based on pupae counts, an exhaustive survey of potential breeding sites has been done in areas with different types of urbanization in Thailand. A GIS has been developed, using the precise localization (GPS) of houses as a basic layer. The characterization of the most productive breeding sites in terms of pupae, the density of human population and socio economical indicators, such as the field description of the type of dwellings (unmanaged urban environment, town houses, residential and administrative areas, villages) were additional layers of information. Most productive BS were similar in the different areas. The containers for water storage produced up to 90% of the pupae which density could reach 0.1–2.6 pupae per person. The correlation between the number of potential BS and the number of pupae is higher (0.9) if we consider groups of neighboring houses (density of attributes). A minimal threshold was defined under which stochastic process in BS colonization may lead to an interruption in pupae production. Spatial patterns in the distribution of pupae allows to identify areas were targeted vector control should be easier and more efficient. This method, combining field survey for the characterization of productive breeding sites and GIS technology to delineate areas with a specific type of urbanization, will help to identify similar environments likely to evolve simultaneously in response to the emergence of epidemic phenomena. Control strategies can therefore target the most productive containers but also key areas in the transmission network, for a better efficiency.

The RNA virus database

Robert Belshaw[1], Tulio de Oliveira[1], Andrew Rambaut[2]

[1]*Department of Zoology, University of Oxford, United Kingdom.*

[2]*Institute of Evolutionary Biology, University of Edinburgh, United Kingdom.*

As concern mounts over the threat posed by RNA viruses, and the amount of genomic information we have for them increases, a single web application providing key genomic resources for all species is timely. We have created the RNA Virus Database to perform six main services for all 700 RNA viral species:

Identify submitted viral nucleotide or amino acid sequences, provide curated whole genome nucleotide alignments – with corresponding phylogenetic trees – for each virus, Align submitted nucleotide sequences to the above, provide amino acid sequences for all viral genes, and allow the user to extract the corresponding region from the above whole genome alignments, provide whole translated genomes for each viral species, show links to the more specialised web sites for the viruses of greatest medical importance (including genotyping tools). We also link to other sites providing further taxonomic or biological information for each virus. We are currently using the database to analyse the deep phylogeny of RNA viruses and the relationship between genome size and genome architecture. We expect the website version of the database to facilitate and encourage research into many other aspects of RNA viral evolution. It is freely accessible at http://virus.zoo.ox.ac.uk/.

Genetic diversity in clinical and environmental samples of *Legionella pneumophila*

Mireia Coscollá[1], Isabel escribano[2], Daniel Muñuzuri[3], Juan Bellido[4], Fernando González-Candelas[1]

[1]*Instituto Cavanilles de Biodiversidad y Biología Evolutiva, University of Valencia, Spain.*

[2]*Hospital Virgen de los Lirios, Spain.*

[3]*Centro de salud Pública de Benidorm, Spain.*

[4]*Centro de salud Pública de Castellón, Spain.*

Legionella pneumophila are waterborne bacteria responsible for legionellosis, an emerging disease causing respiratory illness when a susceptible human host inhales contaminated aerosolized water. Molecular epidemiology and diversity studies of *L. pneumophila* usually compare PFGE and/or AFLP patterns from bacterial cultures derived from respiratory and environmental samples. Recently, a sequence based typing scheme based on six loci has been developed for this species, which is also applied on cultured isolates. But culture of this species on selective media, despite its high specificity, presents low sensitivity (only 5–15% of the samples produce cultured isolates). In order to improve the efficiency of sequence based analysis of *L. pneumophila*, we have developed a protocol to amplify and sequence DNA extracted from uncultured respiratory samples and we have used it to sequence and compare three intergenic regions and internal fragments of six genes of the *L. pneumophila* clinical and environmental origin. Sequences of these nine markers were derived from 40 environmental samples and 40 clinical samples taken from different years and were used to study the genetic diversity and population structure of *Legionella* strains in a Spanish region that includes an area where legionellosis has become almost endemic, with continuous bouts of sporadic cases and several outbreaks affecting tens even hundreds of people. We have also studied the phylogenetic relationships among all these 80 isolates. Phylogenetic analyses have revealed that isolates recovered from patients in years 1999, 2004 and 2005 from this area were almost identical. Moreover, genetic differentiation between clinical isolates from this area and the remaining clinical strains was detected. Statistics measuring genetic differentiation between clinical and environmental data sets also showed a significant differentiation for all 9 markers, both when analyzed independently and when combined. These results are suggestive of an unexpected differentiation between clinical and environmental isolates of *L. pneumophila*.

Borrelia afzelii gene expression in *Ixodes ricinus* ticks

M. Derdáková[2,3], J. Koči[1], K. Peterková[1], M. Kazimírová[3], D. Selyemová[3], M. Labuda[3]

[1]*Institute of Virology, Slovak Academy of Sciences, Dúbravska Cesta 9, Bratislava 84245, Slovakia.*

[2]*Parasitological Institute, Slovak Academy of Sciences, Hlinkova 3, Košice 04001, Slovakia.*

[3]*Institute of Zoology, Slovak Academy of Sciences, Dúbravska cesta 9, Bratislava 84506, Slovakia.*

Spirochetes belonging to the *Borrelia burgdorferi* sensu lato (s.l.) complex have evolved remarkable ability to survive in diverse ecological niches during transmission cycles between ticks and vertebrate hosts by variable gene expression. To understand the events during spirochaete transmission from feeding ticks to hosts, mRNA levels of selected *B. afzelii* genes (*bbk32*, *dbpA*, *ospA*, *ospC* and *vlsE*) were measured by quantitative *real-time* Sybr Green PCR. *B. afzelii* infected *Ixodes ricinus* nymphs fed on laboratory Balb/c mice for 0, 24, 48, and 72 h. The mRNA levels of the constantly expressed *flagellin* gene were used for the relative quantification of selected genes. Differences in gene expression profiles were observed in unfed ticks and during tick feeding. mRNA levels of *bbk32* and *dbpA* showed distinctive decreasing patterns during the first 24 h post-attachment, while *OspC* and *vlsE* mRNA levels increased significantly during the feeding process. In contrast, *ospA* levels decreased for the 48 h of tick feeding and slightly increased by 72 h. More detailed and comprehensive studies on regulation of gene expression in different borrelia genospecies on the vector–host interface would aid to develop effective strategies in preventing pathogen transmission.

Keywords: *Borrelia burgdorferi* sensu lato; *Ixodes ricinus*; Pathogen transmission; Gene expression; *bbk32*; *dbpA*; *ospA*; *ospC*; *vlsE*

First report on lectin-related gene in Iranian main malaria vector *Anopheles stephensi*

Farnoosh Farhoomand[1,2], Saber Gholizadeh[1], Sedigheh Zakeri[1], Navid Dinparast Djadid[1]

[1]*Malaria Research Group (MRG), Biotechnology Department, Pasteur Institute of Iran, P.O. Box 1316943551, Pasteur Avenue, Tehran, Iran.*

[2]*Biology Department, Science and Research Branch, Islamic Azad University, Iran.*

Malaria is the major vector-borne infectious disease. Nearly 40% of the world's population is at risk, and close to 2 million persons (mostly children under the age of five) die every year. Many insects posses lectins with distinct sugar specificities

which play important recognition and protective roles in immune defense against microbial pathogens and parasitic protozoans. Interactions between parasites and vector gut walls may be mediated by the carbohydrates on the surface of parasites and lectin in the vector gut. Fibrinogen-related proteins are a kind of lectin, so in this study we used primers that flanking a sequence of fibrinogen gene based on *Anopheles gambiae* genome. *Anopheles stephensi* specimens were collected from Sistan and Baluchistan Province, followed by DNA extraction, amplification of a 380 bp fragment and sequencing by using FBN9 primers. Two samples originated from Sarbaz and Nikshahr districts had 100% similarity with each other and a GC count of 56.84%. However, its nucleotide similarity with this fragment in *Anopheles gambiae* is about 90%, while amino acid sequences had 92% similarity in these two main vector species of malaria in Africa and Asia. In order to study the role of these genes in interaction between *Plasmodium* and *Anopheles* and to find candidate molecules in designing transmission blocking vaccine, it is necessary to do further research to obtain more information on the structure of lectin related genes in this important but neglected *Anopheles* vectors.

Keywords: *Anopheles stephensi*; Lectin; Fibrinogen; Iran

Analysis of genetic heterogeneity of POL variants in the HIV-1 circulating in Nivisibirsk (RUSSIA)

N.M. Gashnikova[1], O.O. Sevastyanova[1], V.A. Ternovoy[1], V.G. Orlovsky[2], A.G. Pokrovsky[1]

[1]*Federal State Research Institution State Research Center of Virology and Biotechnology "Vector" Rospotrebnadzor, Koltsovo, Novosibirsk Region, Russia.*

[2]*Novosibirsk AIDS Center, Russia.*

Objective: The goal of the work was to assess the heterogeneity of the regions encoding reverse transcriptase and protease of the HIV variants circulating in Novosibirsk and to analyze the emergence of the virus variants carrying drug-resistance mutations.

Patients and methods: The peripheral blood samples of HIV-infected patients from the Novosibirsk AIDS Center were used in the work. The patients were divided into two groups—those received antiretrovirus therapy (ART) and not received it. HIV-1 RNA was isolated from blood sera. DNA nucleotide sequences were determined in a Beckman CEQ2000XL (Beckman Coulter, USA) automated sequencer. Mutations were assayed for their connection with drug-resistance using the Stanford University HIV Database Drug Resistance.

Results: Overall, the fragments for the regions of HIV-1 reverse transcriptase were obtained and analyzed for 84 HIV-1 variants; of protease, for 54 variants. Phylogenetic analysis of the obtained reverse transcriptase sequences demonstrated that the HIV variants cluster into two main clades, one belonging to HIV-1 subtype B and the other, to subtype A. The mutations belonging to the secondary resistance mutations described for HIV-1 subtype B were detectable in all the samples. The mutations connected with the drug resistance to protease inhibitors were recorded in the following codons: I13V (40% of the samples studied), H69K (70%), M36I (100%), V77I (50%), and I93L (100%). The detected occurrence rate of mutation A62V in reverse transcriptase gene amounted to 81%. Analysis of the samples for the presence of primary mutations rendering HIV-1 resistant demonstrated that the HIV-1 variants isolated from the not received ART patients had no primary resistance mutations. The mutations determining resistance to a wide range of drugs were found in 78% HIV variants from the ART patients. In the studied HIV variants isolated from the patients receiving reverse transcriptase inhibitors, the most frequently met mutations occur at codons 215, 184, 41, and 210 in the case of administration of nucleoside inhibitors and at codons 103, 181, 184, 188, and 190 in the case of non-nucleoside inhibitors. In the case of treatment with protease inhibitors, the resistance mutations emerge at a considerably lower rate.

Conclusion: As no resistant variants were isolated from the patients that did not receive ART, we may assume that the resistant variants had not distributed in Novosibirsk. The HIV variants isolated from ART patient treated for a period over 1 year developed multiple drug resistance. This work demonstrates that it is of the utmost importance to monitor the HIV-1 drug resistance in the clinical practice. Administration of the anti-retrovirus therapy in our country is only at its very beginning; therefore, emergence of HIV-1 drug resistant variants is possible in the nearest future.

