The Molecular Evolution of Trypanosomes

J.R. Stevens and W. Gibson

The absence of a fossil record has meant that the evolution of protozoa has remained largely a matter for speculation. Recent advances in molecular biology and phylogenetic analysis, however, are allowing the 'history written in the genes' to be interpreted. Here, Jamie Stevens and Wendy Gibson review progress in reconstruction of trypanosome phylogeny based on molecular data from rRNA and protein-coding genes.

Since the first broad molecular study of eukaryote evolution[1], phylogenetic analysis (Box 1) of kinetoplastid flagellates has become successively more focused, initially on studies concerning the origins of parasitism in the group[2,3], and subsequently on detailed analyses of evolutionary relationships among *Trypanosoma* and *Leishmania* spp[4–7]. As the level of focus has deepened, the number of species representing each genus in successive studies has also increased and there has been a progression of ideas concerning the evolutionary relationships between the species. This process is well illustrated by the trypanosomes, where the conclusions of initial studies including only one or two species have been altered gradually by inclusion of more taxa in subsequent studies.

Evolving trees

Figure 1 compares four phylogenetic trees based on analysis of 18S rRNA genes. The first tree[8] includes *Trypanosoma brucei*, *T. cruzi* and a third trypanosome species from a fish (Fig. 1a); in common with many other early studies[3,9,10], this tree indicates that the genus *Trypanosoma* is paraphyletic. Increasing the number of species under study to seven still leaves *T. brucei* outside the main trypanosome clade and outside the trypanosomatid clade containing *Leishmania* and *Crithidia*[4] (Fig. 1b). The inclusion of four more trypanosome species showed for the first time that the genus *Trypanosoma* might, in fact, be monophyletic; addition of more outgroup taxa considerably strengthened this result[5] (Fig. 1c). Subsequent trees, including 24 trypanosome species[6] and 47 trypanosome species[11] (Fig. 1d), both unequivocally supported monophyly of trypanosomes, and it seems unlikely that the addition of further taxa will alter this conclusion, at least for the 18S rRNA gene. Indeed, it is now generally recognized that the conclusions of many early (1991–1996) 18S rRNA-based phylogenetic studies of trypanosomes[3,4,8–10] were significantly affected by a combination of insufficient taxa, an imbalance in the spread of included taxa and large intrageneric interclade evolutionary rate differences[5,6,11].

The progressive definition of an 'aquatic clade', comprising trypanosome species isolated from both marine and freshwater fish and amphibia, can also be followed by comparing these trees. Little information can be gleaned from the single isolate included in Fig. 1a, but this clade has clearly emerged in Figs 1b and 1c, with just seven and 11 trypanosome species, respectively, and its possibly ancient divergence from other clades is demonstrated by Fig. 1d. In this tree, two other clades are also clearly defined and well supported (see Box 1 'bootstrap support'). The *T. brucei* clade consists of the Salivarian tsetse-transmitted trypanosomes from Africa; *T. evansi* and *T. equiperdum*, although non-tsetse transmitted and not restricted to Africa, also belong here by virtue of their close morphological and genetic similarity to *T. brucei*. Importantly, this clade is characterized by the phenomenon of antigenic variation[6]. Most of the members of the *T. cruzi* clade originate from South American mammals, with some interesting exceptions: two species of European bat trypanosomes and one as yet unnamed species of kangaroo trypanosome from Australia. The significance of these 'oddments' in the *T. cruzi* clade is considered further below.

Thus, the evolutionary trees have themselves 'evolved' and spawned a progression of ideas about trypanosome evolution. Initial trees, which showed trypanosomes to be paraphyletic[4,8], suggested that parasitism and the digenetic life cycle had arisen more than once in the trypanosome lineage. The unequivocal evidence of monophyly revealed by later trees[5,6,11] clearly contradicts this, but still supports the idea that parasitism and digenetic life cycles evolved independently in several trypanosomatid lineages (Fig. 1d). The hypothesis of coevolution of trypanosomes and their vectors could not be addressed meaningfully until sufficient taxa were included[6,11], revealing some obvious clade and vector associations; for example, trypanosomes in the aquatic clade are probably all transmitted by aquatic leeches and members of the *T. brucei* clade share transmission by tsetse (Fig. 1d).

We anticipate that analysis of additional trypanosome species from birds, reptiles and various mammals will begin to clarify the unresolved evolutionary relationships evident in the lower half of the tree shown in Fig. 1d. The rRNA data have not provided sufficient resolution to determine the exact branching order of these groups, and the tree shows an eight-way polytomy. Interestingly, the aquatic clade forms the first branch from the trypanosome lineage in Fig. 1d, providing some evidence in support of host–parasite coevolution, although the relatively low bootstrap value (62%) indicates that other hypotheses might be considered. We might have reached the limit of the resolving power of the ribosomal marker over this time scale and other markers might be more informative; alternatively, there might have been an explosive divergence of trypanosome species over a very short time period, which will be difficult to resolve with any marker.

