23 Gold, A.M., Despommier, D.D. and Buck, S.W. (1990) **Partial characterization of two antigens secreted by the larvae of *Trichinella spiralis*.** *Mol. Biochem. Parasitol.* 41, 187–196

24 Arasu, P. *et al.* (1994) **Molecular cloning of antigens targeted by protective antibodies in rapid expulsion of *Trichinella spiralis*.** *Mol. Biochem. Parasitol.* 65, 201–221

25 Su, X.Z., Prestwood, A.K. and McGraw, R.A. (1991) **Cloning and expression of complementary DNA encoding an antigen of *Trichinella spiralis*.** *Mol. Biochem. Parasitol.* 45, 331–336

26 Zarlenga, D.S. and Gamble, H.R. (1990) **Molecular cloning and expression of an immunodominant 53 kDa excretory-secretory antigen from *Trichinella spiralis* muscle larvae.** *Mol. Biochem. Parasitol.* 42, 165–174

27 Sugane, K. and Matsura, T. (1990) **Molecular analysis of the gene encoding an antigenic polypeptide of *Trichinella spiralis* infective larvae.** *J. Helminthol.* 64, 1–8

28 Vassilatis, D.K. *et al.* (1996) **Developmental expression of a 43 kDa secreted glycoprotein from *Trichinella spiralis*.** *Mol. Biochem. Parasitol.* 78, 13–29

29 Despommier, D.D. and Laccetti, A. (1981) ***Trichinella spiralis*: proteins and antigens isolated from a large-particle fraction derived from the muscle larvae.** *Exp. Parasitol.* 51, 279–295

30 Jasmer, D.P. (1990) ***Trichinella spiralis*: altered expression of muscle proteins in trichinosis.** *Exp. Parasitol.* 70, 452–465

31 Denkers, E.Y., Wassom, D.L. and Hayes, C.E. (1990) **Characterization of *Trichinella spiralis* antigens sharing an immunodominant carbohydrate-associated determinant distinct from phosphocholine.** *Mol. Biochem. Parasitol.* 41, 241–249

32 Wisnewski, N. *et al.* (1993) **Characterization of novel fucosyl-containing and tyvelosyl-containing glycoconjugates from *Trichinella spiralis* muscle stage larvae.** *Mol. Biochem. Parasitol.* 61, 25–35

33 Despommier, D.D. and Muller, M. (1976) **The stichosome and its secretion granules in the mature larva of *Trichinella spiralis*.** *J. Parasitol.* 62, 775–785

34 Takahashi, Y. and Araki, T. (1989) **Morphological study of the stichocyte granules of *Trichinella spiralis* muscle larvae.** *Jpn. J. Parasitol.* 38, 77–85

35 Sanmartin, M.L. *et al.* (1991) **Anatomical location of phosphocholine and other antigens on encysted *Trichinella* using immuno-histochemistry followed by Wheatleys' trichrome stain.** *Parasitol. Res.* 77, 301–306

36 Despommier, D.D. (1974) in Trichinella *and Trichinosis* (Kim, C., ed.), pp 239–254, Intext Educational Publications, New York

37 Capo, V., Silberstein, D.S. and Despommier, D.D. (1986) **Immunocytolocalization of two protection-inducing antigens of *Trichinella spiralis* during its enteral phase in immune and non-immune mice.** *J. Parasitol.* 72, 931–938

38 Parkhouse, R.M. and Almond, N.M. (1985) **Stage-specific antigens of *Trichinella spiralis*.** *Biochem. Soc. Trans.* 13, 426–428

39 Ortega-Pierrez, M.G. *et al.* (1996) **Workshop on a detailed characterization of *Trichinella spiralis* antigens: a platform for future studies on antigens and antibodies to this parasite.** *Parasite Immunol.* 18, 273–284

40 ManWarren, T.R. *et al.* (1997) **Invasion of intestinal epithelia *in vitro* by the parasitic nematode *Trichinella spiralis*.** *Infect. Immun.* 65, 4806–4812

41 Katz, F. (1960) **The oral transplantation of intestinal stages of *Trichinella spiralis*.** *J. Parasitol.* 46, 500–504

42 Rowlands, J.C. and Gustafsonn, J.A. (1997) **Aryl hydrocarbon receptor-mediated signal transduction.** *Crit. Rev. Toxicol.* 27, 109–134

43 Kageyama, R. *et al.* (1995) **Regulation of mammalian neural development by helix–loop–helix transcription factors.** *Crit. Rev. Neurobiol.* 9, 177–188

44 Muscat, G.E., Downes, M. and Dowham, D.H. (1995) **Regulation of vertebrate muscle differentiation by thyroid hormone: the role of the myoD gene family.** *BioEssays* 17, 211–218

45 Molkentin, J.D. and Olson, E.N. (1996) **Defining the regulatory networks for muscle development.** *Curr. Opin. Genet. Dev.* 6, 445–453

46 Simeonidis, S. *et al.* (1997) **Cloning and functional characterization of mouse IκBε.** *Proc. Natl. Acad. Sci. U. S. A.* 94, 14372–14377

47 Lee, D.L. *et al.* (1991) ***Trichinella spiralis*: antigenic epitopes from the stichocytes detected in the hypertrophic nuclei and cytoplasm of the parasitized muscle fibre (nurse cell) of the host.** *Parasitology* 102, 117–123

48 Jasmer, D.P. (1993) ***Trichinella spiralis* infected skeletal muscle cells arrest in G2/M and cease muscle gene expression.** *J. Cell Biol.* 121, 785–793

Beyond Strain Typing and Molecular Epidemiology:

Integrated Genetic Epidemiology of Infectious Diseases

M. Michel Tibayrenc

In the past 20 years, genetic and molecular methods for characterizing pathogen strains have taken a major place in modern approaches to epidemiology of parasitic and other infectious diseases. Here, Michel Tibayrenc explains the main concepts used in this field of research, with special emphasis on the approaches developed in his team, and suggests future avenues to explore.