Keywords: HIV-1; Reverse transcriptase; Protease; Genetic heterogeneity; Drug-resistance mutations

The association of host HLA type and the sequence variation of *Gag* gene of HIV-1 CRF_01AE subtype

Goragoch Gesprasert[1], Koya Ariyoshi[2], Patraraporn Oratai-wan[1], Nuanjuan Wichukchinda[3], Wattana Auwanit[3], Prasert Auewarakul[4], Panita Pathipvanich[5], Arunee Thitithanyanont[1]

[1]*Department of Microbiology, Faculty of Science, Mahidol University, Thailand.*

[2]*AIDS Research Center, National Institute of Infectious Diseases, Tokyo, Japan.*

[3]*Thai National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Nonthaburi, Thailand.*

[4]*Department of Microbiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Thailand.*

[5]*Day Care Center, Lampang Hospital, Lampang, Thailand.*

Human immunodeficiency virus or HIV is one of the most well-known viruses. As other viruses, the survivability of HIV is hanged in the balance with two arms. First side is the selection pressures, such as an immune response and anti-retroviral drug, and another is the capability of the virus for escaping from these selection pressures. Because HIV has an extraordinary of genetic diversity and the ability that can infect and replicate in the cellular component of the immune system, these seem to make HIV undefeatable. However, the immune pressure also leaves the scare, as a traceable history, on the viral genome. To determine the effect of host immune response, we analyzed the viral genetic diversity of 370 HIV-1 CRF_01AE Gag sequences obtained from 116 HIV-1 infected couples, and examined the association between host HLA and

the sequence polymorphism. We found that the rate of sequence variability was depended on the region of *Gag* gene. The part of *Gag* gene that had highest mean sequence variability was p2 spacer region followed by p6, matrix, p1, nucleocapsid, and the most conserved capsid region, respectively. We had identified 52 variable sites that were associated with 12 host HLA alleles. Most associations were observed in the positions that had high polymorphism. The variable sites that associated with the same HLA were clustered in the specific region of Gag, especially HLA-A. HLA-A*02, A*11, and A*24 associated variable sites were clustered in the matrix protein, whereas the variable sites that associated with HLA-B were distributed in all parts of Gag protein. Fifteen variable sites located within HLA-specific CTL epitopes. Although most of variable sites did not locate within HLA-specific CTL epitopes, we found that some of them were correlated with the variable sites that located within HLA-specific CTL epitopes. This correlation might be a compensatory mechanism of the escape mutation. Reversion of the mutated to wild type amino acid was commonly found in many variable sites, when the virus was transmitted from the HLA-presented patient to their wives or husbands that did not have that HLA. The rate of reversion depended on the strength of the association. Our result showed that the HLA-restricted immune response and the reversible escape mutation were the basic mechanism that caused the sequence variation of HIV-1.

Molecular identification of *Anopheles superpictus* (Diptera: Culicidae) complex from Iran and Pakistan

Saber Gholizadeh[1], Sedigheh Zakeri[1], Seyda Azra Qamar[1,2], Navid Dinparast Djadid[1]

[1]*Malaria Research Group (MRG), Biotechnology Department, Pasteur Institute of Iran, P.O. Box 1316943551, Pasteur Avenue, Tehran, Iran.*

[2]*Zoology Department, Karachi University, Pakistan.*

Malaria is one of the most important parasitic and vector-born diseases in southeast provinces of Iran. Eight out of 19 species of Anopheles in Iran incriminated as malaria vectors. *Anopheles superpictus* usually in its distributed regions acts as secondary vector. This species is more prevalent in northern Pakistan and Afghanistan, through the Caucasus into southern Tadzikhstan and middle Asia, and through the Middle East and Asia Minor to Southeastern Europe. In this study, *An. superpictus* specimens collected from Sistan & Baluchistan, Ilam and Khorassan provinces in Iran and Karachi in Pakistan, followed by morphological identification, DNA extraction and PCR amplification of ITS2 region. Total size of amplified fragment in sequenced specimens was 525 bp with ITS2 region of 352–373 bp and GC count of 53–56.%.This is the first world report on identification of ITS2 region in *An. superpictus* as a complex species including two suspected species A and B from Iran and Pakistan. Detailed analysis and differences in these two suspected species will be discussed.

Keywords: *Anopheles superpictus*; Suspected species A and B; ITS2; Iran; Pakistan

Polymorphism in the P-glycoprotein (Pgp) gene from *O. volvulus* Mexican isolates

S. Gonzlalez Guzman, A. Monroy-Ostria, A. Ramírez Ramírez

Departamento de Inmunología, Escuela Nacional de Ciencias Biológicas, IPN, Carpio y Plan de Ayala, Col. Casco de Santo Tomás, C.P. 11340, México, D.F., Mexica.

Onchocerciasis is caused by the filarial parasite *Onchocerca vovulus*, it is transmtited by *Simulium* spp. It is the second infectious disease that leds blindness in the world. Ivermectine is the drug administrated to the population in risk. Nematodes of veterinary importance under pressure with ivermectine develop resistance to this drug. Simple strand and conformational polymorphism analysis (SSCP) and PCR-RFLP of the associated multidrug-resistance P-glycoprotein (Pgp) gene, from *O. volvulus* Mexican isolates from Oaxaca and Chiapas foci were done. By SSCP was found genetic variation in three out of six samples from Chiapas, but there was not variation in samples from Oaxaca. The pattern of the PCR-RFLP of Pgp were analyzed by using the NTSYS-PC program and the corresponding phenogram was built with the UPGMA method. The corresponding phenogram showed three groups, one formed with samples from Oaxaca, other with the samples from Chiapas and the last is formed with samples from both, Oaxaca and Chiapas. The role of the geography and idiosyncrasy of people living in Chiapas and Oaxaca foci is high important in variation. Oaxaca has no migration and excellent control of ivermectine administration, Chiapas state has high illegal migration without a control of dosages of ivermectin and there are not geographical barriers.

Random amplified polymorphic DNA technique for the identification of *Leishmania* species recovered from *P. Papatasi* in north-east of I.R. Iran

H. Hajjaran[1], M. Mohebali[1], M.R. Yaghoobi-Ershadi[2], A.A. Akhavan[2]

[1]*Department Medical Parasitology and Mycology, School of Public Health and Institute of Public Health Research, Tehran University of Medical Sciences, Tehran, Iran.*

[2]*Department Medical Mycology, School of Public Health and Institute of Public Health Research, Tehran University of Medical Sciences, Tehran, Iran.*

Zoonotic cutaneous leishmaniasis (ZCL) is an important endemic disease in Sabzevar district, north-east of Iran. It has been shown that *L. major* is the species encountered and *Rhombomys opimus* is the main animal reservoir host. In this study, promastigotes were isolated from infected *P. papatasi*, inoculated subcutaneously into the base tail of BALB/c mice separately. The isolates were successfully reisolated from the BALB/c and then cultured in Schneiderâ's media supplemented by 10% FBS, subsequently Leishmania specie were determined by RAPD-PCR with four oligonucleotides primers including AB1-O7, A4, 327 and 329.The identification was also confirmed by RFLP-PCR method with nagt gene by Dr. K.P. Chang from Chicago University, USA. Out of five

Leishmania spp. isolated from *P. papatasi*, all of them were identified as *L. major* by above techniques. The results showed that *L. major* is the principal agent of ZCL and *P. papatasi* is the main vector of the disease in studied area in Iran.

What evolutionary ecology tells us about dengue control

Amy Henry, Kira Krend

Department of Zoology, UH Manoa, 2538 McCarthy Mall, Edmondson 152, Honolulu, HI 96822, USA.

Many current vector-borne disease control efforts narrowly focus on mosquito control. Often, these projects lack a broader, evolutionary ecology viewpoint of the disease dynamics. We propose a transdisciplinary approach to analyze disease transmission by evaluating interactions between the vector, pathogen, human and environment within a defined ecosystem. Included in this paper is summary of classic ecological concepts vital to understanding population, community, and ecosystem biology, as well as key ideas regarding the evolutionary ecology of pathogens and disease, with a focus on dengue and its vectors. Mosquito control and avoidance methods are summarized, noting the environments and situations in which they could effectively reduce or eliminate the risk of dengue infection. A protocol was designed to assist in creating an efficient dengue control program including guidelines on how to select a study site and identify the study population and ecosystem, as well as monitoring techniques to evaluate program success.

Development of immunological test kit for detection of porcine reproductive and respiratory syndrome in swine

Surasak Jittavisutthikul[1], Amara Thongpan[3], P. Lertwatcharasarakul[2], O. Booddee[2], W. Chumsing[2], S. Jala[2], Worawidh Wajjwalku[2]

[1]*Center for Agricultural Biotechnology, Kasetsart University, Thailand.*

[2]*Faculty of Veterinary Medicine, Kasetsart University, Thailand.*

[3]*Department of Genetics, Faculty of Science, Kasetsart University, Thailand.*

Two types of virus, European type (EU PRRSV) and North American type (US PRRSV), are known to cause porcine reproductive and respiratory syndrome (PRRS). An enzyme-linked immunosorbent assay (ELISA) for the simultaneous detection of serum antibodies against these two PRRS types was developed. Fusion nucleocapsid proteins of EU PRRSV and US PRRSV (USEU-N protein) expressed in *Escherichia coli* were partial purified and used as antigens. Determined by checkerboard titration optimal condition was obtained using USEU-N protein at 1:1600 and serum at 1:40. The optimal cut-off value for developed USEU PRRS ELISA was found to be 0.4, having sensitivity and specificity at 97.5% and 100%, respectively. Comparison was made with IDEXX® HerdCheck PRRS ELISA using two graph-ROC program testing with 200 positive sera and with 200 negative sera. The degree of agreement (κ value) was highly obtained at 0.7652. The kit is considered reliable for routine diagnostic, epidemiological surveys and outbreak investigations.

Keywords: PRRS; Recombinant protein; ELISA

Characterization of avian influenza virus (H5N1) from Asian open-billed storks in Thailand

Watcharoot Kanchongkittiphon, Goragoch Gesprasert, Patharapan Lersritwimanmaen, Phattaraporn Orataiwun, Suwimon Wiboon-Ut, Arunee Thitithayanont

Department of Microbiology, Faculty of Science, Mahidol University, Bangkok 10400, Thailand.

Emergence of a highly pathogenic avian influenza virus (H5N1) was started in Thailand since late 2003. Asian open-billed stork is one of the waterfowl species infected by H5N1 virus. To investigate whether it could be a reservoir or carrier of this virus, during February 2004–September 2006, 1959 cloacal swabs of this waterfowl from central part of Thailand were collected for H5N1 surveillance. All together, 30 H5N1 viruses were isolated. To characterize the molecular epidemiology, 244 full-length genes were sequenced. All viruses had multiple basic amino acids at the hemagglutinin (HA) cleavage site. Uniquely, 16 neuraminidase (NA) sequences had no glycosylation at position 235 which was associated with the presence of IERRRKKR in the cleavage site of HA. Additionally, all isolates processed at least 1 human specific residue at position 79 of PB1-F2. Phylogenetic analysis was demonstrated that these viruses originated from the Gs/Gd/1/96-like lineage and formed distinct sub-lineages corresponded to Thailand isolates. Our study presents molecular basis and provides insight to understand the evolution of H5N1 viruses as well as assists in planning for pandemic influenza surveillance.

