Nuclear rDNA-based Molecular Clock of the Evolution of Triatominae (Hemiptera: Reduviidae), Vectors of Chagas Disease


Departmento de Parasitología, Facultad de Farmacia, Universidad de Valencia, Av. Vicente Andrés Estellés s/n, 46100 Burjassot, Valencia, España *Centro de Investigaciones sobre Enfermedades Infecciosas, Instituto Nacional de Salud Pubblica, Coenavaca, Morelos, 62508 México **UMR IRD/ORSTOM-CNRS 9926, Institut de Recherche pour le Développement, Montpellier Cedex 1, France ***Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1 E7HT, UK

The evolutionary history and times of divergence of triatomine bug lineages are estimated from molecular clocks inferred from nucleotide sequences of the small subunit SSU (18S) and the second internal transcribed space (ITS-2) of the nuclear ribosomal DNA of these reduviids. The 18S rDNA molecular clock rate in Triatominae, and Prosoorhynchian Hemiptera in general, appears to be of 1.8% per 100 million years (my). The ITS-2 molecular clock rate in Triatominae is estimated to be around 0.4-1% per 1 my, indicating that ITS-2 evolves 23-55 times faster than 18S rDNA. Inferred chronological data about the evolution of Triatominae fit well with current hypotheses on their evolutionary histories, but suggest reconsideration of the current taxonomy of North American species complexes.

Key words: Chagas disease - Triatominae vectors - nuclear rDNA - 18S gene - ITS-2 spacer - molecular clock - evolution

Until the 60s, analysis of fossils was the only source of information about the time when ancestors of extant organisms lived. Since then, molecular genetic studies have provided a series of so-called molecular clocks, that may be used to estimate the evolutionary history and time of divergence of organisms. This has become particularly useful for species that have a poor fossil record, as well as for lineages of organisms which show little morphogenesis through time, or which exhibit a great deal of phenotypic variability.

The molecular clock is based on the assumption that processes such as DNA replication, transcription, protein synthesis, and metabolism are similar in all organisms and that the proteins and RNAs that carry out these "housekeeping functions" should be highly conserved. Over time, nucleotide substitutions in housekeeping genes have occurred and DNA and protein sequences have changed. The changes have tended to preserve the function of the gene rather than modify it, so that changes in the DNA code may occur that have little or no effect on function. For example, changes in DNA sequences can occur but, because of the degenerate genetic code, the third codon base can be altered without affecting which amino acid is designated. Changes in the code also can occur without changing the function of the protein if amino acid changes occur in the region(s) that do not affect the function of the molecule, or one amino acid is replaced by a similar amino acid.

The molecular clock hypothesis assumes that nucleotide substitutions in a given sequence occur at a constant rate to provide a method for measuring time of divergence from the differences between two sequences (Zuckerkandl & Pauling 1965). Phylogenetic history thus can be reconstructed, and the time when branches occurred among the lineages leading to modern species can be resolved.

However, evidence from analyses of different protein sequences shows that rates of change do vary among different proteins and lineages. For example, cytochrome-c has an acceptable clock-like behaviour over the span of organisms studied,
while copper-zinc superoxide dismutase behaves as a more erratic clock (Ayala 1986). The genetic code makes it possible to calculate the minimum number of nucleotide substitutions that would be required to change a codon for one amino acid to a codon for another. Some amino acid replacements can be made with the change of a single nucleotide, but others would require a minimum of two or three nucleotide substitutions (Wilson et al. 1977). The molecular clock should therefore be calibrated with data that are independently derived, preferably with fossil evidence (Wilson et al. 1987).