The molecular epidemiology of pathogens remains a controversial field, viewed as a panacea by some people, but as a useless tool by others. According to the definition of the Centers for Disease Control (CDC) in Atlanta[1], molecular epidemiology means: 'the various techniques derived from immunology, biochemistry, and genetics for typing or subtyping pathogens'. This definition has the merit of clarity; nevertheless, in my opinion, it is too restrictive, for it overly emphasizes the technical side, to the detriment of theoretical considerations. In this article, I will make four important points: (1) analysis of the genetic polymorphism of pathogens reveals much more than strain typing for epidemiological follow-up; (2) the medical questions raised by genetic polymorphism are very similar, whatever kind of pathogen is considered, be it a parasitic protozoan, a fungus, a bacterium or even a virus or a helminth[2–4]; (3) the methods and concepts of evolutionary genetics (population genetics and phylogenetic analysis) are essential for understanding the origin and predictable properties of the genetic polymorphism of pathogens[2–4]; and (4) the future

Michel Tibayrenc is at the Centre d'Etudes sur le Polymorphisme des Microorganismes (CEPM), UMR CNRS/ORSTOM 9926, ORSTOM, BP 5045, 34032 Montpellier Cedex 1, France.
Tel: +33 4 67 41 62 07, Fax: +33 4 67 41 62 99,
e-mail: Michel.Tibayrenc@cepm.mpl.orstom.fr

  0169–4758/98/$19.00 PII: S0169-4758(98)01286-1

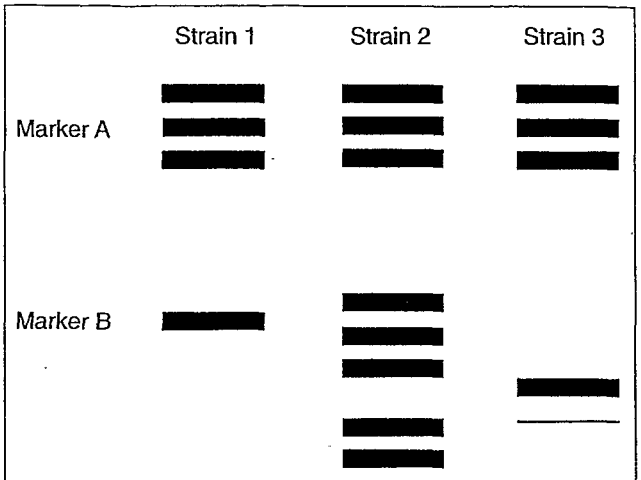


Fig. 1. Diagrammatic representation of the genetic diversity of three pathogen strains analysed for two different sets of markers (A and B). If only the variability of the marker A is considered, the three strains appear identical, and could wrongly be attributed to the same clonal descent; marker B reveals additional variability within this 'clonal lineage'.

of this field lies in an integrated approach, considering jointly the genetic polymorphism of the pathogen, the host and vector (in the case of vector-borne diseases), and their interactions ['the integrated genetic epidemiology of infectious diseases' (IGEID)][5–7].

The relevance of genetic characterization of pathogens to applied research

Genetic characterization of pathogens allows their identification at the strain (epidemiological tracking), subspecies and species (taxonomy) levels, and exploration of the impact of genetic diversity on relevant biomedical parameters such as virulence or resistance to drugs ('downstream studies')[3,4]. Identification of strains, subspecies and species is a well-known application, whereas downstream studies have not yet become the custom. Nevertheless, in my opinion, this is a very promising pay-off from pathogen genetic analysis. Most biomedical studies rely on a 'typologist' approach, where one or a few strains of a given pathogen are considered representative of the whole species. I believe that wide genetic diversity is commonplace in many pathogen species and is a major evolutionary strategy, and this is bound to have a profound impact on their medically relevant characteristics[3,4].

Why is evolutionary genetics indispensable?

Through two examples, I will show that a purely empirical analysis of pathogen genetic polymorphism can be very misleading. First, an intuitive approach to exploring the spread of theileriosis epidemics might be to characterize a sample of *Theileria* strains by a given set of molecular markers, for example by DNA fingerprinting. Those strains that appear identical using this molecular technique will be considered as tracking a given epidemiological spread. However, two important notions of evolutionary genetics are neglected in this empirical method. First, it infers, consciously or (more often) unconsciously, that the genotypes characterized by this DNA fingerprinting are stable enough to be tracked for months or years. Now, if genetic recombination is very frequent, the strains' genotypes will be very unstable, in proportion to the recombination rate. This explains the relevance of the sexuality/clonality debate[8–11] in applied microbiology. When parasites are considered, the risk is especially high for species that undergo an obligatory sexual cycle (*Plasmodium*, *Theileria*, etc.). The second trap concerns the level of resolution of the molecular method used, linked to the crucial notion of the 'molecular clock' (speed of evolution of the marker used). Let us accept that the genotypes characterized by a given marker are stable enough to permit epidemiological tracking. Is it ascertained that identical genotypes represent the same 'strain'? Probably not. If one takes a more discriminative method, 'identical' strains will probably prove to be no longer identical (Fig. 1).

Second, the general phylogenetic diversity of *Trypanosoma cruzi* might be explored by characterizing a set of strains by a given set of genetic markers, and a computer program might then evaluate the genetic relationships among the strains by counting the band mismatches between any two strains, to yield a 'dendrogram' as is shown in Fig. 2. Those strains that are genetically very similar will be clustered close to each other, whereas distantly related strains will be remotely clustered. One will infer that this dendrogram is a phylogenetic tree, representative of the evolutionary past of *T. cruzi*. The bad news is that such an approach developed without a careful analysis with appropriate phylogenetic methods can lead to disaster. Whatever the set of data considered, the computer will yield such a dendrogram. For example, if I take DNA samples from each of the members of my research group, and perform any method of DNA fingerprinting on them I will be able to generate such a dendrogram, which obviously has no phylogenetic significance and merely reflects the individual variability of the human species. The difference between the two examples (*T. cruzi* and the members of my research group) is that in the first case one infers the presence of discrete evolutionary lineages, whereas members of my group represent individual variants of the same evolutionary lineage, the human species. This crucial point cannot be gratuitously inferred and has to be verified by convenient evolutionary methods.

What, therefore, is the job of the geneticist?