Identification of *Trypanosoma* sp. based on hypervariable region of 18S rDNA

Sintawee Khoochareontaworn, Kosum Chansiri

Department of Biochemistry, Faculty of Medicine, Srinakharinwirot University, Bangkok 10110, Thailand.

The complete nucleotide sequences of 18S rDNA (2100 bp) from 42 isolates of trypanosomes were aligned against those of *Euglena viridis* and *Khawkinea quartana* by using ClustalX program. The DNA parsimony (DNA PAR) phylogenetic tree showed five hypervariable regions (HV) within 18S rDNA sequences. Upon analysis of each region, it was found that DNA PAR tree of the HV region 2 (272 bp) was corresponded to DNA PAR tree of 18S rDNA in classification of *Trypanosoma* sp. The HV regions 1 and 3 (242 bp and 314 bp) the distinction of section Salivaria from other *Trypanosoma* sp. whereas HV regions 4 and 5 (292 bp and 402 bp) could discriminate only section Stercoraria from the others. Based on the analysis, it can be concluded that DNA PAR tree inferring from HV regions of 18S rDNA could represent the tree of 18S rDNA. Thus, this could be applicable for identification and discrimination of *Trypanosoma* sp. which will be useful for epidemiological study and the control of the disease.

Genetic diversity of *Trypanosoma evansi* in buffalo based on internal transcribed spacer (ITS) regions

Sintawee Khuchareontaworn[1], Phirom Singhaphan[1], Nareerat Viseshakul[2], Kosum Chansiri[1]

[1]*Department of Biochemistry, Faculty of Medicine, Srinakharinwirot University, 114 Sukhumvit 23, Klongteoy Wattana, Bangkok 10110, Thailand.*

[2]*Department of Parasitology, Faculty of Veterinary Medicine, Chulalongkorn University, Patumwan Road, Bangkok 10330 Thailand.*

The nucleotide sequences of 18S rDNA and internal transcribed spacer (ITS) regions were used for studying the relationships of *Trypanosoma evansi* isolate from a buffalo. The sequences were analyzed and compared to 18S rDNA and the ITS regions of the other *Trypanosoma* spp. Maximum likelihood phylogenetic trees were constructed using *Leishmania major* as the outgroup. The tree of 18S rDNA indicated that *T. evansi* (buffalo B18) isolate was closely related to those of Taiwan and *T. brucei* stock. The ITS tree showed the genetic diversity among 32 clones of *T. evansi* (B18) within a single host. This data will be useful for epidemiological and dynamic studies for designing the rational control programs of the disease.

Keywords: *Trypanosoma evansi*; Phylogenetic; Nucleotide sequence; Small subunit rDNA; 8SrDNA; Internal transcribed spacers (ITS); Buffalo

Genetic diversity as an indicator to the activity of *Echinococcus multilocularis* on the French region of the Ardennes

Jenny Knapp, Jean-Mathieu Bart, Marie-Hélène Guislain, Francis Raoul and Renaud Piarroux

Echinococcus multilocularis is a cestode responsible for one of the most important zoonosis in the Northern hemisphere, the Alveolar Echinococcosis disease. The parasite involves in its live cycle mainly foxes as definitive hosts (DH) and rodents as intermediate host (IH). Human can accidentally act as IH and develop the disease after a long incubation period (10–15 years). In France, about 15 new cases are recorded each year. The parasite is chiefly present in the East and the Centre of the country. Since 1984, 7 human cases were described in the department of the Ardennes, and then could be considered as an emergent endemic focus. Moreover the prevalence in foxes reached 53%. Due to the long incubation period, the number of human cases could not reflect the current activity of the parasite in this area. The genetic diversity of the parasite could reflect its activity in the Ardennes. In this aim, we have studied two microsatellite DNA targets, the tandem repeat multilocus microsatellite EmsB and the single locus microsatellite NAK1 on a panel of 25 red foxes, presenting different worm burdens. These two targets allowed us to defined 6 main genotypes among 145 adult worms. A fox, presenting the more important worm burden ($n = 73380$ worms), harboured individuals with a heterozygote genotype for the NAK1 locus, suggesting cross-fertilization event. From the microsatellite analysis, we proved to be able to detect genetic variability in restricted areas, and thus to track dynamic transmission of *E. multilocularis* at local scale.

The origin of European and North American infectious salmon anemia virus (ISAV) revisited by relaxed molecular clock analysis

Tsan-Yuk Lam, Chung-Chau Hon, Zhenggang Wang, Fanya Zeng, Frederick Chi-Ching Leung

Department of Zoology, The University of Hong Kong, 5N-01, 5/F North Wing, Kadoorie Biological Science Building, Hong Kong SAR, China.

An emerging infectious disease, called infectious salmon anemia (ISA), among farmed salmon has raised the concern of its impact on salmon-farming industry. The causative agent was identified to be a segmented RNA virus and named as infectious salmon anemia virus (ISAV). Earlier phylogenetic analyses classified the ISAV, predominating in European and North American aquacultural farms separately, into two genetically distinct strains: European (EU) and North American (NA) type. It was suggested that human traffic and trading have transported their common ancestor between these continents, and led to their local emergences. A recent evolutionary study on ISAV hemagglutinin-esterase (HE) and fusion protein (FP) gene suggested, under global molecular clock assumption, atypically low substitution rates (10^{-6} and 10^{-5} substitutions/site/year, respectively) comparing to other RNA viruses. This result has new implications on the viral epidemiology and the hypothesis of origin: EU and NA isolates may have separated long before any human traffic could transport the virus. We examined the genetic sequences of HE gene from both EU and NA strains of ISAV. Our result demonstrated a large disparity of the substitution rates with in the phylogeny, which disrupted the molecular clock assumption, was caused by the presence of a small group of virus isolates with extremely low rate. The presence of these isolates in the phylogeny can mislead the substitution rate estimation in global clock analysis. By using relaxed molecular clock models implemented in the Bayesian framework which accommodating the rate variation, the time of the most recent common ancestor of EU and NA strains was extrapolated to around 300 years ago. This finding reasserts the original hypothesis suggesting the translocation of ancient ISAV between North America and Europe could be a consequence of human cross-ocean traffic and trading, which had been refuted by the recent global clock analysis.

Keywords: Infectious salmon anemia virus; Relaxed clock; Molecular dating; Penalized likelihood

Foot-and-mouth disease virus isolated from pig in Thailand in 2005 related to Cathay Topotype in Vietnam

P. Lertwatcharasarakul, A. Boonsoongnern, O. Booddee, W. Chumsing, W. Wajjwalku

Faculty of Veterinary Medicine, Kasetsart University, Kamphaensaen, Nakhon Pathom 73140, Thailand.

Cathy Topotype of Food and Mouth Disease Virus Moved From Vietnam to Thailand.

Two samples of pigs infected Food and Mouth Disease Virus (FMDV) in the central part of Thailand in 2005 were identified by RT-PCR and nucleotide sequencing. The nucleotide sequence was analyzed of multiple alignment and constructed dendrogram by MEGA3.1 program. The dendrogram data of partial VP1 sequence showed that two viruses isolated in Thailand closely related with the Cathay topotype isolated pig from Vietnam in 2004. Moreover, 3A non-structural protein of both isolations showed ten amino acid deletions (position 93–102). In addition, the virus was not found in cattle in the endemic area. This is the warning system for control the outbreak of FMD in Thailand because this is the first reported of this topotype in this country.

Keywords: FMDV; Cathay topotype; Phylogenetic tree

Identification of *Babesia canis vogeli* from domestic dogs in Nakhon Phathom, Thailand in 2005

P. Lertwatcharasarakul, C. Salakij, C. Singthusingha, P. Arunvipas, G. Nunklang, J. Rattanakunuprakarn, S. Ounsiri, S. Phatthanakunanan, S. Jala W. Wajjwalku

Faculty of Veterinary Medicine, Kasetsart University, Kamphaensaen, Nakhon Pathom 73140, Thailand.

The reverse line blot hybridization (RLB) and a restriction fragment length polymorphism (RFLP) analysis were analyzed *babesia* species in dogs. Eighteen blood samples of dogs were found to be babesia infections by microscopical analysis of blood smear at Kasetsart University Veterinary Teaching Hospital of Kasetsart University in Nakhon Phathom province in 2005. The blood samples were performed partial 18s rRNA gene amplification by polymerase chain reaction (PCR) technique, and identified babesia species by species-specific probe of TBD-RLB KIT (Isogen®). All PCR products were also analyzed with RFLP with TaqI restriction enzyme to confirmed Babesia species. The PCR combined with species specific probe (TBD-RLB KIT) and RFLP with TaqI restriction enzyme were showed *Babesia canis vogeli* infection in these eighteen domestic dogs. The partial 18S rDNA spanning the V4 region of three samples were chosen for analysis and revealed the identical sequences with *Babesia canis vogeli* (accession no. AY072925).

Keywords: *Babesia canis* vogeli; 18s rRNA; Reverse line blot hybridization; RFLP

Production of monoclonal antibodies against 3AB non-structural protein of foot and mouth disease virus

P. Lertwatcharasarakul, O. Booddee, W. Chumsing, W. Wajjwalku

Faculty of Veterinary Medicine, Kasetsart University, Kamphaensaen, Nakhon Pathom 73140, Thailand.

Recombinant 3AB protein of a foot and mouth disease virus (FMDV) was expressed from *Escherichia coli* and immunized in mouse after gene amplification and gene cloning into pQE30 vector (Qiagen®). The immunized-mouse was collected spleen to fusion with myeloma cells strain P3-X63-Ag8.653. Hybridoma propagated in HAT medium was evaluated by recombinant 3A and 3B enzymed linked immunosorbent assay (ELISA). The results showed that positive clones produced antibodies against 3A and 3B proteins and also detected FMDV infected BHK-21 cell by immunoperoxidase monolayer assay (IPMA). The monoclonal antibodies against 3AB protein will be used for serological development to discriminate between infected and vaccinated animal.