Other markers

In addition to the 18S rRNA gene, a variety of other markers, including other rRNAs and protein-coding genes, have been used for evolutionary studies of trypanosomes and other kinetoplastids. Where both 18S

Jamie Stevens is at the School of Biological Sciences, University of Exeter, Exeter, UK EX4 4PS. **Wendy Gibson** is at the School of Biological Sciences, University of Bristol, Bristol, UK BS8 1UG.
Tel: +44 117 928 8249, Fax: +44 117 925 7374,
e-mail: w.gibson@bristol.ac.uk

 0169-4758/99/$ – see front matter PII: S0169-4758(99)01532-X

Box 1. Searching for the 'Correct' Phylogenetic Tree

Alignment of sequences

Sequence alignment and the associated problem of identifying true homology between variable sites and portions of sequences remains one of the most problematic areas in molecular phylogenetic analysis. Alignment can be performed by one or a combination of three main approaches: (1) on the basis of secondary structural and functional domains, such as secondary structure in ribosomal sequences[37]; (2) using one of a range of specialist alignment programs with various weighting options and gap penalties[38]; and (3) by eye, often in relation to previously aligned sequences.

Increasing the number of taxa might be accompanied by problems of hypervariability at some sites and saturation of nucleotide changes at others, resulting in a reduction of informative sites suitable for inclusion in phylogenetic analyses. Sites that are informative between closely related taxa might introduce 'noise' at higher phylogenetic levels, resulting in a loss of definition and reduced bootstrap support (see below); such sites might be excluded from broad analyses, provided a sufficient number remain to be able to perform a meaningful analysis.

Methods of phylogenetic analysis

There are three main methods of phylogenetic analysis in widespread use – distance methods, parsimony and maximum likelihood analysis – the relative merits of which have now been explored directly by a range of simulation studies[39,40]. Although parsimony and maximum likelihood methods require greater computing power than distance methods, this is unlikely to be so limiting in the future.

In distance methods, a pairwise matrix of genetic distances, or similarities, between sequences is calculated first. The resulting matrix of distances is then used to construct a tree by one of the many available least squares clustering methods, such as the unweighted pair-group method using arithmetic averages (UPGMA), neighbour-joining[41] and Fitch-Margoliash methods[42], which attempt to fit the distances to a hypothesized phylogenetic tree. The conversion of taxon/character data into pairwise distance measures inevitably involves a reduction in information content compared with the original sequence data. However, reduction in computational complexity often permits the analysis of larger data sets than is possible with parsimony or maximum likelihood methods and, moreover, certain distance methods (eg. neighbour-joining) have been shown by simulation studies to perform well against such complex methods[43].

Parsimony methods[44] focus on finding the shortest phylogenetic tree(s) to fit the data, that is those that require the smallest number of steps. For molecular data, these steps are individual nucleotide or amino acid substitutions. The advantages of parsimony are: (1) all informative characters are considered, rather than summarized by conversion to a pairwise distance; (2) all possible solutions (most parsimonious trees) can be combined into a consensus tree; and (3) a range of related search strategies allow even very large data sets to be analysed. However, note that parsimony works only on variant characters and thus information from the often large amounts of conserved sequence common to all taxa being analysed is excluded. The basic assumption of parsimony – that evolution proceeds economically – can also be questioned. Moreover, this method is not suitable for all situations[20], particularly where rapid and/or dissimilar change occurs within a data set.

Maximum likelihood is arguably the most powerful approach to phylogenetic analysis currently available. Supported by solid statistical principles[45], such methods calculate the probability of a given tree yielding the observed data. Every position of each sequence is considered in each calculation, according to well-defined parameter values; thus, at present, constraints on computation time quickly render analysis impractical for more than about 30 sequences. The possibility of assessing bootstrap support (see below) for a given data set is similarly limited by computational requirements.

Outgroups

The definition of an outgroup and the associated placement of the tree's root sets the ingroup in evolutionary context. The outgroup might comprise closely related taxa, which can be presumed from prior biological knowledge to form a sister group or be assumed to be ancestral. Thus, free-living bodonid species have generally proved suitable outgroups for rooting trypanosomatid trees using a range of ribosomal and protein-coding genes[3,4,6,8,15,16,26] and, in turn, the phylogenetic position of *Bodo caudatus* has been independently verified by comparison with the even more distantly related species *Euglena gracilis*[8].