The geneticist's task is to identify and delimit the convenient units of research, which I have proposed calling 'discrete typing units' (DTUs)[6,7], to illustrate the fact that they can be specifically labelled by convenient genetic markers or 'tags' (see Figs 2 and 3). These DTUs correspond to discrete evolutionary units ('clades' in the cladistic jargon) and the tags can be equated to 'synapomorphic characters' (in the cladistic jargon: derived characters that are shared by all members of a given clade or a given set of clades, and only by them – see 'Conceptual tools' below). In applied research, DTU analysis is relevant mostly at a microevolutionary scale[4,6,7], ie. around and under the species level. Above this level, there are no real problems of identification. For example, distinguishing *Plasmodium falciparum* from *Leishmania braziliensis* does not need a sophisticated evolutionary analysis, whereas trying to distinguish *Trypanosoma brucei gambiense* from *T. brucei rhodesiense* does.

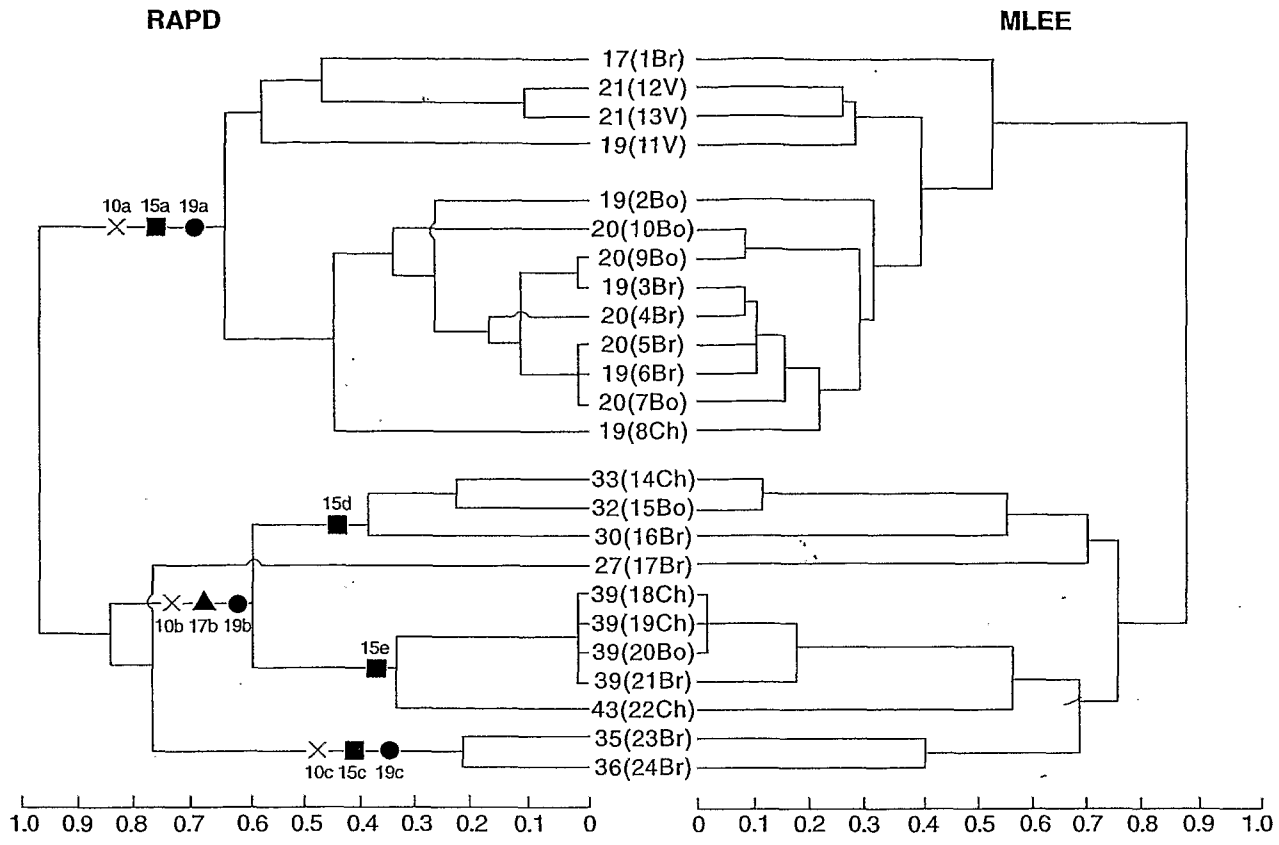


Fig. 2. Two dendrograms reflecting the genetic relationships assessed by isoenzymes (MLEE, right) and random primer amplified polymorphic DNA (RAPD, left) in *Trypanosoma cruzi*. Both isoenzymes and RAPDs show the existence of two main phylogenetic subdivisions[3,14] or 'discrete typing units' (DTUs)[6,7]. Some lesser DTUs can be seen on the RAPD dendrogram. Symbols indicate RAPD characters that correspond to DTU-specific 'tags'[6,7]. Fair agreement between the two dendrograms shows that isoenzyme and RAPD variability are linked in *T. cruzi*, which is a manifestation of linkage disequilibrium (see text). Moreover, this supports the view that the two dendrograms can be equated to real phylogenetic trees (modified from Ref. 40).

Confronted with a given species, first the geneticist must see whether this species corresponds to a real DTU, with clear-cut borders and specific tags. In other words: is the species monophyletic (composed of only one phylogenetic line whose members all share a common and exclusive ancestor), and is it identifiable by specific synapomorphic characters? Second, the geneticist must look for lesser DTUs within the species, and go down as far as possible to lower levels of phylogenetic divergence, so that epidemiologists, clinicians, immunologists, vaccine and drug designers, and so on have more discrete units to work on. The geneticist must look for tags (synapomorphic characters) for the whole species (when it is monophyletic and therefore corresponds to a unique DTU) and the lesser DTUs that subdivide it. To ascertain the reliability of them, it is crucial that the search for specific DTU tags comes after complete phylogenetic analysis, and not in a blind manner[7]. The geneticist