Keywords: FMDV; 3AB protein; Monoclonal antibody

Spoligotypes of *Mycobacterium tuberculosis* strains from Brazilian tuberculosis patients

Harrison Magdinier Gomes[1], Atina Ribeiro Elias[1], Maranibia Aparecida Cardoso Oelemann[1], Márcia Aparecida da Silva Pereira[1], Fátima Fandinho Onofre Monte[1], Luís Cláudio de Oliveira Lazzarini[2], Ana Grazia Marsico[2], Afrânio Lineu Kritski[2], Luciano dos Anjos Filho[2], Paulo Caldas[3], Ângela Werneck[3], Lia Gonçalves Possuelo[4], Patrícia Cafrune[4], Maria Lúcia Rossetti[4], Norma Lucena[5], Marcílio Fernandes Baliza[5], Maria Helena Feres Saad[6], Hebe Rodrigues Cavalcanti[6], Clarisse Queico Fujimura Leite[7], Rossana Coimbra de Brito[8], Maria Luiza Lopes[9], Karla Lima[9], Maisa Souza[9], Rita de Cássia Trindade[10], Júlia Salen[11], Cristophe Sola[12], Nalin Rastogi[12], Philip Noel Suffys[1]

[1]*Laboratory of Molecular Biology Applied to Mycobacteria, Department of Mycobacteriosis, Oswaldo Cruz Institute, Rio de Janeiro, Brazil.*

[2]*Institute of Thoracic Disease, Federal University of Rio de Janeiro, Brazil.*

[3]*Reference Center Prof. Hélio Fraga, Ministery of Healthy, Rio de Janeiro, Brazil.*

[4]*Federal University of Rio Grande do Sul, Brazil.*

[5]*Centro de Pesquisa Aggeu Magalhães, Brazil.*

[6]*Department of Mycobacteriosis, Oswaldo Cruz Institute, Rio de Janeiro, Brazil.*

[7]*Department of Pharmacy State University of São Paulo, Araraquara, Brazil.*

[8]*State University of Rio de Janeiro, Rio de Janeiro, Brazil.*

[9]*Evandro Chagas Institute, Pará, Brazil.*

[10]*Federal University of Sergipe, Brazil.*

[11]*Instituto de Pesquisas da Amazônia, Brazil.*

[12]*Institute Pateur de Guadeloupe, France.*

Brazil is in 14th place among the 23 countries that concentrate 80% of Tb cases in the world, with 50 millions of infected and 6000 deaths annually. The goal of this study is to evaluate *Mtb* strains diversity on a national level. Brazil is a Republic composed by 26 states, a Federal District and has a population of around 170 million. Seventy percent live in the eleven states of which Mtb isolates were genotyped (Amazonas, Ceara, Goias, Minas Gerais, Para, Parana, Pernambuco, Rio Grande do Sul, Rio de Janeiro, Sao Paulo, and Sergipe). A total of 2000 Mtb isolates was submitted to spoligotyping and 42% of the isolates were from the Southeastern region, the richest and most industrialized Brazilian region. A considerable variability of spoligotypes was observed and upon comparison with already described genotype families, 515 were classified as the Latin American and Mediterranean (LAM) family, 20% as the T

family and 13% as the Haarlem family. Within the LAM family, 27% was LAM 9 and within the Haarlem family, 55% belonged to the Haarlem 3 class; 69% of the T family was T1. Another 16% were recognized as other families, including the X family, S family, East African and Indian Family and the Beijing family (3 strains in Rio de Janeiro city). Unrecognized profiles (U) and profiles that were classifieds only by the type number in the SpolDB4 database were also observed. Besides genotypes already described in literature, 15% were classified as new types and include profiles belonging to LAM, Haarlem, T, and other families. Interestingly was the absence of LAM class 7, and 10. Furthermore, no genotypes characteristic for other species of the MTBC were observed. Our data demonstrate the high prevalence of some strain families that are probably deeply rooted in the phylogeny of brazilian Mtb, probable consequence of the colonization process started in the beginning of XV century. Representative of many others families were found in a small frequency (Manu, Beijing, Class LAM 11 and 8); this could represent cases resulting from modern migratory flux between countries and continents.

Transcriptional levels of *Helicobacter pylori cagA* and *vacA* genes in Lebanese patients with gastritis and peptic ulcer disease

Ghassan M. Matar[1], Aline Khayat[1], Mireille M. Kattar[3], Ayman N. Tawil[3], Benjamin Gold [4], Assaad M. Soweid[2]

[1]*Department of Microbiology and Immunology, Faculty of Medicine, American University of Beirut, Beirut, Lebanon.*

[2]*Departments of Internal Medicine, Faculty of Medicine, American University of Beirut, Beirut, Lebanon.*

[3]*Departments of Pathology and Laboratory Medicine, Faculty of Medicine, American University of Beirut, Beirut, Lebanon.*

[4]*Division of Pediatric Gastroenterology and Nutrition, Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, USA.*

Background: The prevalence and clinical relevance of *Helicobacter pylori cagA* and *vacA* virulence genes in the pathogenesis of disease phenotype was assessed by a novel approach for this organism consisting of gastric mucosal *H. pylori* gene transcription levels determination and comparisons made according to disease phenotype.

Materials and methods: Gastric mucosal biopsies were collected from patients with peptic ulcer disease (PUD), gastritis, and normal mucosa in an academic medical center in Lebanon. *H. pylori* was detected in these biopsies by rapid urease (CLO^{TM}) test and PCR amplification of the *ureaseA* gene. *H. pylori* virulence genes, their transcription and transcription levels were determined respectively by PCR, RT-PCR and real time RT-PCR.

Results: Forty-five percent of patients were *H. pylori* positive by PCR of the *ureaseA* gene, 37.5% of whom had *cagA* and 59.4% had *vacA*. The *vacA* s1a allele was more prevalent in our study population and appeared to be associated with increased virulence. The *cagA* and *vacA* genes were detected and transcribed more frequently in PUD patients and the transcription levels of both *cagA* and *vacA* genes were observed to be higher in endoscopically apparent disease phenotypes (i.e., PUD, gastropathy) than in controls suggesting that they likely contribute to disease pathogenesis.

Conclusions: This study provided insight into the virulence potential of *H. pylori* encountered in Lebanese patients with gastroduodenal pathology. The results of our investigation in this regard need to be corroborated through larger studies.

The study of TH_1 and TH_2 cytokines profile (FNγ, IL-12, IL-4, IL-10) in PBMCs of patients with MDR-TB and newly diagnosed treated cases tuberculosis

N. Miandehi, M. Shekarabi, D. Mansoori, M.T. Kheiri, M. Mirsaeedi

Department of Immunology, Iran University of Medical Sciences, Tehran, Iran, Influenza Unit, Pasteur Institute of Iran, Tehran, Iran.

Multi-drug-resistant strains of *Mycobacterium tuberculosis* seriously threaten TB control and prevention efforts. Studies of the immune response mechanisms are essential. In this article, the status of cytokines release is examined after stimulating the mononuclear cells of patients. 16 MDR patients, 14 newly diagnosed TB cases were selected according to clinical/radiological data. 10 apparently healthy PPD negative individuals selected as control. Blood was obtained and PBMCs were isolated by differential centrifugation over ficoll-Hi Paque and plated at 2×10^5 cells per well. PPD and PHA were then added to proper wells and were cultured at 37 °C. Supernatants were harvested and frozen at −70 °C. Cytokine concentrations wee measured by ELISA. Concentrations of IL-10 AND lL-4 in supernatants from tuberculosis patients (responsive to treatment and MDR) does not differ significantly, while IL-12 and IFNγ was much higher in patients compared to control group. The present finding is not compatible with the data reported by John F. McDyer and J.S. Lee and it seems in our patients groups no marked imbalance in Th_1/Th_2 activity is noticeable.

Independent evolution of pyrimethamine resistance in *Plasmodium falciparum* in Melanesia

Toshihiro Mita[1], Kazuyuki Tanabe[2], Nobuyuki Takahashi[1], Kiyoshi Kita[3], Srivicha Krudsood[4], Sornchai Looareesuwan[4], Akira Kaneko[5], Anders Bjorkman[5], Takatoshi Kobayakawa[1]

[1]*Department of International Affairs and Tropical Medicine, Tokyo Women's Medical University, 9-1Kawada-cho, Shinjuku-ku, Tokyo 162-8666, Japan.*

[2]*Laboratory of Malariology, Research Institute for Microbial Diseases, Osaka University, Osaka, Japan.*

[3]*Department of Biomedical Chemistry, Graduate School of Medicinem University of Tokyo, Tokyo, Japan.*

[4]*The Asian Centre of International Parasite Control, Faculty of Tropical Medicine, Mahidol, University, Thailand.*

[5]*Department of Medicine, Malaria Research Laboratory, Karolinska institute, Stockholm, Sweden.*

A single focus origin of pyrimethamine-resistant *Plasmodium falciparum* is suggested to have spread from Southeast Asia to

Africa. We have compared the genetic profile of pyrimethamine-resistance in Melanesia, where unique chloroquine resistance developed independently, including the polymorphism of the dihydrofolate reductase gene (*dhfr*) and microsatellite haplotypes flanking *dhfr* in a total of 285 isolates from different regions of Melanesia (Papua New Guinea, Vanuatu and Solomon) and Southeast Asia (Thailand and Cambodia). Nearly all isolates (92%) in Melanesia harbored a *dhfr* double mutant (CNRNI at positions 50, 51, 59, 108 and 164), whereas 98% of isolates were either triple (CIRNI) or quartet (CIRNL) mutants in Southeast Asia. Microsatellite analysis revealed two distinct lineages of the *dhfr* double mutants in Melanesia. One lineage had the same microsatellite haplotype as reported in Southeast Asia and Africa, suggesting the spread to Melanesia from Southeast Asia. The other lineage had microsatellite haplotype not found elsewhere. This study therefore provides evidence for independent at least partly unique evolution of *P. falciparum* pyrimethamine-resistance in Melanesia, in contrast to the apparent common evolution in Southeast Asia and Africa.

Variability among *Taenia solium cysticerci* from Mexico state of Mexico

A. Monroy-Ostria[1], A.C. Hinojosa-Juárez[1], M. Sandoval Balanzario[3], G. Zúñiga Bermudez[2]

[1]*Departamento de Inmunología, Escuela Nacional de Ciencias Biológicas, IPN, Carpio y Plan de Ayala, Col. Casco de Santo Tomás, C.P. 11340, México, D.F., Mexico.*

[2]*Departamento de Zoología, Escuela Nacional de Ciencias Biológicas, IPN, Carpio y Plan de Ayala, Col. Casco de Santo Tomás, C.P. 11340, México, D.F., Mexico.*

[3]*Neurocirugía, Hospital de Especialidades del Centro Medico Nacional "La Raza" México, Mexico.*

In Mexico, neurocysticercosis has an incidence of 0.2–3.4%. High human neurocysticercosis mortality rates are found in the State of Mexico, situated around Mexico City, comprising both urban areas with a high human population and rural areas with traditional breeding of pigs that often lacks appropriate hygienic conditions. Mitochondrial COI, ribosomal ITS1, and 28S rDNA from 23 *T. solium* cysticerci isolates from pigs from several districts in Mexico State and *Cysticercus racemosus* and *C. cellulosae* from patients with neurocysticercosis were PCR-RFLP analyzed with several restriction enzymes. The PCR-RFLP data were analyzed with the Li Nei ($S_{ij} = 2\ n_i - j / n_i + n_j$) index, the similarity matrix was constructed among the different isolates of cysticerci using Ntsys-PC Version 2.0 software and the corresponding dendrogram with the UPGMA program. The statistical analysis was done with the Manteel test. These analyses showed three groups of cysticerci isolated from pigs and three groups from human brain. The dendrograms demonstrated that there is intraspecific variability in *T. solium* isolates from Mexico State. Sequencing and polymorphism analysis showed that the phylogenetic tree of ITS1 has two related groups, the first Mexican and the second Phylipino. The consensus tree of the COI gene shows four groups and the highest nucleotide diversity of the three analyzed sites.

Acknowledgements

A. Monroy-Ostria and G. Zúniga-Bermúdez are supported by COFAA and EDI, IPN, Mexico. A.C. Hinojosa-Juárez was sponsored by CONACYT, Mexico.