Bootstrap support

The 'correctness' of a phylogenetic tree cannot be interpreted without statistical support for the evolutionary relationships presented. Bootstrap analysis[46] involves re-sampling the data to determine the percentage of replicate trees supporting given relationships. Debate surrounding the non-linear nature of bootstrap support is still considerable, although clarification of what such support means and how it can be interpreted continues to be improved[47].

and (partial) 28S rRNA sequences have been used, conclusions relating to *Trypanosoma* largely agree, in that studies using few taxa show paraphyly[3,4,9,10,12,13], whereas a study including 11 *Trypanosoma* spp[5] indicates monophyly. However, studies based on the glyceraldehyde phosphate dehydrogenase (GAPDH) gene have consistently shown *Trypanosoma* to be monophyletic[14–17], with two to five species, indicating that this gene might be a more reliable phylogenetic marker over this time scale. Similarly, studies using 9S and 12S mitochondrial rRNA genes[2], elongation factor 1α[18], trypanothione reductase and α-tubulin[16], and phosphoglycerate kinase[19], using at most five trypanosome species, also indicate the genus to be monophyletic.

Should analysis of the 18S rRNA gene be abandoned in favour of other markers? Certainly the lack of sufficient taxa in many early 18S studies[3,4,13] appears to have allowed the phenomenon of 'long branch attraction'[20,21] to occur, whereby *T. brucei* has been pulled towards various outgroup taxa by a high, but unconnected in evolutionary terms, level of substitutions[22,23]. As shown above, this problem can often be resolved by inclusion of more taxa. A more serious problem highlighted by the most recent 18S rRNA gene studies[5,6,11] is the high rate of substitution in the *T. brucei* clade compared with other clades. The extent to which unequal evolutionary rates between clades might have distorted the topology of the tree is as yet unknown and, clearly, a pressing question for future analyses[22,24].

Additional gene markers will undoubtedly help to unravel the higher level polytomies within *Trypanosoma* apparent in even the most recent 18S rRNA