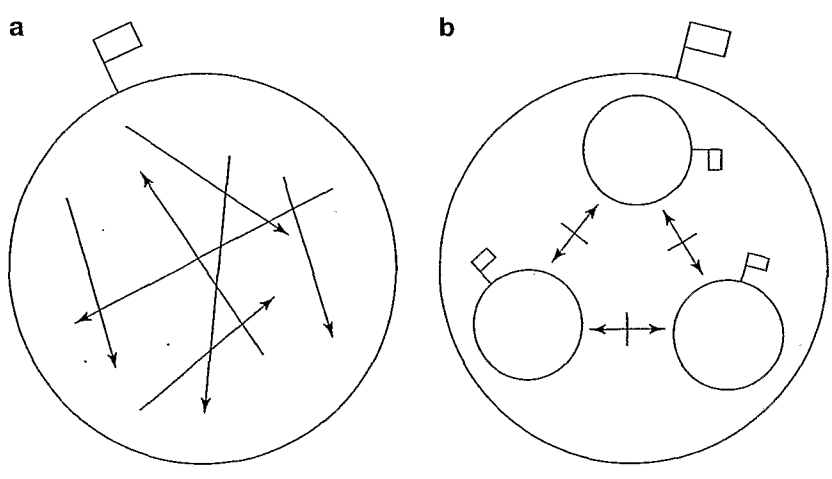


Fig. 3. Diagrammatic representation of nonstructured (a) versus structured (b) models of pathogen populations[3]. In the nonstructured model, genetic exchange (symbolized by arrows) is frequent enough to prevent the maintenance of discrete genetic lines. The operational unit of research is the whole species only. In contrast, in the structured model, the species is subdivided into discrete genetic entities between which gene flow (symbolized by double arrows) is inhibited. In the two models, the species are supposed to be monophyletic and, therefore, correspond to 'discrete typing units' (DTUs)[6,7] that can be specifically identified by appropriate genetic markers or 'tags' (symbolized by small flags). In the structured model, the species is further subdivided into lesser DTUs, each characterized by specific tags. These lesser DTUs can be taken as units of research as well as the whole species.

must then explore the possible biomedical differences (virulence, resistance to drugs, etc.) which can be exhibited by the DTUs that structure the species under study, and analyse the genetic background of these differences. Lastly, under the level of phylogenetic divergence that makes it possible to identify DTUs reliably, the geneticist can look for statistical indications of clonality with the means of population genetics (linkage disequilibrium analysis – see below). Indeed, even if clear DTUs cannot be delimited, it remains important for epidemiological tracking to see whether the genotypes of the pathogen under study are stable (propagate clonally) or unstable (undergo very frequent genetic recombination). An unanswered question here is the level of resolution required to have a fair presumption of common clonal descent between two isolates. Obviously, many studies that rely on a limited set of genetic markers run the risk of very misleading conclusions (wrong clonal identity).

The huge task thus summarized obviously has to be carried out in close collaboration with other specialists. Field epidemiologists provide the field data indispensable for a reliable interpretation of genetic variability. Biochemists, immunologists, molecular biologists, and so on, are highly involved in downstream studies (virulence, resistance to drugs, etc.). Evolutionary analysis, by clearly identifying the relevant units of analysis (DTUs), provides the framework for these downstream biomedical studies, at least for biologists who want to: (1) work on reliably delimited species; and (2) take into account a crucial parameter that is too often neglected, the genetic variability of pathogenic agents.

Conceptual tools

Two complementary approaches are available for the geneticist's job: population genetics and phylogenetic analysis[3,4,6–9]. Population genetics gives a 'snapshot' of the present population structure of the species under study and the level of gene flow (or absence of gene flow) among populations. Population genetics of pathogens has focused mainly on the clonality/sexuality debate[8–11] for the reasons given above. All statistics take a panmictic situation (in which genetic exchange occurs at random) as the null hypothesis. Statistical departures from panmictic expectations are then taken as an indication that gene flow is inhibited or suppressed in the population under study. This can be due to physical obstacles (isolation by time or geographical distance or both) or biological obstacles (either cryptic biological speciation or clonality). The main statistical approach used is linkage disequilibrium (LD) or nonrandom association of genotypes occurring at different loci (Fig. 2). To take a very simple example, in a human population, if the observed frequency of the AB blood group is 0.5 and the observed frequency of the Rh+ group is 0.3, the expected frequency of the double genotype AB/Rh+ is $0.5 \times 0.3 = 0.15$, as AB and Rh blood groups are governed by distinct genes that recombine independently. If this frequency is significantly higher, this means that the two genes are 'linked' (exhibit LD). LD analysis gives a presumption (proportional to the levels of statistical significance) that: (1) the species under study is structured into DTUs; and (2) under the level of reliably identifiable DTUs, a given population is clonal rather than sexual.

Phylogenetic analysis actually considers similar problems, but on a much broader time-scale. It explores the past of the species under study on an evolutionary scale. Specific approaches have been developed, the best known of which is cladistics[12]. This method is based on the distinction between ancestral (plesiomorphic) and derived (apomorphic) characters. Only those derived characters that are specifically shared by several distinct phylogenetic lines (synapomorphic characters) convey reliable phylogenetic information.

Traps and difficulties

The traps and difficulties are numerous and dangerous, even if a purely empirical approach is discarded. A few notable examples will be discussed here. When population genetics is considered, LD merely indicates that there is an obstacle to gene flow. First, it is necessary to assess whether this could simply be due to physical obstacles (time and/or geographical separation)[9]. If this is not the case, it must be determined whether LD is due to: (1) 'epidemic clonality' (ephemeral appearance of unstable clones in a basically sexual species); (2) long-term clonal evolution; or (3) presence of cryptic biological species[10]. In my opinion[3,4], the relevant border is actually between nonstructured species (sexual species with or without epidemic clonality) and structured species (with either clonal evolution or cryptic speciation). Stable DTUs are seen only in structured species (Fig. 3).

When exploring the notions of clonality and DTUs, one has to think of the possibility of 'non black and white' situations. In many pathogens, it appears that genetic recombination can occur at a limited level, on an evolutionary scale. This is enough to make the task of the geneticist more difficult. The null hypothesis of panmixia is appropriate to settle the hypothesis of preponderant clonality and to indicate the existence of DTUs. However, it is probably not strict enough to establish this existence firmly[7].