The information furnished by genes encoding small subunit rRNAs (16S and 18S rDNA in prokaryotes and eukaryotes, respectively) allows the evaluation of relatively ancient relationships (>100 million years ago — mya) (Woese 1987, Sogin 1989, Hillis & Dixon 1991, Wainwright et al. 1993). In a study of the evolutionary rates of the 18S rRNA gene, Philippe et al. (1994) concluded that rapid adaptive radiations spanning less than 40 million years (my) are generally beyond the limits of resolution of the entire 18S gene. In these genes, overall base-substitution rates (changes of nucleotide sites through time) are relatively slow. Such rates have been used as a molecular clock for estimating time of divergence of ancient lineages, especially among prokaryotes (Ochman & Wilson 1987). However, different regions within some rRNAs show varying temps of base substitution (Forsgren & Cluster 1988). Consequently, different portions within rRNA genes can be used to infer different levels of phylogenetic relationships. For instance, the 18S rDNAs of certain Hemiptera of the suborder Sternorrhyncha include large expansion regions that have relatively fast base-substitution rates (Kwon et al. 1991, Campbell et al. 1994). Differences in nucleotide sequences in such regions can be useful for resolving affiliations of closely related taxa that have diverged relatively recently (<10 mya) (Campbell et al. 1995a,b).

Molecular clocks used to estimate times of divergence can be based on rates of base substitution. The conventional estimate for base-substitution rates of prokaryotic 16S rDNAs is 1.8 x 10^-10 substitutions per site per year (1.8% per 100 my) (Ochman & Wilson 1987) or 2-4% per 100 my using aphid bacterial symbionts and the aphid fossil record (Moran et al. 1993). These estimates have been suggested to apply to eukaryotic 18S rDNA substitution rates, although rates amongst eukaryotes are much more variable than for bacteria (Van De Peer et al. 1993). According to Escalante and Ayala (1995), an evolutionary rate of 0.8% substitutions per site per 100 my is obtained when comparing 18S rRNA gene sequences from increasingly divergent multicellular organisms for which reasonably accurate times of divergence are known.

The evolutionary rate of 0.8% sequence divergence per 100 my for the 18S rDNA is thought to reflect the rate of the whole 18S rDNA, whereas the rate of 2% per 100 my might reflect rapid evolution of variable sequences within the 18S rDNA. The rate of 4% sequence divergence per 100 my seems even more biased to hypervariable sites.

The internal transcribed spacers (ITS-1 and ITS-2) of the nuclear rDNA seem more useful for estimation of more recent events. ITSs evolve more rapidly than the 18S rRNA gene and may therefore provide information about taxa having appeared and evolved in the last 50 my (Mas-Coma 1999). ITSs have also been used for estimating molecular clocks: a base substitution rate for the ITS-2 (usually somewhat more conserved than the ITS-1) of 0.3-0.8% per 1 my has been obtained for different organisms (Despres et al. 1992). However, ITSs may furnish significant data for this purpose only when dealing with taxa showing a relatively low number of nucleotide differences. If nucleotide differences are too numerous, important information may be lost because of the impossibility of detecting an accumulation of nucleotide substitutions in the same sequence position over time.

In Hemiptera, estimates of the divergence of the main lineages based on fossil history and using an 18S rDNA molecular clock, showed differences in substitution rates in 18S rDNAs between Euhemiptera and Sternorrhyncha (Campbell et al. 1995a, b), illustrating the problem of a universal 18S rDNA molecular clock. Percentage nucleotide substitutional differences in 18S rDNAs of exemplar Clypeorrhyncha and Archeorrhyncha shpew a range of 3.7-4.8%. However, the estimated time of divergence of these lineages using the conventional molecular clock rate of 1.8% per 100 my, would be 270-260 mya, which falls within the times of divergence interpreted from fossil evidence that indicates that major lineages of Hemiptera probably diverged between the Lower Permian and Lower Triassic (about 260-220 mya) (Evans 1963, Hennig 1981, Shcherbakov 1992, Sorensen et al. 1995).

The molecular clock approach has been amply discussed and criticized (eg. Sibley & Ahlquist 1984, Wilson et al. 1987). However, it does seem that there is a correlation of evolutionary rates among closely related evolutionary lineages. This correlation is biologically plausible because factors that may be responsible for divergence of rates among lineages (e.g. population size, generation time, fidelity of DNA replication) may themselves be quite similar among closely related lineages (Thorne et al. 1998). Thus, within given taxonomic
groups and defined categories of genetic markers, the concept of a molecular clock can provide a tool for estimating divergence, above all if it is used to calculate relative times rather than absolute times of divergence. The aim of the present paper is to estimate the evolutionary history and times of divergences of different lineages of Triatominae (Hemiptera: Reduviidae) based on molecular clocks inferred from nucleotide sequences of the small subunit rDNA (also known as 18S) and the second internal transcribed spacer (ITS-2) from the nuclear ribosomal DNA of these bugs.