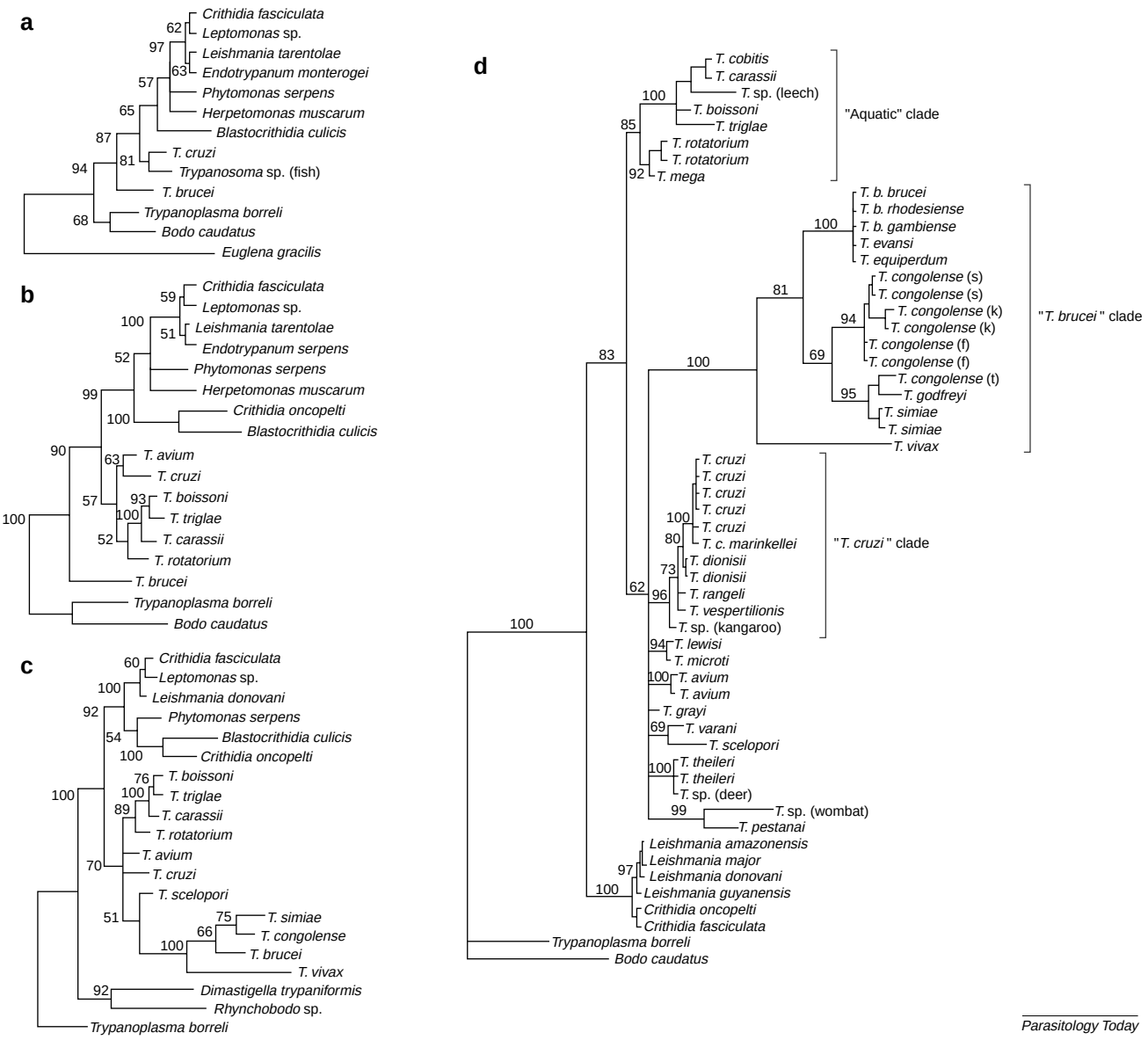

Fig. 1. Comparison of four phylogenetic trees based on bootstrapped maximum parsimony analysis of the 18S rRNA gene, each tree containing progressively more *Trypanosoma* species. In (a) (Ref. 8) and (b) (Ref. 4), the genus *Trypanosoma* is paraphyletic, whereas in (c) (Ref. 5) and (d) (Ref. 11) it is shown to be monophyletic. Note that although the relative branch lengths within each tree are correct, branch lengths between trees cannot be compared directly. *Trypanosoma* sp. (fish) (a) is the same sample as *T. carassii* in (b) and (c) (Ref. 4). In (d), *T. congolense* subgroups are denoted by f (forest), k (kilifi), s (savannah) and t (tsavo). We thank the authors for kindly allowing us to use their trees; (c) and (d) are reproduced, with permission, from Refs 5 and 11, respectively.

gene-based phylogenies (eg. Fig. 1d). Despite the inclusion of increasing numbers of species, recent work[25] indicates the sensitivity of such trees (and perhaps parsimony) to different outgroup taxa and the effect on tree topology. For example, the addition of two *Phytomonas* 18S rRNA sequences to the analysis significantly reduced phylogenetic definition within the upper level of the *Trypanosoma*, such that the aquatic clade no longer diverged earlier than other *Trypanosoma*, resulting in a nine-way polytomy[25]. This result underlines the important influence that the choice of outgroup taxa (Box 1) might exert on phylogenetic analyses and resultant evolutionary conclusions. Interestingly, other studies have also revealed phylogenetic problems associated with inclusion of *Phytomonas*[17,26]; for example, difficulties in resolving the relationship of the *Phytomonas* and *Herpetomonas* lineages by any of a variety of tree reconstruction methods[17].

Thus, although it now seems certain that *Trypanosoma* is monophyletic, which indeed makes biological sense, it seems equally certain that the taxonomic status of the genus will not be resolved fully until the phylogenetic relationships of various closely related sister genera are also resolved fully.

Dating the trees

Phylogenetic trees for trypanosomes are of interest for what they can reveal about the evolution of parasitism and other characteristics, such as antigenic variation or sexual reproduction, in the group. Interpretation of the tree in relation to other events on the evolutionary time-scale depends on conversion of