In the case of phylogenetic analysis, the use of genetic and molecular characters for phylogenetic studies remains controversial. As an example, cladistics[12] has been designed mainly for morphological characters and must be used cautiously with genetic methods.

Lastly, population genetics and phylogenetic analysis are based on statistics, at least to some extent. Like any statistical method, they are subject to a statistical 'type II' error[3] (impossibility of rejecting the null hypothesis, not because it is verified, but because the test is not powerful enough owing to insufficient information). For example, impossibility of rejecting the null hypothesis of panmixia is frequently taken as evidence that the species is sexual[10]. Obviously, this depends on the richness of information available (number of strains, amount of genetic diversity).

Some notable results

Many informative results are available concerning the population structure and evolution of pathogens. There would have been much more if the ocean of data, generated by strain typing studies based on an empirical approach, had been analysed in evolutionary genetic terms.

In the case of parasitic protozoa, the clearer results concern the genera *Trypanosoma* and *Leishmania*. Only a few examples are discussed here. *Trypanosoma cruzi*, the agent of Chagas disease, constitutes a clear example of clonal evolution[13]. *Trypanosoma cruzi* natural clones are distributed in two main phylogenetic lineages (upper DTUs; Fig. 2)[3,14]. Lesser DTUs can be identified within each of the main DTUs (S. Brisse, pers. commun.; see also Fig. 2). *Trypanosoma cruzi* exhibits a strong link between phylogenetic diversity and biomedical variability analysed by experiments (growth in culture, virulence in mice, *in vitro* drug resistance, transmissibility by triatomine bugs)[3,15]. It has been proposed that the different biological properties of *T. cruzi* natural clones include variation in tropism for human target tissues and that this heterogeneity may have significant pathogenetic relevance for Chagas disease ('clonal-histotropic model')[16]. Despite predominant clonal evolution, there are some indications for hybridization events in *T. cruzi*[17] (S. Brisse, pers. commun.).

Evidence for structuration and DTUs are much less clear in the case of *T. brucei*, although this species shows obvious evidence of clonality[8]. *Trypanosoma brucei gambiense* 'group 1' (Ref. 18) seems to correspond to a reliable DTU, identifiable by a specific tag, namely a kinetoplast DNA probe[19], whereas there is more dispute as to whether *T. brucei rhodesiense* can be considered a DTU[20]. Lastly, the stability of *T. brucei* natural clones in the long term is still under debate[21,22] and it is possible that the impact of genetic recombination is not the same in different populations of this parasite[22].

Speciation in the genus *Leishmania* has been widely explored by phylogenetic methods. It is interesting to note that many *Leishmania* species formerly described from eco-epidemiological criteria can be equated to reliable DTUs[23,24]. However, it is not clear whether the epidemiological and clinical specificities classically inferred for these *Leishmania* species will be verified for all strains of a given DTU. In terms of population structure, *Leishmania* appear to be basically clonal[8], although clear indications for natural hybrids have been recorded in both Old World[25] and New World *Leishmania*[26,27].

The population structure of *Plasmodium falciparum* has been, and still is, under hot debate[2,8,9,28–31]. Some kind of uniparental propagation, probably as a result of self-fertilization, could take place in certain ecosystems[31]. Nevertheless, no indication for clear-cut DTUs within *P. falciparum* has been recorded so far, and it is apparent that *P. falciparum* is not a case of long-term clonal evolution.

The population structure and phylogenetic diversity of *Toxoplasma gondii*, *Entamoeba histolytica* and *Giardia intestinalis* show interesting similarities[32]. These three species exhibit high levels of LD, and have been inferred to be clonal[8,9,33,34]. Moreover, they are clearly subdivided into two or more DTUs. In the case of *T. gondii*, one of the DTUs is linked to virulence in mice[33], whereas in the case of *E. histolytica*, one of the DTUs is linked to virulence in humans[35,36]. However, owing to limited genetic variability within each of the DTUs and the problem of type II error caused by this[3], it is difficult to decide whether the DTUs of these three species correspond to either clonal lineages or cryptic biological species[32].

Many other pathogens have been analysed extensively with evolutionary genetic methods, and it is impossible to summarize all these results here. The bulk of the literature concerns bacteria[11]. Although a clonal population structure is a common picture in bacteria, this 'clonal paradigm' has been challenged in the case of several species[10], including *Neisseria gonorrhoeae* and *Bacillus subtilis*. In the case of bacteria, geneticists have focused more on population genetic problems. Much remains to be done in terms of intraspecific phylogenetic analysis and the search for DTUs[7].

An informative model: population structure of HIV-associated pathogens

HIV infection provides a natural 'experiment' to evaluate the impact of immune defenses on the genetic diversity and population structure of pathogens. The first results obtained from co-infections with HIV and either *L. infantum*[37] or *T. cruzi* (L. Perez, pers. commun.) indicate that the parasite's population structure is not drastically modified by immunodepression.

The usefulness of a comparative approach

The results summarized above illustrate the value of evolutionary studies in applied research on parasitic and other infectious diseases. Whatever the kind of pathogen considered, be it a parasitic protozoan, a fungus, a bacterium, or even a virus or a helminth, of either medical, veterinary or agronomical interest, it is necessary to define and identify strains, subspecies and species in a reliable manner to give a rational framework for downstream biomedical studies. Only an evolutionary genetic approach can do that, through the notions of DTUs[6,7] and clonality[8–11]. A synthetic and comparative approach[2–4], involving common conceptual and technical tools, will boost progress in this field of research by: (1) enlightening the specificities of each model and at the same time, the common laws that govern the evolution of pathogens; and (2) making it possible to design versatile tools that can be used whatever the pathogen considered.

Avenues for the future

Evolutionary studies on the host, the pathogen, and in the case of vector-borne diseases, the vector, have developed separately so far. Each of these fields has made impressive progress in recent years and is now ready to undergo a higher level of integration. The new synthetic approach, which I have called integrated genetic epidemiology of infectious diseases (IGEID)[5–7], will make it possible to take into account jointly the respective impact of the genetic variability of the host, the pathogen and the vector on the transmission and pathogenicity of infectious diseases. More specifically, it will be possible to focus on the interactions between these three parameters, in other words, to explore in depth the coevolution phenomena of the three links of the epidemiological transmission chain. From an evolutionary point of view, these three links can be fruitfully equated to specific characters of each other.

