International Journal of Systematic and Evolutionary Microbiology (2001), **51**, 731–735 Printed in Great Britain

NOTE

Construction and bootstrap analysis of DNA fingerprinting-based phylogenetic trees with the freeware program FreeTree: application to trichomonad parasites

Vladimír Hampl,[1] Adam Pavlíček[2] and Jaroslav Flegr[1]

[1] Department of Parasitology, Faculty of Sciences, Charles University, Viničná 7, 128 44 Prague 2, Czech Republic

[2] Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Prague, Czech Republic

Author for correspondence: Jaroslav Flegr. Tel: +4202 21953289. Fax: +4202 24919704. e-mail: flegr@cesnet.cz

The Win95/98/NT program FreeTree for computation of distance matrices and construction of phylogenetic or phenetic trees on the basis of random amplified polymorphic DNA (RAPD), RFLP and allozyme data is presented. In contrast to other similar software, the program FreeTree (available at http://www.natur.cuni.cz/~flegr/programs/freetree or http://ijs.sgmjournals.org/content/vol51/issue3/) can also assess the robustness of the tree topology by bootstrap, jackknife or operational taxonomic unit-jackknife analysis. Moreover, the program can be also used for the analysis of data obtained in several independent experiments performed with non-identical subsets of taxa. The function of the program was demonstrated by an analysis of RAPD data from 42 strains of 10 species of trichomonads. On the phylogenetic tree constructed using FreeTree, the high bootstrap values and short terminal branches for the *Tritrichomonas foetus/suis* 14-strain branch suggested relatively recent and probably clonal radiation of this species. At the same time, the relatively lower bootstrap values and long terminal branches for the *Trichomonas vaginalis* 20-strain branch suggested more ancient radiation of this species and the possible existence of genetic recombination (sexual reproduction) in this human pathogen. The low bootstrap values and the star-like topology of the whole Trichomonadidae tree confirm that the RAPD method is not suitable for phylogenetic analysis of protozoa at the level of higher taxa. It is proposed that the repeated bootstrap analysis should be an obligatory part of any RAPD study. It makes it possible to assess the reliability of the tree obtained and to adjust the amount of collected data (the number of random primers) to the amount of phylogenetic signals in the RAPD data of the taxon analysed. The FreeTree program makes such analysis possible.

Keywords: fingerprinting, FreeTree software, *Tritrichomonas*, *Trichomonas*, *Tetratrichomonas*

The advent of molecular taxonomic techniques offered a solution for many problems which were out of reach of classical taxonomic methods and approaches. Currently, the methods of construction of phylogenetic trees (based on molecular data) are widely used not only in systematic and comparative biology, but also in ecology, ethology, sociobiology and epidemiology. The methods of molecular taxonomy can be divided into two groups: single-locus methods and multilocus methods. The result of any single-locus method, such as DNA sequencing, microsatellite analysis or single-stranded conformational polymorphism (SSCP) analysis, is the so-called 'gene tree', the topology of

This paper was represented at the XIIIth meeting of the International Society for Evolutionary Protistology in České Budějovice, Czech Republic, 31 July–4 August 2000.

FreeTree is available from IJSEM Online (http://ijs.sgmjournals.org).

Abbreviations: RAPD, random amplified polymorphic DNA; SSCP, single-stranded conformational polymorphism.