Variation in the O-150 *Onchocerca* specific repeat sequence family of *Onchocerca volvulus* Mexican isolates

A. Monroy-Ostria[1], A. Ramírez-Ramírez[1], G. Sanchez-Tejeda[2], J. Mendez-Galvan

[1]*Departamento de Inmunología, Escuela Nacional de Ciencias Biológicas, IPN. Carpio y Plan de Ayala, Col. Casco de Santo Tomás, C.P. 11340, México, D.F., Mexico.*

[2]*Centro de Vigilancia Epidemiologica, SSA, Benjamín Franklin No. 132, Col. Escandon, Mexico, D.F., Mexico.*

Onchocerciasis has been the world's second leading infectious cause of blindness. It is endemic to Africa, the Arabian Peninsula, and the Americas. Around 120 million people worldwide are at risk of onchocerciasis. In Mexico there are two endemic foci one in Oaxaca and the other in Chiapas States. The disease is caused by the nematode *Onchocerca volvulus.* DNA from adult worms of *Onchocerca volvulus* from Oaxaca (24 onchocercomata) and Chiapas (28 onchocercomata), Mexico were used as templates to amplify and sequencing members of the O-150 *Onchocerca* specific repeat sequence family. The O-150 sequences of Mexican *O. volvulus* isolates were aligned and compared with the O-150 sequences of savanna and rain forest strains of Africa, one strain from Guatemala and one from Brazil. The statistical analysis was done and the corresponding trees were built. It was found higher variation in the Mexican isolates sequences than the Brazilian and African. Mexican and Guatemala isolates are more related to the African rain forest and the Brazilian with the African Savanna strains.

Acknowledgements

A. Monroy-Ostria was supported by COFAA and EDI, IPN, Mexico. A. Ramírez-Ramírez was sponsored by CONACYT, Mexico.

Reference

Meredith, S.E.O., Lando, G., Gbakma, A.A., Zimmerman, P.A., Unnasch. T.R., 1991. *Onchocerca volvulus*: application of the polymerase chain reaction to the identification and strain differentiation of the parasite. Exp. Parasitol. 73, 335–344.

Study on iron uptake regulation in *Agrobacterium tumefaciens*

Patchara Ngok-ngam[1], Rojana Sukchawalit[2], Skorn Mongkolsuk[1,2]

[1]*Department of Biotechnology, Faculty of Science, Mahidol University, Rajadhevee, Rama VI Road, Bangkok 10400, Thailand.*

[2]*Laboratory of Biotechnology, Chulabhorn Research Institute, Lak Si, Bangkok 10210, Thailand.*

Iron is critical for bacterial growth, but problems arise from the toxicity of excess iron; thus, iron uptake is subject to tight control. The most widely found and best studied iron-responsive regu-

lator in Gram-negative bacteria is the ferric uptake regulator Fur. In recent years, it has become apparent that iron regulation in rhizobia differs from that in many other bacteria. New regulators, RirA were identified which appear to mediate functions that in other bacteria are accomplished by Fur. Agrobacterium tumefaciens causes crown gall disease of a wide range of dicotyledonous plants. The A. tumefaciens Fur protein does not exhibit significant role on iron regulation. Mutation of rirA shows the sensitive phenotype to iron, hydrogen peroxide, tert-butyl hydroperoxide, menadione and streptonigrin which resulted from iron overload. The result form reverse transcriptase PCR (RT-PCR) indicated that Fur regulates sitABCD (an ABCs-type transporter). RirA plays an important role in regulation of iron responsive genes which are siderophore production and suf operon; an [Fe-S] cluster formation. Blast searched indicate putative binding box of RirA; IRO-like motif lie on the siderophore and suf gene. The ability to infect plant is affected by mutation of rirA shows in decreasing of both virB and virE promoter activity and slightly reduced tumor formation.

Keywords: *Agrobacterium tumefaciens*; Iron regulation; Fur; RirA

Occurrence of two heterophyid metacercariae from freshwater fish in reservoirs

C. Nithikathkul[1,2], C. Wongsawad[1,2]

[1]*Department of Biology, Faculty of Science, Chiang Mai University, Thailand.*

[2]*Department of Biological Science, Faculty of Science and Technology, Huachiew Chalermprakiet University, Thailand.*

According to current information, the matacercaria of heterophyid trematode found in freshwater fishes, especially the cyprinoid group. A parasitological investigation was made using 241 fishes from 12 species: *Barbodes schwanenfeldi*, *Labiobarbus siamensis*, *Barbodes gonionotus*, *Osteochilus hasselti*, *Henicorhynchus siamensis*, *Cyclocheilichthys armatus*, *Cyclocheilichthys apogon*, *Mystacoleucus marginatus*, *Cirrhinus cirrhosus*, *Notopterus notopterus*, *Oxycleotris marmoratus* and *Pristolepis fasciatus.* The fishes were caught over the summer period (February 2006–May 2006) in the Chiang Mai water reserviors (Mae Ngud and Mae Kwong) and the Chiang Rai water reservoirs (Nong Luang and Mae-Kataa). The prevalence of heterophyid metacercariae infection at Mae Ngud, Mae Kwong, Nong Luang and Mae Kataa were 15.04%, 5.39%, 12.37% and 0.07% respectively. The highest prevalence of heterophyid metacercariae infection in *H. siamensis* in Mae Ngud was 17.46%. The freshwater fish, *C. cirrhosus*, *N. notopterus*, *O. marmoratus*, and *P. fasciatus* were not found to be infected with heterophyid metacercariae.

Recovery of *Haplorchis taichui* and *Stellantchasmus falcatus* from small intestines of *Gallus gallus domesticus*

C. Nithikathkul[1,2], C. Wongsawad[1,2]

[1]*Department of Biology, Faculty of Science, Chiang Mai University, Thailand.*

[2]*Department of Biological Science, Faculty of Science and Technology, Huachiew Chalermprakiet University, Thailand.*

Parasitic zoonoses are widespread in Southeast Asia, particularly Thailand. Among fish-borne trematode infection caused by many species. It is a universally observed characteristic of parasites that they infect a restricted group or a few restricted groups of hosts. This phenomenon has been known as ''host specificity'' *Stellantchasmus falcatus*, one of the minute intestinal flukes of fish-eating birds and mammals, was first described by Onji and Nishio (1915). *Haplorchis taichui* is a minute intestinal fluke (MIF) that parasitizer the small intestinal of birds and mammals including human (Faust and Nishigori, 1926). An experimental study was performed to observe the recovery of minute intestinal flukes, *H. taichui* and *S. falcatus* from small intestines of chicks (*Gallus gallus domesticus*). Metacercariae of *Haplorchis taichui* were isolated from *Labiobarbus siamensis* and *Henicorhynchus siamensis* and metacercariae of *S. falcatus* were isolated from *Dermogenys pusillus* which were collected in the Chiang Mai Province, Thailand, by using 1% acid pepsin solution in a blender. The digested material was incubated in shaking water bath for one and a half hour at 37 °C incubation and subsequently passed through two layers of wet gauze. The digested material was rinsed with 0.85% sodium chloride solution and examined for metacercariae under the stereomicroscope. The identification of metacercariae was carried out by morphological examination based on Sholz et al. (1991) and Wongsawad et al. (2000) under a compound microscope. The 25 of one day-old chicks were orally force fed with a dose of fifty metacercariae of *H. taichui* and one hundred metacercariae of *S. falcatus*. The worms were recovered from small intestines of chicks by Dearman's apparatus technique. The intestine of the chicks were examined in seven day post-infection (PI). The infection rate of *H. taichui* and *S. falcatus* were 12.32 and 1.44% respectively.

Keywords: *Haplorchis taichui*; *Stellanchasmus falcatus*; *Gallus gallus domesticus*

Phenotypic resistance prediction from genotypes for human immunodeficiency virus type 1 (HIV-1) protease inhibitors using neural networks

Ekawat Pasomsub[1], Boonserm Kijsirikul[2], Wasun Chantratita[1]

[1]*Department of Clinical Pathology, Faculty of Medicine, Ramathibodi Hospital, Mahidol University, Thailand.*

[2]*Department of Computer Engineering, Chulalongkorn University, Thailand.*

The 598 HIV-1 protease sequences and their corresponding phenotypic fold change values for six drugs were retrieved from the Stanford HIV RT and Protease Database for neural network modeling using software Clementine Version 7.0. The results were compared with those from the rule-based method from the Stanford HIV RT and Protease Database and the support vector machine method from the Geno2Pheno. The amino acid input pattern encodings gave higher total correlation coefficient values than the binary input pattern encodings, which ranged from 0.83 to 0.93 and the best total correlation coefficient value was 0.93 from ''AA Rb'' input pattern encoding of the ritonavir resistance dataset. The neural network system provided a high correlation coefficient of 0.96 and high accuracy of 95%, both

of which were higher than the other two systems, when compared with experimental phenotypic testing values. Regarding consensus based prediction; neural network system predicted values also showed better results (97%) than the other two systems.

Keywords: Virtual phenotype; Neural networks; HIV-1 protease inhibitors; Resistant mutations; Fold change; Inhibition concentration

Detection of vancomycin resistance genes and virulence factors in vancomycin-resistant enterococci isolated from patients hospitalized in large university clinical hospital during 2-year period

M. Przybylski Maciej[1,2], F. Marchel Halina[2], F. Rudnicka Jolanta[2], F. Wroblewska Marta, M. Luczak Miroslaw[1,2]

[1]*Medical University in Warsaw, Department of Microbiology, Chalubinski 5 Srt., 02-04 Warsaw, Poland.*

[2]*Independent Public Central Clinical Hospital, Microbiology Department, Banacha 1A Str., 02-097 Warsaw, Poland.*

We analyzed group of 98 VRE strains isolated from patients hospitalized in large university clinical hospital in Warsaw, Poland during 2-years period (2000–2002). Strains were isolated from blood, wound, peritoneal cavity, bile and feces. We tested VRE strains for their ampicillin and glycopeptides susceptibility, type of van genes and virulence factors genes carried. Strains were identified with phenotype-based methods and checked for species-specific *Enterococcus faecium* and *Enterococcus faecalis* ddl genes by PCR. MICs for ampicillin and glycopeptides were determined. vanA, vanB, vanC and vanD genes were detected by PCR. Virulence factors genes were detected using PCR method with primers specific for five genes from cytolysin complex, gelatinase, aggregation substance, enterococcal proteins (EfaA) and surface protein genes. Among 98 strains of VRE, the most prevalent was *E. faecium* (82,7%), according to biochemical identification results. Other VRE species embraced *E. faecium*, *E. durans*, *E gallinarum* and two strains of *Enterococcus* spp. Discrepancy between phenotypic and genetic identification results was noted in case of 11.2% of strains. Resistance to teicoplanin was detected in 96.9% of strains. Three VRE strains were ampicillin susceptible. The most prevalent glycopeptide resistance gene was vanA (92.9%). Other types of glycopeptide resistance genes were vanB (three strains) and vanD (one strain). In two VRE strains, we were unable to detect any van gene. The most prevalent virulence factor gene was efaA (88.8% of strains). Other virulence factors genes were also present: esp in 72.4% of strains, gelE (12.2%) and agg (10.2%). Various genes of cytolysin complex were found in 11.2% of strains, but there was no strain carrying complete set of cyl genes. We characterized two groups of epidemic strains, and seven groups of VRE not responsible for epidemic infections. Only in case of esp gene we found significant correlation between site of infection and carried virulence factor.