branch points into dates to estimate time of divergence of different clades. The molecular clock approach[27] assumes that changes in a given sequence accumulate at a constant rate, and thus that the difference between two sequences is a measure of the time of divergence. From a post-genomics standpoint these notions look almost quaint and, indeed, the approach has been amply discussed and criticized over the years[28–30]. Nevertheless, within given taxonomic groups and defined categories of genetic marker, the concept of a molecular clock can provide a useful tool for dating divergence. For example, using an estimate of 0.85% substitutions per 100 million years (my) derived from rRNA analysis of Apicomplexa[31], the divergence of Salivarian trypanosomes from other trypanosomes was dated at about 300 million years before present (mybp)[6].

A second way in which times of divergence can be estimated relies on congruence of host and parasite phylogenies. Thus, parasite trees can be calibrated by reference to known time points within host phylogenies, which have been independently dated from the fossil record. This assumes that existing associations between hosts and parasites reflect past associations. Using this approach, the divergence of fish from higher vertebrates (400 mybp) and the divergence of birds from rodents (220 mybp) were used to estimate the split of Salivarian trypanosomes from other trypanosomes at 260 and 500 mybp, respectively[6].

Putting all these dates on a geological time-scale, the most recent estimate (260 mybp) places the divergence of the Salivaria in the Permian, at a time when reptiles were the most advanced vertebrates. Thus, the Salivaria would have diverged long before even the most primitive ancestors of their present hosts had appeared. Perhaps by considering the trypanosome phylogeny in the context of known biogeographical events, a more realistic estimation of divergence could be obtained. This approach to phylogeny calibration is known as vicariance biogeography[32] and several studies of trypanosomatids have drawn on this technique. For example, using the breakup of Africa and South America (Fig. 2) to date the divergence of *Leishmania* and *Trypanosoma*[2], to corroborate the split between Old and New World *Leishmania*[3] and, most recently, to date the divergence of *T. brucei* and *T. cruzi*[11].