Table 1. List of trichomonads strains

Species/strain	Origin	Isolated by:*
Trichomonas vaginalis		
TV 10-02	Prague, Czech Republic	Kulda (1973)[1]
TV 73-87	Prague, Czech Republic	Kulda (1973)[2]
TV 71-96	Prague, Czech Republic	Kulda (1973)[2]
TV 79-49	Prague, Czech Republic	Kulda (1973)[2]
TV 7-37	Prague, Czech Republic	Kulda (1973)[1]
TV 85-08	Prague, Czech Republic	Kulda (1973)[2]
TV 14-85	Prague, Czech Republic	Kulda (1973)[1]
TV 67-77	Prague, Czech Republic	Kulda (1973)[1]
TV 17-48	Prague, Czech Republic	Kulda (1973)[1]
HL-4MT	Liberec, Czech Republic	Těmín (1986)[1]
FF28	Bratislava, Slovakia	Demeš (1987)[2]
C:1-NIH, ATCC 3001	Washington, DC, USA	Jacobs (1956)[3]
JH31A, ATCC 30236	Baltimore, MD, USA	Hollander (1963)[1]
CP1	Peking, China	Tachezy (1987)[2]
JTCRYO	Rio de Janeiro, Brazil	Silva Filho (1982)[4]
CDC-85, ATCC 50143	Columbus, OH, USA	Lossick (1980)[5]
RU357, ATCC 50139	Pennsylvania, PA, USA	Sondheimer (1982)[5]
TALL-MT	Tallin, Estonia	Tompel (1987)[2]
BO	Gothenburg, Sweden	Forsgren (1978)[6]
IR78, ATCC 50138	Vienna, Austria	Meingassner (1978)[7]
Trichomonas gallinae		
TGK	Prague, Czech Republic	Tachezy (1994)[2]
Tritrichomonas foetus		
LUB	Lublin, Poland	Stepkowski (1965)[8]
KVc-1†	Žalmanov, Czech Republic	Lípová (1962)[9]
LIL-1	Lublin, Poland	Stepkowski (1970)
CO-1	Colorado, USA	Kulda (1967)[9]
CB-1	Ithaca, NY, USA	Kulda (1967)[9]
UTO	Utah, USA	McLoughin (1967)[9]
130	San Cristobal, Cuba	Kulda (1966)
B93	Bayamo, Cuba	Kulda (1966)
Tritrichomonas suis		
SU-H3B	Halle, Germany	Kulda (1988)
PC-9	Prague, Czech Republic	Kulda (1964)
RND	Doksany, Czech Republic	Kadlec (1974)
C19F, ATCC 30169	Logan, UT, USA	Hibler (1959)
11S, ATCC 30168	Ames, IA, USA	Buttrey (1956)
1N, ATCC 30167	Ames, IA, USA	Buttrey (1956)
Tritrichomonas nonconforma		
R114	San Vincente, Cuba	Kulda (1965)[2]
Tritrichomonas augusta		
T37, ATCC 30077	Stillwater, OK, USA	Twohy (1959)
Tritrichomonas mobilensis		
M776, ATCC 50116	Bolivia, S. America	Pindak (1985)[10]
Pentatrichomonas hominis		
HOM V-3	Da-Nang, Vietnam	Tolarová (1988)[2]
Trichomitus batrachorum		
BUB	Veselí nad Lužnicí, Czech Republic	Kulda (1983)[2]
Hypotrichomonas acosta		
L3, ATCC 30069	California, USA	Honigberg (1948)

Table 1 (*cont.*)

Species/strain	Origin	Isolated by:*
Tetratrichomonas gallinarum		
M3	Uhlířské Janovice, Czech Republic	Suchánková, Kulda (1970)[11]

* References: [1]Kulda *et al.* (1982); [2]Vaňáčová *et al.* (1997); [3]Reardon *et al.* (1961); [4]Silva Filho *et al.* (1986); [5]Lossick *et al.* (1986); [6]Forsgren & Forssman (1979); [7]Meingassner & Thurner (1979); [8]Kulda *et al.* (1999); [9]Kulda & Honigberg (1969); [10]Culberson *et al.* (1986); [11]Kulda *et al.* (1974).

† KVc-1 is a clone derived from KV-1, ATCC 30924.

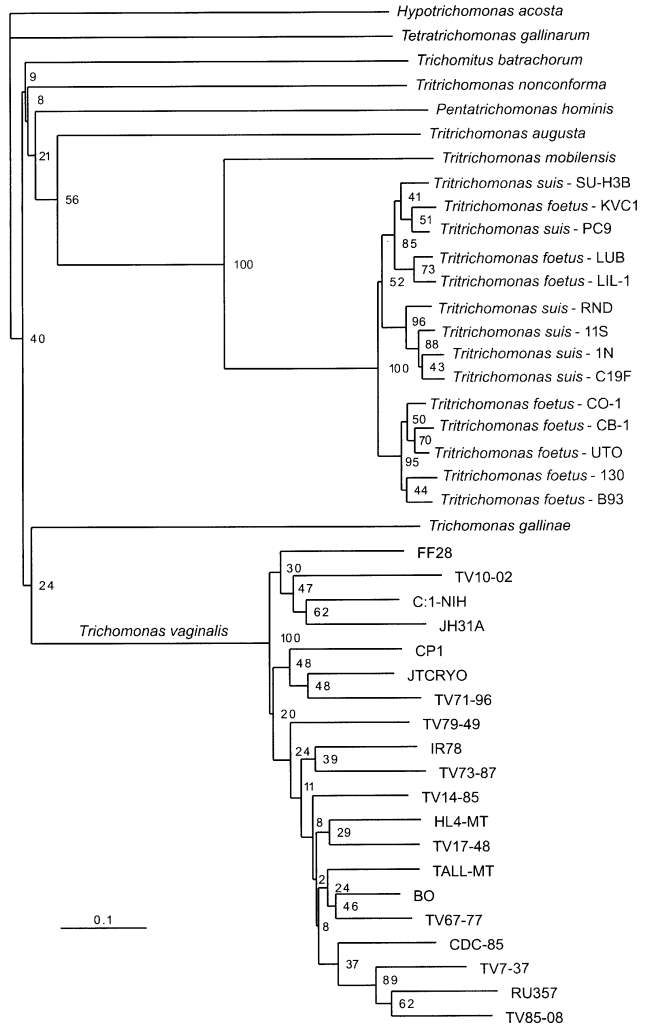

Fig. 1. Phylogenetic tree of trichomonads constructed on the basis of 731 RAPD traits, by the neighbour-joining method.