MATERIALS AND METHODS

The complete sequences of the small subunit 18S and the ITS-2 of the nuclear ribosomal DNA of several species of Triatominae were obtained by direct PCR cycle sequencing methods and the dideoxy chain-termination method. The results furnished by primary and secondary structure analyses, and the phylogenetic trees obtained by parsimony and distance methods confirm that both are useful and complementary markers for relationships in Triatominae, covering from high taxon (family, subfamily, tribe) to even low levels (species, subspecies, populations) (Bargues et al. 2000).

Several Triatominae species were selected for DNA extraction and nuclear ribosomal DNA sequencing. Analyses included in the present paper concern sequences which can be found in the GenBank and EMBL databases under the following accession numbers:


RESULTS

Results inferred from the 18S rDNA - A total of 35 base substitutional differences appear in the 1918 nucleotide position alignment when comparing the 18S rDNA sequences among the Triatominae studied. This gives a base substitution rate of 1.8%.

Although base substitution rates in 18S rDNAs are not equal among all hemipteran taxa (Campbell et al. 1995a), a comparison of pairwise mean distances based on average number of substitutions per site shows that the Triatominae have a substitution rate approximately equal (0.96-1.03 times) to that of other prosorrhynchan taxa with respect to a mutual sister group (i.e., the peloridiomorphan Hackeriella veitchi - Campbell & Ouvrard 1997 in GenBank, Accession No. AF004766). This correspondence also indicates that the base-substitution rate of euhemipteran 18S rDNA does not depart significantly from the conventional rate ascertained by Ochman and Wilson (1987), as already found by Campbell et al. (1995a). In contrast to Sternorrhyncha (see Campbell et al. 1994, 1995a), the normal base substitution rate of Triatominae corresponds to factors such as the absence of expanded variable regions which in the secondary structure may give rise to the appearance of additional base substitutions for stability reasons, their relatively long generation time, and their low number of generations per year (see Martin & Palumbi 1993).

Hence, the base substitution rate for prosorrhynchan 18S rDNA seems equivalent to the rate of 1.8 x 10^-2 substitutions per site per year estimated by Ochman and Wilson (1987) for synonymous 16S and 18S rRNAs, and appears closer to the lowest rate among the extremes noted for 18S rDNA by Escalante and Ayala (1995).

Consequently, the 18S rDNA molecular clock rate in Triatominae appears to be of 1.8% per 100 my. On this base, our calculated time of divergence between the ancestors of Triatomini and those of the Rhodniini can be estimated at 48.9-64.4 mya.

The emergence of the predatory reduviids before or during the early Cretaceous period (135-70 mya) is consistent with their current wide diversity in continents then connected as Gondwanaland. The calculated time for the separation of Gondwanaland is about 100-90 mya (Pitman et al. 1993). Africa and South America began to divide at their southern end about 120 mya but did not become completely separated until less than 100 mya (Smith et al. 1981). Divergence of the ancestors of the Triatomini and Rhodniini is then put at the time when South America was already separated from Africa.

Calculations based on the above-mentioned 18S rDNA base-substitution rate indicate a divergence
of the ancestral T. phyllosoma complex from T. infestans at 22.8-23.5 mya. Similar calculations indicate that divergences between D. maxima and T. infestans, T. dimidiata and the T. phyllosoma complex may be dated at 28.6-31.9, 20.2-20.4 and 17.4 mya, respectively.