From this study, the divergence of the Salivarian clade is dated around 100 mybp, when Africa became isolated from the other continents (Fig. 2). This is based on the observations that the *T. brucei* clade consists exclusively of African mammalian tsetse-transmitted species and that trypanosome species from African amphibia and reptiles are unrelated (*T. mega*, *T. grayi*, *T. varani*; Fig. 1d). At this time, the ancestors of many extant mammalian groups were present,

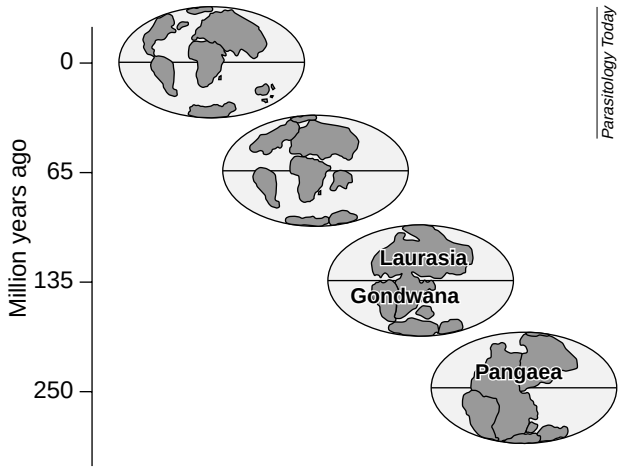

Fig. 2. Sketch maps illustrating the effects of continental drift over geological time[34].

but had not yet begun major diversification and it is easy to envisage subsequent coevolution of this clade with African hosts. Interestingly, Lambrecht[33] arrived at a similar evolutionary scenario considering only palaeoecological data. The composition of the *T. cruzi* clade – mostly mammalian trypanosome species from South America – also agrees with this interpretation. Moreover, the inclusion of an Australian marsupial trypanosome in the clade (*Trypanosoma* sp. kangaroo; Fig. 1d) reinforces the idea that this clade had a New World origin; the split of South America from Antarctica and Australia is thought to have occurred later than the separation from Africa[34]. The only trypanosomes from this clade found in the Old World are those infecting bats, mammals that were able to colonize many regions of the world because of their ability to fly.

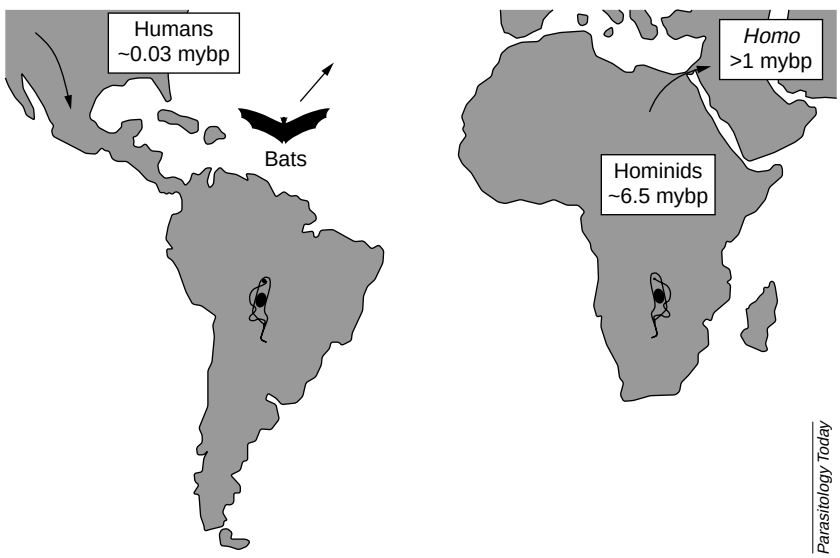

Fig. 3. Hypothesized evolutionary histories of *Trypanosoma cruzi* in South America and *T. brucei* in Africa. Estimated dates for the first appearance of the hominid lineage and migration out of Africa from Ref. 35, prior to the entry of triatomine bugs into the South American disease cycle, *T. cruzi* may have been transmitted as a monogenetic parasite of opossums via anal gland secretions and urine (C.J. Schofield, pers. commun.).

Thus, the phylogenetic evidence suggests that *T. brucei* and *T. cruzi* had very different evolutionary histories with humans. In Africa, *T. brucei* could have shared the long period of primate evolution (~15 my), with eventual emergence of the genus *Homo* ~3 my ago[35], presumably in continuous contact with tsetse (Fig. 3). In contrast, human contact with *T. cruzi* would not have occurred before human migration into the Americas, which is generally dated no earlier than 30 000–40 000 years ago (Fig. 3). Humans might then have become infected as a simple addition to the already extensive host range of *T. cruzi*, which includes other primates[36]. Taking the example of malaria, where several mechanisms of genetic resistance have been selected in the susceptible human population, a prolonged period of struggle between trypanosome and host should also have led to selection for increased host defences. It is tempting to speculate that the long evolutionary history of humans with Salivarian trypanosomes explains our present innate resistance to infection with all but two subspecies of *T. brucei*.