Keywords: VRE; Hospital acquired information; Virulence factor

Understanding the interaction and the structure–activity correlation of efavirenz derivatives and WT and K103N HIV-1 RT by molecular docking and 3D-QSAR approaches

Pornpan Pungpo[1], Oradee Pankwang[2], Patchreenart Saparpakorn[2], Supa Hannongbua[2]

[1]*Faculty of Science, Ubonratchathani University, Ubonratchathani, 34190, Thailand.*

[2]*Faculty of Science, Kasetsart University, Bangkok, 10900, Thailand.*

Molecular docking and 3D-QSAR analyses were performed to understand the interaction between a series of efavirenz derivatives with WT and K103N HIV-1 RT. To model the potential binding modes of efavirenz derivatives in the binding pocket of WT and K103N HIV-1 RT, molecular docking approaches by using GOLD and Autodock 3.0 programs were performed. The results show that the docking results obtained from both methods reveal a good ability to reproduce the X-ray bound conformation with rmsd less than 1.0 Å for both WT and mutant enzymes. The docking calculations of all efavirenz derivatives in the data set were, consecutively, performed to elucidate their orientations in the binding pockets. The results derived from docking analysis give additional information and further probes the inhibitor–enzyme interactions. The correlation of the results obtained from docking models and the inhibitory activities validate each other and lead to better understanding of the structural requirements for the activity. Therefore, these results are informative to improve the development of more efficient HIV-1 RT inhibitors, especially, active against mutant enzyme. Based on the molecular alignment of conformations obtained from molecular docking procedures, the high predictive 3D-QSAR models were produced by using CoMFA and CoMSIA approaches. The CoMFA models reveal the importance of steric and electrostatic interactions through contour maps. The resulting CoMSIA models enhance the understanding of steric, electrostatic, hydrophobic, electron donor and acceptor requirements for ligands binding to the K103N HIV-1 RT. Consequently, the results obtained from structure-based and ligand-based design approaches can be integrated to identify the structural requirements of HIV-1 RT inhibitors in the class of efavirenz compounds. The principle derived from the present study provides a beneficial guideline to design and predict new and more potent compounds active against K103N HIV-1 RT.

Characterization of *sod* genes involved in oxidative stress response and tumor formation in *Agrobacterium tumefaciens*

Panatda Saenkham[2], Warawan Eiamphungporn[2], Paiboon Vattanaviboon[1], Skorn Mongkolsuk[1,2]

[1]*Laboratory of Biotechnology, Chulabhorn Research Institute, Lak Si, Bangkok 10210, Thailand.*

[2]*Department of Biotechnology, Faculty of Science, Mahidol University, Bangkok 10400, Thailand.*

Phytopathogenic bacteria are always exposed to reactive oxygen species (ROS) generated from aerobic metabolisms as well as from the host plant first line defense. To avoid oxidative

damages, bacteria produce an array of ROS-scavenging enzymes to detoxify them before reaching harmful level. Superoxide dismutase (SOD) capable of catalyzing the dismutation of superoxide anions to H_2O_2 and molecular oxygen. The removal of superoxide blocks the secondary reactions which lead to formation of the highly reactive hydroxyl radicals. Here, we cloned and characterized multiple SODs from *Agrobacterium tumefaciens*, the causative agent of crown gall disease in plants. Analysis of its genome revealed three genes encoded iron superoxide dismutases (FeSODs). Sod1 and Sod3 are cytoplasmic while Sod2 is a periplasmic isoenzyme. *sod1* was expresses at relatively high level and appears to be increased during stationary phase. The expression of *sod2* is a phase-dependent and produced at detectable level at stationary phase. *sod3* is a member of SoxR regulon whose expression could be strongly induced upon exposure to superoxide anion. Deprivation of *sod1* markedly alleviated resistance to superoxide generator. Moreover, the strain increases tolerance to H_2O_2 due to a compensatory expression of catalase-peroxidase. Inactivation of *sod2* or *sod3* alone slightly affected ability of bacteria to cope superoxide toxicity. The SOD null mutant is extremely sensitive to killing treatment with superoxide generator and attenuated the ability to cause tumor on plant leaves. All the evidences indicate that superoxide dismutases are not only critical enzymes responsible for protection of superoxide anion but also required for a virulence of *A. tumefaciens*.

Keywords: *Agrobacterium tumefaciens*; Oxidative stress; *sod*

RNA interference in the malaria parasite and its possible use in treatment

Maryam Saleh[1,2], Virander Singh Chauhan[1,2]
[1]*Pasteur Institute of Iran, Tehran, Iran.*
[2]*ICGEB, New Delhi, India.*

Introduction: Malaria is still a major parasitic disease despite efforts spanning more than a century to eradicate or control it because of the poor understanding of the functions of the malaria parasite. One of the main reasons for the slow progress in the development of new anti-malarial or for an effective malaria vaccine has been the poor understanding of the functions of most of the malaria parasite proteins. Recently RNA interference (RNAi) has emerged as a powerful tool to understand the gene function in variety of organisms and to inhibit the gene expression on the level of single gene, gene families and the entire genome. The major goals of this study is to develop RNAi technologies for functional genomic studies in malaria and follow it, for next studies in future, is to understand molecular mechanisms of malaria drugs and host-parasite interactions.

Method: *Plasmodium falciparum*, and to a much lesser extent *Plasmodium vivax,* are the main causes of disease and death from malaria.

The genome sequencing project of the human malaria parasite, *P. falciparum* has identified 5300 proteins, of which 60% (3208) have not been assigned any function. Even though gene targeting by homologous recombination has been successfully used to understand the functions of a number of parasite proteins, it still has limitations. Following methodology has been used for carrying out RNAi for the two cystein protease genes of *Plasmodium falciparum*:

1. Selection of dsRNA/siRNA sequences: the length requirement of dsRNA has been recommended greater than 500 bp.
2. dsRNA preparation: using PCR with appropriate RNA polymerase.
3. Preparation of siRNA (Donze and Picard, 2002).
4. Treatment of parasite with dsRNA/siRNA.
5. FACS analysis of dsRNA/siRNA treated GFP parasite lines.
6. Detection of reduction in protein by Western blotting.

Results: This study earlier carried out an in-depth study to establish RNAi in *P. falciparum in vitro* for the two cystein protease genes (*falcipain* 1 and 2) of the parasite. Using dsRNAs corresponding to cystein protease genes of *P. falciparum*, we demonstrated that *falcipains* play an important role in hemoglobin degradation. In this process, specific dsRNA elicits the degradation of cognate mRNA.

Trypanosoma cruzi infection reactivation after immunosupression is correlated with the parasite genetic diversity

D.M. Santos[1], R.R. Silva[2], T.A.F. Martins[2], F.B. Queiroz[1], I.S. Caldas[1], V.M. Veloso[2], C.M. Carneiro[1,3], M. Lana[1,3], M.T. Bahia[1,2]
[1]*Núcleo de Pesquisa em Ciências Biológicas, Universidade Federal de Ouro Preto, Brazil.*
[2]*Departamento de Ciências Biológicas, Universidade Federal de Ouro Preto, Brazil.*
[3]*Departamento de Análises Clínicas, Universidade Federal de Ouro Preto, Brazil.*

The factors involved in the reactivation of chagasic infection are not clear enough and may be related to selective host immune depletion or parasite genetic diversity. To evaluate the role of the parasites genetic in *T. cruzi* infection reactivation induced by ciclophosphamide immunosupression, groups of 32 Swiss mice were inoculated with *T. cruzi* clonal stocks classified as *T. cruzi* I (Cuicacl1, P209cl1, Gambacl1, SP104cl1) and *T. cruzi* II (Bug2148cl1, MNcl2, IVVcl4, MVBcl8) were used. Infected animals were treated with ciclophosphamide when was with subpatent parasitemia still during the acute phase (AP), and chronic phase (CP). Animals infected with *T. cruzi* I stocks showed 82,6% and 47,5% of parasitemia reactivation during the AP and CP, respectively, being observed 0%, 100%, 100% and 100% of parasitemia reactivation in animals inoculated with clones SP104cl1, Gambacl1, Cuicacl1 and P209cl1, respectively in the AP, and 80%, 40%, 50% and 20% in CP. On the other hand, animals infected with *T. cruzi* II showed only 4.1% of parasitemia reactivation when immunosuppressed during the AP and 0% in CP. However the heart and skeletal muscle lesions of animals infected by *T. cruzi* I were similar to those observed in controls infected and not immunosupressed group (CI). By the way an increase of encephalic lesions in animals immunosupressed during the CP in relation to CI was observed. Although parasitemia reactivation was not observed

in animals infected with *T. cruzi* II clones, an increase of inflammatory process in the heart and skeletal muscle, but not in the brain, was observed among animals infected with Bug2148cl1. These results showed that the genetic diversity of *T. cruzi* has an important role on the reactivation of the infection after immunosupression and corroborates the working hypothesis subjacent to the model clonal theory in *T. cruzi*.

Identification of *Theileria* in endangered serow and endangered eld's deer in Thailand

P. Sanyathitiseree[1], P. Lertwatcharasarakul[1], N. Thougtip[1], S. Phatthanakunanan[1], S. Jala[1], W. Tanwattana[2], J. Thaewnern[2], W. Wajjwalku[1]

[1]*Faculty of Veterinary Medicine, Kasetsart University, Kamphaensaen, Nakhon Pathom 73140, Thailand.*

[2]*K Kheow Open Zoo, Chonburi, Thailand.*

The blood samples of eld's deer and in the zoo and poaching serow were colleted for DNA extraction. Polymerase Chain Reaction (PCR) was performed with the 18s rRNA specific primer for babesia and theileria. The PCR products were determined species specific by species-specific probe of TBD-RLB KIT (Isogen®). Two PCR products of Eld's deer and also of the serow hybridized to *Theileria*/*Babesia* common probe but not found hybridization to species specific probe position. Moreover, nucleotide sequences of PCR products were determined by using BigDye Terminal Cycle Sequence Kit and analyzed with blast program which closely related with *Theileria* sp. The sequence of the parasite from the serow, *Theileria* sp. (Khao Yai), showed most similarities with *Theileria* sp. (OT1) of the ovine reported in Spain and from elder's deer, *Theileria* sp. (Khao Khaew) was highest similarity with *Theileria* sp. of deer in Japan. The completely sequence of 18srRNA will be done soon. This is the first report of theileriosis in the wildlife serow and elder's deer in the zoo in Thailand.

Keywords: Theileria; 18s rRNA; Reverse line blot hybridization

Using a climate dependent matrix model to predict mosquito abundance: Application to *Aedes*(*Stegomyia*)*africanus* and *Aedes*(*Diceromyia*) *furcifer* (Diptera: Culicidae), two main vectors of the yellow fever virus in West Africa

Brigitte Schaeffer[1], Bernard Mondet[2], Suzanne Touzeau[1]

[1]*INRA, Jouy-en-Josas, France.*

[2]*IFP/IRD-UR 178, Pondicherry, India.*

Mosquitoes, acting as vectors of diseases, are particularly involved in the transmission of viruses. Thus the abundance of some tree-hole breeding species strongly depends on weather, specially rainfall. The aim of this paper is to provide a tool to predict vector abundance. In order to describe the dynamics of these mosquito populations, we developed a matrix model integrating climate fluctuations. The population is structured in five stages: two egg stages (immature and mature), one larval stage and two flying stages (nulliparous and parous adult females). We considered the water availability in breeding-sites as the main environmental factor affecting the mosquito life cycle. The model represents the evolution of the mosquito abundance in each stage over time, in connexion with water availability. This model was used to simulate the abundance trends over three years of two mosquito species, *Aedes africanus* (Theobald) and *Aedes furcifer* (Edwards), vectors of the yellow fever in Côte d'Ivoire, West Africa. Water dynamics in the tree-hole was reproduced from daily rainfall data. The results we obtained show a good match between the simulated population and the field data over the time period considered.