Results inferred from the rDNA ITS-2 - In Triatominae, results inferred from ITS-2 sequences have generally been consistent with those obtained using 18S sequences. As in other organisms, the ITS-2 sequence tends to evolve more rapidly, allowing better differentiation between closely related species of Triatominae (Bargues et al. 2000). The nucleotide differences between the ITS-2 sequences of Triatomini and Rhodniini are too numerous (85.8% variable positions in the alignment of the nine Triatomini and three Rhodniini species studied) to furnish significant information. However, the divergence of the T. phyllosoma complex from T. infestans (estimated at 22.8-23.5 mya from the 18S rDNA molecular clock) may be used to estimate the ITS-2 molecular clock rate because only 107 total nucleotide differences (59-60 differences between these taxa in the pairwise distance matrix of PAUP 4.0b 3a - Swofford 1999) appear in the 476 bp-long alignment (22.5% total variable positions) including T. infestans and the phyllosoma complex members (T. dimidiata excluded) obtained with CLUSTAL V (Higgins et al. 1992). Similarly, the divergence of D. maxima and T. infestans (estimated at 25.6-31.9 mya from the 18S rDNA molecular clock) may also be used for the same purpose, with only 103 total differences (58 differences between these taxa in the pairwise distance matrix of PAUP 4.0b 3a) in the 486 bp-long alignment (21.2% total variable positions) including both species obtained with CLUSTAL V.

The comparison of pairwise mean distances based on average number of substitutions per site shows that the species of the genus Triatoma have a substitution rate approximately equal (0.94-1.01 times) to that of other Triatomini taxa with respect to a mutual sister group (i.e., the Rhodniini represented by R. prolixus). Hence, the base substitution rate for the rDNA ITS-2 of Triatomini seems equivalent to the rate of 41.1-99.4% per 100 my or 0.4-1% per 1 my, that is 41.4-99.4 x 10^-10 substitutions per site per year. This means that, in Triatomini, ITS-2 evolves 23-55 times faster than the 18S rDNA. This base substitution rate for the rDNA ITS-2 of Triatomini appears to be similar to that of 0.3-0.8% per 1 my known for other organisms (Despres et al. 1992). Unfortunately, no valid estimations may be inferred in Rhodniini since there are as yet too few 18S and ITS-2 sequences to provide significant results.

On the base of an ITS-2 molecular clock rate in Triatomini of 41.1-99.4% per 100 my, we can estimate the times of divergence between D. maxima and T. infestans, T. dimidiata populations and the T. phyllosoma complex members at 19.5-34.1, 11.8-21.2 and 11.1-19.5 mya, respectively, which agrees well with the estimations obtained for the same divergences by the 18S molecular clock (28.6-31.9, 20.2-20.4 and 17.4 mya, respectively).

The time of divergence between T. infestans and T. dimidiata (Yucatan population excluded) can be estimated at 24.6-38.3 mya, which is similar to the 24.5-37.3 mya and 22.8-23.6 mya estimations for the divergence between T. infestans and the phyllosoma complex obtained by means of the ITS-2 and 18S molecular clock rates, respectively. The divergence between the phyllosoma complex and T. dimidiata (Yucatan population excluded) is estimated at 5.8-12.9 mya. The T. dimidiata population of Yucatan diverged some 5.9-10.5 mya from the other T. dimidiata populations.

No nucleotide differences were found between the T. dimidiata populations from Ecuador and those from Honduras, but the time of divergence between populations of T. dimidiata from southern Mexico (Yucatan population excluded), Honduras and Nicaragua is estimated at 0.72-3.86 mya. These different forms had previously been denoted as subspecies (Usinger 1941, 1944) before being synonymized by Lent and Wygodzinsky (1979). Similarly, the divergence of members of the phyllosoma complex is estimated at only 0.74-2.28 mya, which also seems consistent with the subspecific ranking originally proposed by Usinger (1944).