Conclusion

Sufficient data have now been gathered to demonstrate unequivocally that trypanosomes are monophyletic. Phylogenetic trees can serve as a framework to reinterpret the biology, taxonomy and present day distribution of trypanosome species, and provide insights into the coevolution of trypanosomes with their vertebrate hosts and vectors. Different methods of dating the divergence of trypanosome lineages can give rise to radically different evolutionary scenarios. The use of one such biogeographically based approach has provided new insights into the coevolution of the pathogens *T. brucei* and *T. cruzi* with their human hosts.

Acknowledgements

The financial support of The Wellcome Trust is gratefully acknowledged. We are indebted to the many colleagues who have contributed to this article by providing materials or ideas.

References

1 Sogin, M.L., Elwood, H.J. and Gunderson, J.H. (1986) **Evolutionary diversity of eukaryotic small subunit rRNA genes.** *Proc. Natl. Acad. Sci. U. S. A.* 83, 1383–1387
2 Lake, J.A. *et al.* (1988) **Evolution of parasitism: kinetoplastid protozoan history reconstructed from mitochondrial rRNA gene sequences.** *Proc. Natl. Acad. Sci. U. S. A.* 85, 4779–4783
3 Fernandes, A.P., Nelson, K. and Beverley, S.M. (1993) **Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism.** *Proc. Natl. Acad. Sci. U. S. A.* 90, 11608–11612
4 Maslov, D.A. *et al.* (1996) **Phylogeny of trypanosomes as inferred from the small and large subunit rRNAs: implications for the evolution of parasitism in the trypanosomatid protozoa.** *Mol. Biochem. Parasitol.* 75, 197–205
5 Lukes, J. *et al.* (1997) **Analysis of ribosomal RNA genes suggests that trypanosomes are monophyletic.** *J. Mol. Evol.* 44, 521–527
6 Haag, J., O'Huigin, C. and Overath, P. (1998) **The molecular phylogeny of trypanosomes: evidence for an early divergence of the Salivaria.** *Mol. Biochem. Parasitol.* 91, 37–49
7 Croan, D.G., Morrison, D.A. and Ellis, J.T. (1997) **Evolution of the genus *Leishmania* by comparison of DNA and RNA polymerase gene sequences.** *Mol. Biochem. Parasitol.* 89, 149–159
8 Maslov, D.A. and Simpson, L. (1995) **Evolution of parasitism in kinetoplastid protozoa.** *Parasitol. Today* 11, 30–32
9 Gomez, E. *et al.* (1991) **What is a genus in the Trypanosomatidae family? Phylogenetic analysis of two small rRNA sequences.** *Mol. Biol. Evol.* 8, 254–259
10 Landweber, L.F. and Gilbert, W. (1994) **Phylogenetic analysis of RNA editing: a primitive genetic phenomenon.** *Proc. Natl. Acad. Sci. U. S. A.* 91, 918–921
11 Stevens, J.R. *et al.* (1999) **The ancient and divergent origins of the human pathogenic trypanosomes, *Trypanosoma brucei* and *T. cruzi*.** *Parasitology* 118, 107–116
12 Briones, M.R. *et al.* (1992) ***Leishmania tarentolae* taxonomic relatedness inferred from phylogenetic analysis of the small subunit ribosomal RNA gene.** *Mol. Biochem. Parasitol.* 53, 121–127
13 Maslov, D.A. *et al.* (1994) **Evolution of RNA editing in kinetoplastid protozoa.** *Nature* 368, 345–348
14 Hannaert, V. *et al.* (1992) **Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate dehydrogenase in *Leishmania mexicana*.** *Mol. Biochem. Parasitol.* 55, 115–126
15 Wiemer, E.A.C. *et al.* (1995) **Molecular analysis of glyceraldehyde-3-phosphate dehydrogenase in *Trypanoplasma borreli*: an evolutionary scenario of subcellular compartmentation in Kinetoplastida.** *J. Mol. Evol.* 40, 443–454
16 Alvarez, F., Cortinas, M.N. and Musto, H. (1996) **The analysis of protein coding genes suggests monophyly of *Trypanosoma*.** *Mol. Phylog. Evol.* 5, 333–343
17 Hannaert, V., Opperdoes, F.R. and Michels, P.A.M. (1998) **Comparison and evolutionary analysis of the glycosomal glyceraldehyde-3-phosphate dehydrogenase from different Kinetoplastida.** *J. Mol. Evol.* 47, 728–738
18 Hashimoto, T. *et al.* (1995) **Phylogenetic place of kinetoplastid protozoa inferred from a protein phylogeny of elongation factor 1α.** *Mol. Biochem. Parasitol.* 70, 181–185
19 Adje, C.A., Opperdoes, F.R. and Michels, P.A.M. (1998) **Molecular analysis of phosphoglycerate kinase in *Trypanoplasma borreli*.** *Gene* 217, 91–99
20 Felsenstein, J. (1978) **Cases in which parsimony or compatibility methods will be positively misleading.** *Syst. Zool.* 27, 401–410
21 Hendy, M.D. and Penny, D. (1989) **A framework for the quantitative study of evolutionary trees.** *Syst. Zool.* 38, 297–309
22 Noyes, H. (1998) **Can *Trypanosoma* trees be trusted?** *Parasitol. Today* 14, 49–50
23 Maslov, D.A. and Lukes, J. (1998) **Searching for a tree that can be trusted.** *Parasitol. Today* 14, 334
24 Noyes, H. and Rambaut, A. (1998) **Reply.** *Parasitol. Today* 14, 335
25 Stevens, J., Noyes, H. and Gibson, W. (1998) **The evolution of trypanosomes infecting humans and primates.** *Mem. Inst. Oswaldo Cruz* 93, 669–676
26 Marche, S. *et al.* (1995) **Characterization and detection of plant trypanosomatids by sequence analysis of the small subunit ribosomal RNA gene.** *Mol. Biochem. Parasitol.* 71, 15–26
27 Zuckerkandl, E. and Pauling, L. (1965) in *Evolving Genes and Proteins*, pp 97–166, Academic Press
28 Sibley, C.G. and Ahlquist, J.E. (1984) **The phylogeny of the hominoid primates as indicated by DNA–DNA hybridization.** *J. Mol. Evol.* 20, 22–25
29 Wilson, A.C., Ochman, H. and Prager, E.M. (1987) **Molecular time scale for evolution.** *Trends Genet.* 3, 241–247
30 Fitch, W.M. (1976) in *Molecular Evolution*, pp 160–178, Sinauer
31 Escalante, A.A. and Ayala, F.J. (1995) **Evolutionary origin of *Plasmodium* and other Apicomplexa based on rRNA genes.** *Proc. Natl. Acad. Sci. U. S. A.* 92, 5793–5797
32 Nelson, G. and Rosen, D.E. (1981) *Vicariance Biogeography: A Critique*, Columbia University Press
33 Lambrecht, F.L. (1980) **Palaeoecology of tsetse flies and sleeping sickness in Africa.** *Proc. Am. Philos. Soc.* 124, 367–385
34 Cox, C.B. and Moore, P.D. (1993) *Biogeography. An Ecological and Evolutionary Approach* (5th edn), Blackwell Science
35 Lewin, R. (1993) *Human Evolution. An Illustrated Introduction* (3rd edn), Blackwell Scientific Publications
36 Hoare, C.A. (1972) *The Trypanosomes of Mammals*, Blackwell Scientific Publications
37 Neefs, J-M. *et al.* (1990) **Compilation of small ribosomal subunit RNA sequences.** *Nucleic Acids Res.* 18, 2237–2243
38 Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice.** *Nucleic Acids Res.* 22, 4673–4680
39 Huelsenbeck, J.P. (1995) **The performance of phylogenetic methods in simulation.** *Syst. Biol.* 44, 17–48
40 Wiens, J.J. and Servedio, M.R. (1998) **Phylogenetic analysis and intraspecific variation: performance of parsimony, likelihood and distance methods.** *Syst. Biol.* 47, 228–253