In the case of the host, genetic susceptibility to infectious diseases has been demonstrated in several models, including parasite infections[38,39]. This susceptibility may not be the same for different strains of the same pathogen species, especially in those pathogen species with considerable strain diversity, such as *T. cruzi*[13].

Finally, different populations of the same species of vector, or different vector species for the same species of pathogen, could have different transmissibility potential and different impacts on strain virulence.

A benefit of IGEID would be to encourage distinct populations of geneticists to work together and to compare their tools in order to favour a desirable synergy towards convergent goals. For example, analyses of LD specifically designed for pathogens[8,9] could be used to explore the links between genetic background and susceptibility to infectious diseases in humans, or to analyse the population structure of vectors and hosts.

Chagas disease constitutes an appropriate model to develop an IGEID approach. As a matter of fact, many data are already available on the vector's and the parasite's genetic variability, and *T. cruzi* presents clear-cut DTUs, which is not the case for *P. falciparum*, for example. Moreover, a complete chagasic cycle is very easy to maintain experimentally, which is the case neither for African trypanosomes, nor for *Leishmania* or *Plasmodium*. An IGEID approach could develop the following complementary lines of research. In experiments[15], the respective impact of the genetic variability of the vector, the parasite and the host (laboratory animals) could be conveniently explored by keeping two of these parameters fixed, allowing only the third to vary. In the field, again, these three parameters should be explored, and the data should be analysed in the same way (one parameter allowed to vary, the two other kept as homogeneous as possible). When humans are considered, susceptibility to Chagas disease and to its different clinical forms should be analysed, at the level of both the individual and the population (either within or between ethnic groups), for, in some cases, the population rather than the individual will represent the relevant evolutionary unit. Lastly, when using genetic variability as an epidemiological marker, one would obtain much more powerful information if the variability of both triatomine bugs and *T. cruzi* were taken into account in the same time, in the same area. Evolution of the vector and of the parasite and the spread of both populations occur at different speeds. Therefore, this double information could constitute a valuable 'epidemiological magnifying glass'. Needless to say, these approaches illustrated through the example of Chagas disease are fully applicable to any infectious disease.

In this perspective, the compared approach advocated for evolutionary genetics of pathogenic microorganisms[2–4] should be extended to IGEID. Particular cases enlighten each other, and make it possible to clarify both specificities and common features. For example, there is little doubt that comparisons of Chagas disease and leishmanioses integrated genetic epidemiologies will help delimit the respective role of the parasite's, the host's and the vector's genetic variabilities and of their interactions for both diseases. Substantial conceptual progress and effort savings can be hoped for from IGEID, whose bases can be summarized as follows: (1) broad use of evolutionary genetic concepts; (2) integration of evolutionary genetics of pathogens, vectors and hosts. Emphasis on phenomena of interaction and coevolution; (3) better delimitation of the relevant evolutionary units. In the case of pathogens: DTUs. In the case of vectors and hosts: either individuals or populations; and (4) comparative approach involving different kinds of infectious diseases.

References

1 Centers for Disease Control (1994) in *Addressing Emerging Infectious Disease Threats. A Prevention Strategy for the United States*, p. 27
2 Tibayrenc, M. and Ayala, F.J. (1991) **Towards a population genetics of microorganisms: the clonal theory of parasitic protozoa.** *Parasitol. Today* 7, 228–232
3 Tibayrenc, M. (1995) **Population genetics of parasitic protozoa and other microorganisms.** *Adv. Parasitol.* 36, 47–115
4 Tibayrenc, M. (1996) **Towards a unified evolutionary genetics of microorganisms.** *Annu. Rev. Microbiol.* 50, 401–429
5 Tibayrenc, M. (1997) in *Molecular Epidemiology and Evolutionary Genetics of Pathogenic Microorganisms*, 2nd International Workshop CDC/ORSTOM/CNRS, pp 9–10
6 Tibayrenc, M. *Analysis of Microbial Population Structures – an Evolutionary Approach, Medical Mycology Protocols, Methods in Molecular Medicine Series* (Haynes, K.A., ed.), Humana Press (in press)
7 Tibayrenc, M. (1998) **Genetic epidemiology of parasitic protozoa and other infectious agents: the need for an integrated approach.** *Int. J. Parasitol.* 28, 85–104
8 Tibayrenc, M., Kjellberg, F. and Ayala, F.J. (1990) **A clonal theory of parasitic protozoa: the population structure of *Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas* and *Trypanosoma*, and its medical and taxonomical consequences.** *Proc. Natl. Acad. Sci. U. S. A.* 87, 2414–2418
9 Tibayrenc, M. *et al.* (1991) **Are eukaryotic microorganisms clonal or sexual? A population genetics vantage.** *Proc. Natl. Acad. Sci. U. S. A.* 88, 5129–5133
10 Maynard Smith, J. *et al.* (1993) **How clonal are bacteria?** *Proc. Natl. Acad. Sci. U. S. A.* 90, 4384–4388
11 Lenski, R. (1993) **Assessing the genetic structure of microbial populations.** *Proc. Natl. Acad. Sci. U. S. A.* 90, 4334–4336
12 Hennig, W. (1966) *Phylogenetic Systematics*, University of Illinois Press
13 Tibayrenc, M. *et al.* (1986) **Natural populations of *Trypanosoma cruzi*, the agent of Chagas disease, have a complex multiclonal structure.** *Proc. Natl. Acad. Sci. U. S. A.* 83, 115–119
14 Souto, R.P. *et al.* (1996) **DNA markers define two major phylogenetic lineages of *Trypanosoma cruzi*.** *Mol. Biochem. Parasitol.* 83, 151–162
15 Laurent, J.P. *et al.* (1997) **Impact of clonal evolution on the biological diversity of *Trypanosoma cruzi*.** *Parasitology* 114, 213–218
16 Macedo, A.M. and Pena, S. (1998) **Genetic variability of *Trypanosoma cruzi*: implications for the pathogenesis of Chagas disease.** *Parasitol. Today*, 14, 119–124
17 Carrasco, H.J. *et al.* (1996) **Genetic exchange as a possible source of genomic diversity in sylvatic populations of *Trypanosoma cruzi*.** *Am. J. Trop. Med. Hyg.* 54, 418–424
18 Gibson, W. (1986) **Will the real *Trypanosoma brucei gambiense* please stand up?** *Parasitol. Today* 2, 255–257
19 Mathieu-Daudé, F. *et al.* (1994) **Identification of *Trypanosoma brucei gambiense* group I by a specific kinetoplast DNA probe.** *Am. J. Trop. Med. Hyg.* 50, 13–19
20 Hide, G. *et al.* (1990) **The identification of *Trypanosoma brucei* subspecies using repetitive DNA sequences.** *Mol. Biochem. Parasitol.* 39, 213–226
21 Stevens, J. and Tibayrenc, M. (1996) ***Trypanosoma brucei* s.l.: evolution, linkage and the clonality debate.** *Parasitology* 112, 481–488
22 Hide, G. *et al.* (1994) **Epidemiological relationships of *Trypanosoma brucei* stocks from South East Uganda: evidence for different population structures in human infective and non-human infective isolates.** *Parasitology* 109, 95–111
23 Lanotte, G., Rioux, J.A. and Serres, E. (1986) Approche cladistique du genre *Leishmania* Ross, 1903. A propos de 192 souches de l'Ancien Monde. Analyse numrique de 50 zymodmes identifis par 15 enzymes et 96 isoenzymes. *Leishmania.* Taxonomie et phylogénèse. Applications éco-épidémiologiques (Coll. Int. CNRS/INSERM, 1984), pp 27–40, IMEEE
24 Thomaz-Soccol, V. *et al.* (1993) **Phylogenetic taxonomy of New World *Leishmania*.** *Ann. Parasitol. Hum. Comp.* 68, 104–106
25 Evans, D.A. *et al.* (1987) **Hybrid formation within the genus *Leishmania*?** *Parassitologia* 29, 165–173
26 Darce, M. *et al.* (1991) **Etiology of human cutaneous leishmaniasis in Nicaragua.** *Trans. R. Soc. Trop. Med. Hyg.* 85, 58–59
27 Bañuls, A.L. *et al.* (1997) **Evidence for hybridization by multilocus enzyme electrophoresis and random amplified polymorphic DNA between *Leishmania braziliensis* and *Leishmania panamensis/guyanensis* in Ecuador.** *J. Euk. Microbiol.* 44, 408–411
28 Dye, C. (1991) **Population genetics of nonclonal, nonrandomly mating malaria parasites.** *Parasitol. Today* 7, 236–240