DISCUSSION

Current theory envisages the Triatominae as a polyphyletic assemblage of blood-sucking Reduviidae, with the various lineages probably derived from different predatory ancestors (Scholfield 1988, Carcavallo et al. 1999, Schofield & Dujardin 1999, Dujardin et al. 2000). The evolutionary transition is assumed to have proceeded from free-living predatory forms to nest-dwelling haematophagous forms, and potential intermediates in such a transition are illustrated by facultative haematophagy in many predatory Reduviidae (e.g. Schofield & Dolling 1993) and facultative predation in many Triatominae (e.g. Miles et al. 1981, Sandoval et al. 2000). There is evidence that adaptation to haematophagy has been a relatively recent process, but may have occurred at different times in the different triatomine lineages (see Schofield, this issue). The results presented here cannot be directly
viewed in terms of the evolution of haematophagy since neither the 18S nor ITS-2 genes are known to be involved in blood-sucking behaviour, but they do lend strong support to arguments for a polyphyletic origin of the Triatominae.

Our results indicate divergence between the two main tribes, Triatomini and Rhodniini, of 48.9-64.4 mya, at about the time when South America was beginning to separate from Antarctica and Australia during the lower tertiary period. In South America at this time, there appear to have been relatively few of the nest-building vertebrates that today form the main silvatic hosts of the Triatominae, and so it seems unlikely that haematophagy had yet developed in these insects (see Schofield, this issue). In contrast, a great array of invertebrate prey species was present, and fossil evidence indicates that most of the currently recognised groups of reduviid predators were already developed in the Americas (see Wootton 1981, Schuh & Slater 1995). The idea that Triatomini and Rhodniini have evolved from different predatory lineages is also supported by morphological, morphometric, and physiological evidence (Schofield & Dujardin 1999).

Within the Triatomini, our results also indicate a polyphyletic origin. The divergence between South American Triatoma represented by T. infestans, and North American species represented by T. dimidiata, various members of the phyllosoma complex and D. maxima, is estimated at around 19.5-38.3 mya — considerably before the joining of South and North America through the isthmus of Panama towards the end of the Pliocene (3-5 mya) (Vermeij 1991). This may be considered strong evidence that North American species of Triatoma derived independently from South American species, an idea suspected from cytogenetical comparisons (Panzera et al. 1998) and increasingly supported by morphometric comparisons (Dujardin et al. 2000). The idea of independent lineages within the genus Triatoma suggests the existence of an early predatory form, separated into two populations in North and South America, both of which subsequently developed haematophagous habits to give the North American and South American Triatoma. The greater divergence of Diplatalogaster may relate to its biogeographic isolation at the southern tip of Baja California.

At the intraspecific level, our results point to the need for careful reconsideration of the taxonomy of the North American species complexes of Triatoma. The chromatic variants of T. dimidiata and those of the phyllosoma complex were originally ranked as subspecies by Usinger (1944), but while the former were then synonymised, the latter were elevated to specific rank by Lent and Wygodzinsky (1979). In contrast, our results reveal greater genetic variability between dimidiata populations than between members of the phyllosoma complex, suggesting that subspecific ranking may be more appropriate for all these forms.

Further studies are required to clarify relationships between the various populations of T. dimidiata. This species is distributed along a broad area from southern Mexico, through all countries of Central America, into Colombia, Ecuador and northern Peru. This geographical distribution is suggestive of an ecological or evolutionary cline, but our results point to two clear anomalies for such an idea. First is the substantial divergence of the Yucatan population from other populations of dimidiata, and second is the lack of divergence of Ecuadorian populations from Honduran and other Mesoamerican populations. It may be that the Yucatan region represents the original source, from which dimidiata has spread northwards into Mexico and southwards into Central America, but the lack of divergence of the Ecuadorian populations suggests that these may have been carried in recent times from Central America — for example by accidental human intervention.

ACKNOWLEDGEMENTS

Sequences here analyzed were from triatomine materials in part supplied by Dr E Magallan-Gastelum, Dr F Breniere, Dr B Espinoza, Dr P Lozano-Kasten, Dr PM Salazar-Schettino (Mexico), MF Abad-Franch (Ecuador) and Dr JCP Dias (Brazil). Dr David Swofford generously provided the beta test version of PAUP 4.0b 3a to MD Bargues.

REFERENCES


Carcavallo RU, Jurberg J, Lent H 1999. Phylogeny of...


Usinger R 1941. Notes and description of Neotropical
Tristominae (Hemiptera: Reduviidae). *Pan-Pac Entomol* 17: 49-57.