Keywords: Vector population; Mathematical model; Climate dependency

Human immunodeficiency viruses type 1 circulating in the Comunitat Valenciana (Spain)

V. Sentandreu[1], M.A. Bracho[1], I. Alastrue[2], J. Belda[3], F. González-Candelas[1]

[1]*Cavanilles Institute for Biodiversity and Evolutionary Biology, University of Valencia, Spain.*

[2]*Center for AIDS Information and Prevention, Public Health Department of Valencia, Spain.*

[3]*Center for AIDS Information and Prevention, Public Health Department of Alacant, Spain.*

Human immunodeficiency virus type 1 (HIV-1) mutates rapidly and nucleotide substitutions, deletions, insertions, and rearrangements resulting from recombination events are the main factors contributing to its high degree of genetic heterogeneity. Nucleotide sequence analyses allow the recognition of phylogenetic relationships and the classification of HIV-1 into different subtypes and recombinant forms. The prevalence of multiple HIV-1 subtypes in a single geographic region might result in an increased frequency of mixed infections. Eventually, recombinants between more than one variant can be found. HIV-1 subtype B is the most prevalent variant in Spain, although non-B subtypes have been reported, mainly among African immigrants. Classification of HIV-1 into subtypes is based primarily on the analysis of sequences coding for the *env* gene. However, the *pol*-coding region has also been validated for this purpose and is currently used much more since drug-resistance testing is undertaken routinely at a large scale. Here we describe the main features of HIV-1 subtypes circulating in the Comunitat Valenciana (Valencia and Alacant provinces). We have amplified and sequenced the HIV-1 protease and partial reverse transcriptase (PR-RT) genes from isolates of 75 patients. Samples were obtained from two Centers of AIDS Information and Prevention and 40% of them correspond to immigrants. Subtype identification was performed by phylogenetic analyses, taking as reference a panel of 120 HIV-1 sequences representing all subtypes and recombinant forms described. We have found 60 (80%) subtype B sequences. Non-B subtypes were mainly represented by recombinant forms (14/15, 18.7%) from patients coming from Africa and South America and only one for subtype A. At least one drug-resistance mutation in the *pol* gene was detected in 86.7% of sequences. This study demonstrates that most viruses circulating in Spain (Valencia-Alacant) are in fact inter-subtype recombinants, with CRF02_AG being the most prevalent recombinant form.

Evidence of recombination in hepatitis C virus intrapatient populations

V. Sentandreu, N. Jiménez, M. Torres-Puente, M.A. Bracho, M.J. Gosalbes, A. Moya, F. González-Candelas

Cavanilles Institut of Biodiversity and Evolutionary Biology, Valencia University, Spain.

Hepatitis C virus (HCV) is a major cause of liver disease worldwide and a potential cause of substantial morbidity and mortality in the future. The prevalence of HCV infection is estimated to be 2%, representing about 170 million people in the world. Hepatitis C virus is characterized by a high degree of genetic heterogeneity. Although homologous recombination has been demonstrated in many members of the family Flaviviridae, to which HCV belongs, there are few studies reporting recombination on natural populations of HCV, suggesting that these events are rare *in vivo*. Furthermore, these few studies have focused on HCV recombination between different genotypes/subtypes but there are no reports about the extent of intragenotype or intra-subtype recombination between viral strains infecting the same patient. Given the important implications of recombination for RNA virus evolution, our aim in this study has been to assess the existence and eventually the frequency of intragenic recombination on HCV. For this, we have analyzed two regions of the HCV genome (NS5a and E1–E2) in viruses obtained from patients belonging to two different groups: (i) infected only with HCV (either treated with interferon plus ribavirin or treatment naïve), and (ii) HCV–HIV co-infected patients (with and without treatment against HIV). The complete data set included more than 16000 clonal sequences from 215 samples of 119 patients. Recombination analyses were performed using 6 different methods implemented in RDP3 program. We have detected recombination events (by at least 3 of the 6 methods used) in 12% of the samples, which belonged to all the groups described and to the two genomic regions studied. Consequently, intragenic recombination cannot be disregarded as a potentially important mechanism generating genetic variation in HCV.

Identifying deeper taxonomic relationships among RNA viruses

Beth Shapiro, Robert Belshaw, Andrew Rambuat

Oxford University, Department of Zoology, Oxford OX1 3PS, UK.

Although approximately 200,000 gene sequences (and ∼1000 complete genomes) from RNA viruses are available on GenBank, the origin, phylogenetic relationships and mechanisms of genome evolution of these important infectious agents are largely unknown. This is due in part to the rapid rate of mutation of RNA viruses, which results in very little sequence similarity remaining even between closely related viruses and after relatively little evolutionary time. As a consequence, traditional phylogenetic methods are inappropriate for reconstructing evolutionary relationships between RNA viruses. Here, we use an alignment-free method to assess the relatedness between RNA viruses. We perform pairwise comparisons between all available RNA virus genomes, and between each viral genome and 1000 simulated genomes to assess the extent to which short fragments of amino acid sequences are conserved across various taxonomic levels. We find significantly more conservation of amino acid fragments than is expected by chance when comparing within genera, within subfamilies, and within families. These results suggest that, despite the high rate of mutation, it may be possible to use phylogenetic methods based on identifying sequence homology to reconstruct deeper taxonomic relationships among RNA viruses.

Towards high-throughput molecular diagnosis of Plasmodium: New approaches and molecular markers

Nicolas Steenkeste[1], Sandra Incardona[1], Sophy Chy[1], Sean Hewitt[2], Linda Duval[1], Doung Socheat[3], Tho Sochantha[3], Jean Yves Coppee[4], Pharath Lim[1], Marie-Therese Ekala[1], Odile Mercereau-Puijalon[5], Thierry Fandeur[1,5], Frederic Ariey[1]

[1]*Laboratoire d'Epidémiologie Moléculaire, Institut Pasteur du Cambodge, Phnom Penh, Cambodia.*

[2]*European Commission National Malaria Control Program, Phnom Penh, Cambodia.*

[3]*National Center for Parasitology, Entomology and Malaria Control, Phnom Penh, Cambodia.*

[4]*Plate-Forme 2, Puce à ADN, Génopole de l'Institut Pasteur de Paris, Paris, France.*

[5]*Unité d'Immunologie Moléculaire des Parasites, CNRS URA 2581, Institut Pasteur, Paris, France.*

Malaria epidemiologic studies require diagnosis tools that have to deal with a large scale. The molecular approach offers this flexibility. Thus, two PCR assays based on the SSU rRNA and cytochrome *b* genes were developed in our laboratory. These new ways of diagnosing malaria are designed for being compatible with high-throughput methods, namely Dot Blot and SNP analysis. In Cambodia, a cross-sectional malaria survey was conducted in the Rattanakiri province which is characterized by a high malaria transmission rate and low levels of drug-resistance compared with the rest of the country. In three selected villages, 337 blood spots were collected. Based on these samples, our new diagnosis techniques were compared with two reference methods: microscopy and a nested PCR method published by Singh et al. (1999). Our results confirmed the previously reported high sensitivity and specificity of molecular methods. Indeed, the prevalence of *Plasmodium* infections in the three studied villages increased from 41.5%, using Giemsa-stained thick blood smears, to 76%, using the reference PCR method. The new high-throughput methods resulted in a prevalence of 80% (Dot Blot) and 85% (SNP analysis), respectively. For the majority of samples, species typing was also confirmed by these three methods. Contradictory results were mostly related to detection of minor species (*P. malariae* and *P. ovale*) in mixed infections. However, similar results were obtained in 83% of tested isolates. Molecular large-scale methods provide more accurate informations of the malaria epidemiology in a country or region. In particular, they reveal a much larger distribution of *Plasmodium* infections than previously supposed and thus are useful for a better follow-

up of malaria control measures. The implication of these new tools will be discussed, in particular the possibility to adapt cytochrome b SNP detection on DNA microarrays.

CD40L is a type II membrane protein comprised of 261 amino acids

M. Steiper, S. Parikh, J. Zichello

Hunter College of the City University of New York (CUNY), 695 Park Avenue, New York, 10021, USA.

CD40L is a type II membrane protein comprised of 261 amino acids. CD40L plays a crucial role in the immune system where it is primarily expressed on activated T cells and triggers immunoglobulin class switching. The genetic disease X-linked hypergammaglobulinemia (HIGM1, XHIGM, or XHIM) is caused by mutations in the CD40L gene. Individuals with HIGM1 have are susceptible to recurrent infections to pathogens and a relationship has been shown to exist with malaria (Sabeti et al., 2002). The CD40L gene is under strong and recent natural selection in humans (Sabeti et al., 2002). In this paper, we phylogenetically examine the promoter region of CD40L in primates and other mammals via phylogenetic shadowing. This analysis revealed several regions of the promoter of CD40L that were highly constrained and thereby inferred to be functional. These constrained regions confirmed known regulatory sites that had been studied in vitro. In additional, a highly constrained region with an NF-AT recognition site was also identified. This region would be an excellent target for studies of CD40L regulation in vitro. These analyses also showed that the primate and rodent CD40L do not share a similar set of promoter binding sites, and instead that a 'mouse specific' and a 'primate specific' promoter has evolved, suggesting that this gene is differently regulated in these species.

The SARS-coronavirus nucleocapsid protein: A protein with multifarious activities

Milan Surjit[1], Boping Liu[2], Vincent T.K. Chow[2], Sunil K. Lal[1]

[1]*Virology Group, International Centre for Genetic Engineering & Biotechnology, Aruna Asaf Ali Road, New Delhi 110067, India.*

[2]*Human Genome Laboratory, Microbiology Department, Faculty of Medicine, National University of Singapore, Kent Ridge, Singapore 117597, Singapore.*

The SARS-CoV nucleocapsid (N) protein is a major structural component of the virus capsid and is postulated to play important roles in viral pathogenesis, replication, and RNA packaging. Besides these, the N protein has also been predicted to be involved in a variety of other important functions in the viral life cycle. Here we present and discuss recent data obtained from our laboratory showing the diversity of host pathways that this protein may be involved in. We have tested the capability of N protein to self-associate using its ~140 amino acid interaction domain implying that N may be involved in various regulatory activities in the infected cell. Mammalian cell expression studies further proved our *in-silico* predictions that the 46 kDa N protein is a phosphoprotein. Immunofloure-sence, *in-vitro* phosphorylation and c-DNA substraction techniques subsequently were used to prove that the serine-phosphorylated N was stable and localized in the cytoplasm and co-precipitated with the membrane fraction. N was a substrate of cyclin dependent kinase (CDK), glycogen synthase kinase (GSK3) and casein kinase II (CKII). Phosphorylated N translocated to the cytoplasm by binding 14-3-3 (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein) thus revealing a phosphorylation dependent nucleoplasmic shuttling mechanism. The N protein directly inhibited the activity of the cyclin-CDK complex resulting in hypophosphorylation of retinoblastoma protein with a concomitant down-regulation in E2F1 mediated transactivation. Our data clearly points towards N having a major and multifarious role to play in SARS-CoV life cycle and pathogenesis.