29 Walliker, D. (1991) **Malaria parasites: randomly interbreeding or 'clonal' populations?** *Parasitol. Today* 7, 232–235
30 Tibayrenc, M. and Lal, A. (1996) **Self-fertilization, linkage disequilibrium and strain in *Plasmodium falciparum*.** *Science* 271, 1300
31 Paul, R.E.L. *et al.* (1995) **Mating patterns in malaria parasite populations of Papua New Guinea.** *Science* 269, 1709–1711
32 Tibayrenc, M. (1993) ***Entamoeba, Giardia* and *Toxoplasma*: clones or cryptic species?** *Parasitol. Today* 9, 102–105
33 Sibley, L.D. and Boothroyd, J.C. (1992) **Virulent strains of *Toxoplasma gondii* comprise a single clonal lineage.** *Nature* 359, 82–85
34 Meloni, B.P., Lymbery, A.J. and Thompson, C.A. (1989) **Characterization of *Giardia* isolates using a non-radiolabelled DNA probe, and correlation with the results of isoenzyme analysis.** *Am. J. Trop. Med. Hyg.* 40, 629–637
35 Sargeaunt, P.G., Williams, J.E. and Greene, J.D. (1978) **The differentiation of invasive and non-invasive *Entamoeba histolytica* by isoenzyme electrophoresis.** *Trans. R. Soc. Trop. Med. Hyg.* 72, 519–521
36 Blanc, D.S. (1992) **Determination of taxonomic status of pathogenic and nonpathogenic *Entamoeba* zymodemes using isoenzyme analysis.** *J. Protozool.* 39, 471–479
37 Jiménez, M., Alvar, J. and Tibayrenc, M. (1997) ***Leishmania infantum* is clonal in AIDS patients too; epidemiological implications.** *AIDS* 11, 569–573
38 Blackwell, J.M. (1996) **Genetic susceptibility to leishmanial infections: studies in mice and man.** *Parasitology* 112, S67–S74
39 Hill, A.V. (1996) **Genetic susceptibility to malaria and other infectious diseases: from the MHC to the whole genome.** *Parasitology* 112, S75–S84
40 Tibayrenc, M. *et al.* (1993) **Genetic characterization of six parasitic protozoa: parity of random-primer DNA typing and multilocus isoenzyme electrophoresis.** *Proc. Natl. Acad. Sci. U. S. A.* 90, 1335–1339

The Flip-side of Cytoadherence: Immune Selection, Antigenic Variation and the *var* Genes of *Plasmodium falciparum*

C.J. Sutherland

In areas where Plasmodium falciparum *is endemic, the natural immunity acquired by people exposed to frequent malaria infection is likely to have a differential selective impact upon different parasite genotypes. It has been suggested that the immune response directed against the variant antigen PfEMP1, which is expressed on the infected erythrocyte surface, is a crucial determinant of parasite population structure and favours the existence of distinct strains, or Varotypes. Here, Colin Sutherland summarizes current knowledge of the* var *multigene family, which encodes the PfEMP1 variants, and suggests that this information may allow certain predictions of the strain hypothesis to be tested directly.*

The ability of certain species of *Plasmodium* to cytoadhere enables infected erythrocytes to sequester in the microvasculature of a number of host organs and thus evade splenic clearance. This requires the expression of parasite-encoded cytoadherent molecules on the surface of the infected red blood cell (IRBC), which effectively blows the parasite's cover and presents the host immune system with exposed surface antigens[1]. PfEMP1, a cytoadherent molecule expressed on the surface of human erythrocytes infected with *P. falciparum*, is also a variant antigen in the classical sense: immunologically distinct variants of the antigen are expressed sequentially and reversibly during expansion of a clonal parasite population in experimental models *in vivo*[2] and *in vitro*[3,4]. Therefore, in order to benefit from one survival strategy, cytoadherence, the malaria parasite requires a second, namely antigenic variation, to minimize the host immune response against the cytoadherent neoantigens on the IRBC surface.