41 Saitou, N. and Nei, M. (1987) **The neighbour-joining method: a new method for reconstructing phylogenetic trees.** *Mol. Biol. Evol.* 4, 406–425
42 Fitch, W. and Margoliash, E. (1967) **Construction of phylogenetic trees.** *Science* 155, 279–284
43 Nei, M. (1991) in *Phylogenetic Analysis of DNA Sequences*, pp 90–128, Oxford University Press
44 Hennig, W. (1966) *Phylogenetic Systematics*, University of Illinois Press
45 Felsenstein, J. (1981) **Evolutionary trees from DNA sequences: a maximum likelihood approach.** *J. Mol. Evol.* 17, 368–376
46 Felsenstein, J. (1985) **Confidence limits on phylogenies: an approach using the bootstrap.** *Evolution* 39, 783–791
47 Hillis, D.M. and Bull, J.J. (1993) **An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis.** *Syst. Biol.* 42, 182–192

Wolbachia Bacteria of Filarial Nematodes

M.J. Taylor and A. Hoerauf

The finding that the intracellular bacteria of filarial nematodes are related to the Wolbachia *symbionts of arthropods has generated great interest. Here, Mark Taylor and Achim Hoerauf review recent studies by several groups on the structure, distribution and phylogeny of these endosymbionts, and discuss the potential role for these bacteria in filarial disease and as a target for chemotherapy.*

Intracellular bacteria of filarial nematodes were discovered in the 1970s, with the advent of electron microscopy, in studies on a number of species by McLaren, Vincent, Kozek and colleagues[1–4]. With some impressive foresight, these researchers speculated that these bacteria might be related to the *Wolbachia* symbionts of vector insects[1]; they also suggested that these bacteria might contribute to the pathogenesis of filarial disease and offer a novel target for chemotherapy[3,4]. It is therefore surprising that this fascinating symbiosis was completely ignored by many filarial parasitologists until the application of molecular genetic techniques enabled Sironi *et al.* in 1995 to identify the bacterium in the dog heartworm, *Dirofilaria immitis*, as a close relative of the *Wolbachia* complex[5]. This was particularly intriguing as this group of bacteria had previously been reported only in arthropods and were well known as causative agents of a variety of modifications in host development and reproduction[6]. The studies reviewed here have uncovered further evidence for an important contribution of these bacteria to the biology of filarial parasites, which should ensure that these curious organisms are no longer neglected.

Wolbachia in filarial nematodes

Wolbachia have been detected in the majority of filarial species analysed so far (Table 1), including the major filarial parasites of humans: *Wuchereria bancrofti*, *Onchocerca volvulus* and *Brugia malayi*[5,7–9]. The exceptions, which have consistently been shown by PCR or immunohistology to be free of bacteria, are the rodent filaria *Acanthocheilonema viteae*[7] and the deer parasite *Onchocerca flexuosa*[9,10]. Bacteria can be detected in all developmental stages and can be very abundant in adult worms (Figs 1,2). The organisms are restricted to the hypodermis and reproductive tissues of the female worm, which suggests a vertical mode of transmission through the cytoplasm of the egg[3], in accordance with *Wolbachia* of arthropods. Within a section of tissue the bacteria can be present as an individual bacterium, small groups of bacteria, or large groups that almost entirely fill their cellular environment. The bacteria are typically contained in a host-derived vacuole, and evidence of bacteria undergoing division have been reported in adult female worms[3,8]. Despite the frequent presence of large numbers of bacteria within the lateral cords of adult worms, no obvious pathological effect on nematode tissue has been observed. However, McLaren *et al.* noted that heavily infected embryos appear to have an arrested development[1]. Recently, bacteria have been visualized by immunohistology using antibodies against bacterial catalase[9] and heat shock protein 60 (hsp60) (Refs 9,11; A. Koszarski, Dissertation, Universität Hamburg, 1999) (Figs 2,3). The tissue distribution is similar to that described using electron microscopy; bacteria are not observed in male reproductive tissue.

Phylogenetic analysis by comparison of 16S rDNA (Fig. 4) or *fts*Z sequences shows that all filarial *Wolbachia* are closely related and, in general, form a group separate from the *Wolbachia* of arthropods, which are in turn related to other rickettsial bacteria[5,7,8]. Bandi *et al.*, by analysis of the *fts*Z gene, have shown that *Wolbachia* of filariae segregate into two clusters (C and D), which diverge from the A and B clusters recognized for arthropod *Wolbachia*[6]. Within the C and D filarial *Wolbachia* lineages, bacterial phylogeny is congruent with the nematode host phylogeny[7]. Although studies are limited, evidence so far suggests that each species of worm throughout its geographical distribution is infected with the same 'strain' of bacteria; *D. immitis* obtained from Italy, Japan, Spain, Cuba and the USA, isolates of *B. malayi* derived from Malaysia and Indonesia, and *W. bancrofti* from Papua New Guinea and Tanzania (M.J. Taylor and H. Cross, unpublished) contain identical or virtually identical 16S rDNA and *fts*Z sequences within the same nematode species[5,7,8]. In preliminary crossing experiments, the single base pair difference between 16S rDNA of *B. malayi* and *Brugia pahangi Wolbachia* was used to determine the route of transmission. Crosses of female *B. pahangi* with male *B. malayi* resulted in microfilarial progeny containing the maternally derived 'strain' of bacteria, which supported the view that the principal mode of transmission is via the eggs of female worms[8].

All isolates of filarial parasites studied so far by PCR are infected with bacteria (with the exception of *A. viteae*), yet it is not known how widespread bacteria

Mark Taylor is at the Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK L3 5QA. Achim Hoerauf is at the Bernhard-Nocht-Institut fur Tropenmedizin, Hamburg, Germany.
Tel: +44 151 708 9393, Fax: +44 151 708 9007,
e-mail: mark.taylor@liverpool.ac.uk

 0169-4758/99/$ – see front matter PII: S0169-4758(99)01533-1