Designing and searching of highly effective anti-virus compounds as preparations for prophylaxis of AIDS and Flu pandemic

I. Timofeev, N. Varaksin, A. Serbin, N. Perminova

FSRI SRC VB "Vector", Koltsovo, Novosibirsk region; R&DH, Moscow, Russia.

Annually in the world there are some millions new cases HIV infection, and FLU infections emerging constantly. At present, the whole humanity can be considered as a risk group for AIDS and FLU. Radically efficient medicines, which could allow curing HIV infection, have not been created yet; available preparations only somewhat slow down the development of AIDS. It is necessary creation of new highly effective preparations for prophylaxis and treatment HIV-1/2 infection. To preparations against a FLU the drug-resistance quickly develops, therefore creations of new effective compounds also are actually. On the basis of co-polymers of divinyl ether with maleic anhydride, polymeric matrixes modified with Norbornane have been synthesized and adamantine, and/or peptide imitators chemokine receptors for complex original anti-virus compounds have been designed and synthesized, also. By evaluating the newly synthesized complex compounds *in vitro,* their low toxicity was revealed ($CC_{50} > 2.0$ mg/ml), as well as high anti-FLU and anti-HIV activity (by means of viral procedure: suppression of reproduction of virus (EC_{50} from 0.5 μg/ml), and by means of ELISA measuring inhibition of the production of p24 HIV-1 protein (IC_{50} from 0.1 μg/ml). Maximal level of antiviral efficiency was revealed when preparations were introduced at the stage of virus adsorption or/and was during the whole of cultivation. From our point of view, these compounds can be promising for the development of anti-FLU and anti-HIV microbicide preparations. Questions concerning the development of optimal means of transporting the complex compounds of this class are investigated and the effective medicinal forms also are prepared. In order to eliminate irritating action we included efficient anti-virus compounds into the pH-dependent interpolymeric complex (IPC) which is stable into weakly acidic media and decompose with the release of active anti-virus preparation in neutral and/or alkaline media. The anti-virus efficiency of some IPC proposed has been demonstrated experimentally *in vitro* (up to 99% HIV-1 suppressed), along with a decrease in local toxic action of

preparation included in the IPC in vivo when applied on vaginal mucous membrane of white mice (morphological studies and histology). Thus, now we have high efficient candidates of anti-FLU compounds and anti-HIV microbicide gel for local intra vaginal using for anti-HIV prophylaxis and therapy.

Quantitative real-time PCR applications proposed for investigations of environmental leptospirosis

Mayee Wong, Brett Ellis, Bruce A. Wilcox

Department of Tropical Medicine, Medical Microbiology, and Pharmacology University of Hawaii Manoa 651 Ilalo Steet, Biosciences Building 320, Honolulu, HI 96813, United States

Leptospirosis is a globally emergent infectious disease involving complex ecological processes that as such requires a systems analytic approach in order to elucidate patterns of pathogen transmission within the host reservoir community. An ecosystems-level model of leptospirosis acknowledges the cross-scale interactions that are relevant to the ecology of this disease and recognizes the necessity of a trans-discipline solution that includes consideration of the molecular, cellular, and organismal factors likely to be involved in disease emergence in human populations. We highlight three specific contributions to an understanding of leptospirosis in the environment that can be made by application of a recent advance in molecular techniques, specifically quantitative polymerase chain real-time reaction or Q RT-PCR methods. The use of Q RT-PCR in ecological investigations of pathogenic leptospires allows for: (1) The ability to quantify an environmental disease risk to humans; (2) an estimate of pathogen type abundance which thereby allows for an understanding of pathogen community assembly; and (3) the ability to investigate thresholds of pathogen persistence in the environment. While Q RT-PCR will help little in elucidating the evolutionary relationships amongst leptospiral types of interest, this promising modern molecular technique should be useful for identifying the scale and magnitude of transmission patterns of pathogenic leptospires in the environment.

Cat! Has any role in zoonotic transmission of giardiasis in Tehran/Iran

M. Zarebavani[1], A.A. Farahvash[2], M. Rezaeian[3]

[1]*Tehran University of Medical Sciences, Allied Health Services, Tehran, Iran.*

[2]*Tehran University of Medical Sciences, Tehran, Iran.*

[3]*Tehran University of Medical Sciences School of Public Health and Institute of Public Health research, Tehran, Iran.*

In order to investigate the genotype of *Giardia duodenalis* from domestic and stray cats in Iran, feces were collected from 181 cats, screened by microscopy and examined by molecular method which included DNA extraction, triosephosphate isomerase gene amplification (PCR) and restriction fragment length polymorphism (RFLP). 21 cats were found to be infected with *Giardia* and harboring cysts or trophozoites belonging to assemblage A and subtypes of this assemblage. Also, the human isolate of *Giardia* in Iran has been clearly specified to be assemblage A and its subtypes. These findings suggest that infection of humans by zoonotic genotypes from domestic and stray cats and vice versa, is possible and could be of high significance. This is the first report on the genotype of *Giardia* isolate from cats in Iran.

Workshops/debates

Modern morphometrics, a cheap and advanced tool for medical entomology

J.P. Dujardin

IRD, UR 165 UMR 2724 GEMI, Mahidol University, Building 2, 999 Phuttamonthon 4 Road, Nakhon Pathom 7170, Thailand.

Morphological characters of medically important insects were usually described on a qualitative mode (color, texture, aspect, etc.) with limited quantification, and on a few specimens only. By reconsidering morphological characters quantitatively, at the population level, modern morphometrics brought new and important perspectives. Among them: (i) an identification tool challenging the accuracy of molecular tools, and (ii) a sensible detection tool of current population structure. In this latter application, modern morphometrics seems best suited to detect recent events affecting the population structure, including isolation or environmental changes. In addition, (iii) it is able to evaluate the level of adaptation of an insect to its local environment. Since it is cheap and does not require any special entomological skill, modern morphometrics should be the first line technique associated with entomological surveillance, an important epidemiological question not well negotiated as far.

Phylogeography

J.P Hugot, J.F. Cosson

How old is HIV?

J.P. Hugot, Fran Van Heuverswyn

Infections, Genetics and Evolution: Past and future of the journal. The new webportal Infections, Genetics and Evolution

M. Tibayrenc, B. Straub

Ecoepidemiology of *Aedes aegypti*

Philippe Barbazan

Characterization of a novel *IL7RA* mutation (444_450insA) caused marked reduction in CD127 expression highlighting an important role of interleukin-7 receptor α on T-cell development

Y. Monteerarat[1], O. Piboonpocanun[2], K. Sanpakit[3], W. Chinchang[3,1], K. Tachavanich[3], N. Visitsunthorn[2], V. Thongboonkerd[4], P. Vichyanond[2], V. Viprakasit[3]

[1]*Medical Scholar Programme, Department of Immunology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Thailand.*

[2]*Division of Allergy and Immunology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Thailand.*

[3]*Division of Hemato-Oncology, Department of Pediatrics, Faculty of Medicine, Siriraj Hospital, Mahidol University, Thailand.*

[4]*Medical Molecular Biology Unit, Office for Research and Development, Faculty of Medicine, Siriraj Hospital, Mahidol University, Thailand.*

Susceptibility to infectious diseases involved pathogen, environment and importantly, host defense mechanism combating exogenous factors. Molecular evidence addressing the important of human (host) genetics in predisposition to infectious pathogen was mainly derived from studies in primary immunodeficiency syndrome. Severe combined immunodeficiency (SCID) is one of the most common primary immunological defects in man characterized by blocking in T-cell development and/or functions with variable degree of simultaneous association of B-cell or natural killer (NK) cell dysfunction. Previous studies indicated that there was a genetic heterogeneity underlying SCID including mutation in *IL2RG, IL7RA, JAK3* and *CD3D* gene. Recently, we have molecularly characterized a 2 year-old girl presenting with BCGosis and recurrent serious bacterial infections including sepsis. This patient had distinctive immunological profiles of T^-, B^+, NK^+ SCID. Therefore, we firstly analyzed *IL7RA* and *CD3D* genes, which have been previously shown to cause such phenotype. We identified a novel adenine insertion at an adenine tract located between nucleotide 444–450 of *IL7RA* encoded CD127, resulting in a frameshift and premature stop codon. A quantitative real-time PCR analysis revealed that the relative mRNA expression were markedly reduce (0.037) in the patient and considerably decrease and her parents (0.119, 0.167) compared to normal (1, $n = 8$). This suggested that this mutation hampered mRNA expression, possibly, due to the non-sense mediated decay mechanism (NMD). Using flow cytometric analysis, we demonstrated that there was a significant reduction of CD127 positive-T cells in the patient confirming *in vivo* reduction at the protein level. Finally, this truncated CD127 protein was identified as expected by 2-Dimention electrophoresis-western blotting assay. Our study provided, for the first time, the molecular basis of SCID in Southeast Asian population and characterization of further cases will provide more insights on immunological mechanism controlling interaction between host and pathogen at last.

Specific mechanisms of genomic plasticity in the tick-borne Rickettsiale *Ehrlichia ruminantium*

Roger Frutos[1], Alain Viari[2], Nathalie Vachiery[1], Frédéric Boyer[2], Thierry Lefrançois[1] and Dominique Martinez[1]

[1]*Cirad, TA30/G, Campus International de Baillarguet, 34398 Montpellier Cedex 05, France.*

[2]*Inria Rhône-Alpes, Projet HELIX, 655 Av. de l'Europe, 38330 Montbonnot-Saint Martin, France.*

Ehrlichia ruminantium is the causative agent of a major tick-borne disease of livestock in Africa known as heartwater or cowdriosis. Three genomes corresponding to two different groups of differing phenotype, Gardel and Welgevonden, have been completely sequenced. The three strains display genomes of differing sizes with 1,499,920 bp, 1,512,977 bp and 1,516,355 bp. The genome organization is highly conserved with *A. marginale* whereas no synteny is conserved with the other *Rickettsiales*. 56 unique sequences and 19 truncated genes differentiate the two phenotypic groups but only 10 CDs are associated to major genomes rearrangements (i.e. deletions or extensive mutations). *E. ruminantium* displays a strong strand-specific compositional bias as well as a specific group of membrane proteins. *E. ruminantium* displays a strong GC bias resulting in the presence of different codon usages in leading and lagging strands. Moreover, *E. ruminantium* displays an active specific process of genome expansion/contraction targeted at tandem repeats in non-coding regions and based on the addition or removal of ca. 150 bp tandem units. Two populations of tandems repeats are present: short tandems displaying a period of 12–15 bp, associated to coding sequences, and long tandems displaying a period of 150 bp and associated to non-coding regions. The long tandems are affected by a GC bias whereas the short tandems are not. This specific mechanism of genome plasticity might be related to the low efficiency of vaccines in the field.