Colin J. Sutherland is at the Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London, UK WC1E 7HT.
Tel: +44 171 636 8636, Fax: +44 171 436 5389, e-mail: c.sutherland@lshtm.ac.uk

The genetic basis of antigenic variation: *var* genes

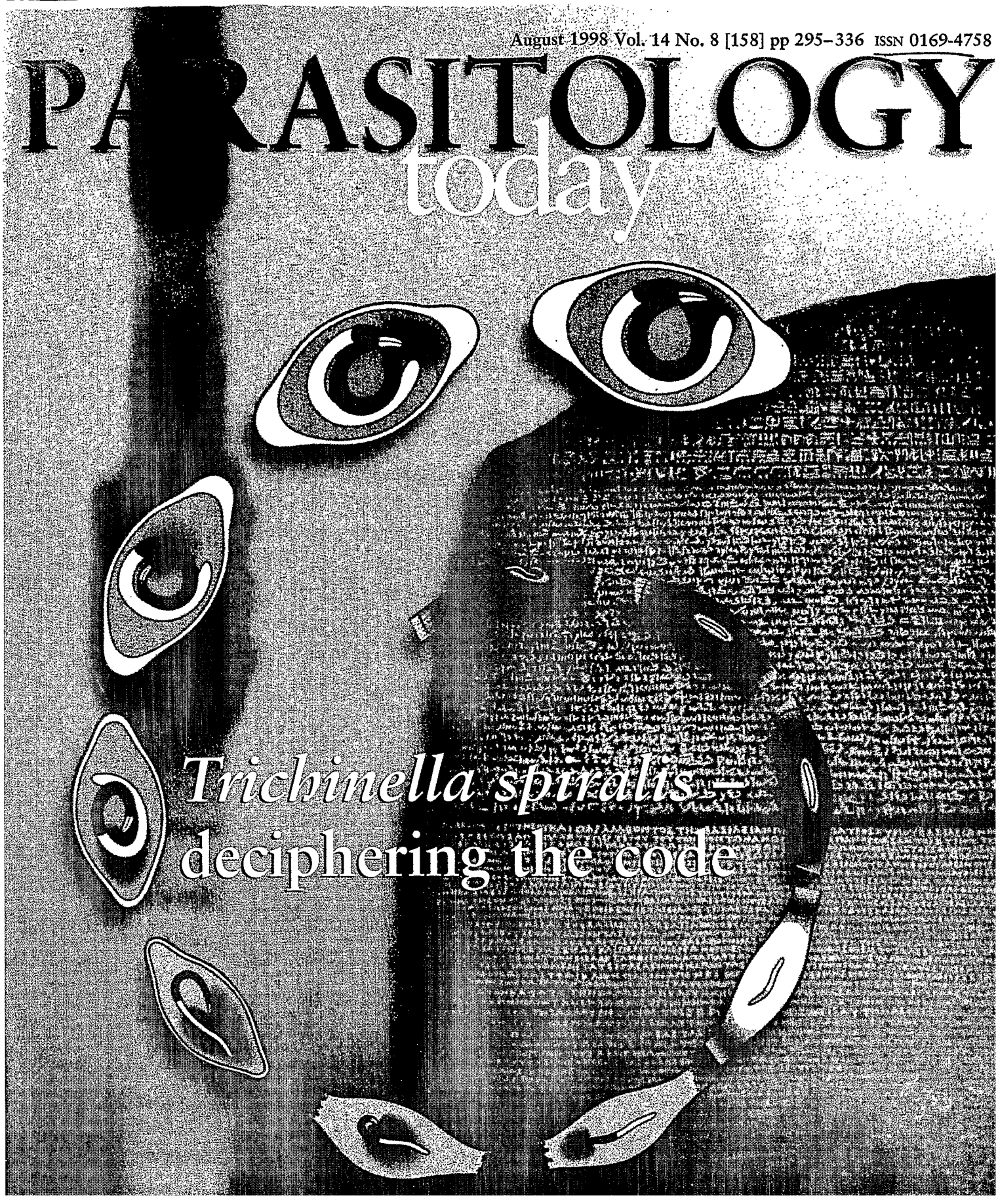
The identification of the genes encoding PfEMP1 has demonstrated that this antigen is actually an extensive family of antigenically and functionally distinct proteins encoded by the large *var* multigene family[4–6]. Probes encoding the relatively conserved acidic terminal segment (ATS) of PfEMP1 hybridize to a large number of loci that are spread throughout the genome on most, but not all, chromosomes[5,7–9]. These loci include the Pf60-related sequences, which comprise a divergent branch of the *var* multigene family[5,8,10,11]. All full-length *var* genes characterized to date share a conserved organization[5,7,12] (see Box 1), yet are highly diverse at the amino acid sequence level[5,11–13]. Thus, although there may be some functional constraint upon surface-exposed cytoadherent domains, as suggested by both the crossreactivity of PfEMP1 phenotypes of parasites from distant geographical locations[14] and the finite number of endothelial receptors for PfEMP1 (Ref. 15), the total number of distinct variants that exist appears to be immense.

Current work in several laboratories is taking advantage of recently developed yeast artificial chromosome (YAC) contigs of *P. falciparum* clone 3D7 (Ref. 16) to describe the chromosomal location and organization of *var* loci. Detailed analyses of chromosomes 4 and 12 (Ref. 8), 2 and 8 (Ref. 9), and 9 and 13 (Ref. 17; D. Holt *et al.*, unpublished) have demonstrated that

  0169–4758/98/$19.00 PII: S0169-4758(98)01276-9

August 1998 Vol. 14 No. 8 [158] pp 295–336 ISSN 0169-4758

PARASITOLOGY today



Trichinella spiralis: deciphering the code