which reflects the evolution of a particular gene. Very often, however, the aim of the analysis is to obtain a species tree, the phylogenetic tree of the taxa under study, or the genealogical tree of the individuals in the population being studied. Under favourable conditions (i.e. with long intervals between speciation events, in the absence of any horizontal gene transfer between species, etc.), the topology of the gene tree can reflect the topology of the species tree. Mostly, however, such single-locus-based species trees contain some errors (Takahata & Nei, 1985; Neigel & Avise, 1986). A substantial number of errors can be avoided by using multilocus methods such as DNA hybridization, randomly amplified polymorphic DNA (RAPD) analysis, and RFLP. These methods provide species trees based on phylogenetically relevant information contained in many loci or (ideally) in the whole genome. Many of these methods, especially those applicable mainly to an analysis of closely related groups of organisms (RFLP, RAPD), are often even cheaper and easier to perform than the usual single-locus methods. There is, however, one serious technical obstacle in the routine application of multilocus methods. For single-locus methods, a broad collection of application software exists for all stages of the data analysis, including programs for the automatic input of data, the construction of dendrograms and the statistical analysis of the reliability of the results. However, for the multilocus methods, such programs are scarce and for some steps of the analysis are even not available. As far as we know, for the construction of phylogenetic/genealogic trees on the basis of RFLP and RAPD data, no program exists that can also perform the bootstrap or jackknife analysis of the robustness of the tree topology. Whilst all published trees constructed on the basis of DNA sequencing presently contain the information about the robustness of the tree topologies (generally the bootstrap values for internal branches of the trees), this fundamental information is never provided for the trees constructed on the basis of multilocus methods. Unfortunately, it is the information about the robustness of a tree that is often critical for biological interpretation of the data.

Here, we present a description of our program, FreeTree, and demonstrate its performance by analysing 42 strains of 10 species of trichomonads, including six strains of *Tritrichomonas suis*, eight strains of *Tritrichomonas foetus* and 20 strains of *Trichomonas vaginalis* (Table 1). Axenic cultures of all organisms (deposited in the culture collection of the Department of Parasitology, Charles University, Prague) were initiated from frozen stabilates cryo-

protected with 5% DMSO and grown in Diamond's tryptone, yeast extract, maltose (TYM) medium (Diamond, 1982) supplemented with 10% heat-inactivated horse serum. The pH of the medium was adjusted to pH 6·2 for *Trichomonas vaginalis* or to pH 7·2 for all other organisms. Trichomonads from mammals and birds were maintained at 37 °C whereas those from amphibians and reptiles were kept at 26 °C. The last three transfers before harvesting were grown in TYM medium without agar. Cells were harvested in the late exponential phase at an approximate density of 1–3×10^6 cells ml^{-1}. Nucleic acids were isolated using a modified guanidium chloride method (Pramanick *et al*., 1975) and were analysed using 18 primers (OPAs 8, 9, 11, 12, 14, 15, 19 and OPFs 1, 3–6, 9, 10, 12, 14, 16; Operon Technologies), as described elsewhere (Vaňáčová *et al*., 1997). The total number of binary RAPD characters was 731. The program FreeTree was used for the construction of a phylogenetic tree (Fig. 1) and for the bootstrap analysis (Nei–Li distances; neighbour-joining tree-construction method; 250 resampled datasets).

The FreeTree program

The FreeTree program was originally designed for the analysis of the results of DNA fingerprinting methods (RFLP, RAPD, arbitrarily primed PCR) or other methods that provide binary character data (presence/absence of the characters) (Pavlíček *et al*., 1999). For such data, the program computes the distance matrix (by seven different methods), constructs the phenetic or phylogenetic tree (by using the unweighted pair group arithmetic average-linkage algorithm or by using the neighbour-joining method), and computes bootstrap, jackknife or operational taxonomic unit-jackknife values for internal branches of the tree. The program FreeTree can also be used for analysis of the data obtained in two or more independent experiments performed with non-identical subsets of taxa, and also for the construction of trees on the basis of frequency data (e.g. the results of isoenzyme analysis). With the frequency data, however, the program cannot test the reliability of the trees. FreeTree is a Windows 95/98/NT program. The bootstrap analysis of large matrices can take a relatively long time on slow computers. The installation files of FreeTree, a manual in MS Word format, and a sample of the input file are available at http://www.natur.cuni.cz/~flegr/programs/freetree, as an autoextracting archive, or http://ijs.sgmjournals.org/content/vol51/issue3/, as a .zip file.

Analysis of trichomonads

On the phylogenetic tree constructed by our program, all strains of *Trichomonas vaginalis* formed one distinct branch with a bootstrap value of 100%. The strains of *Tritrichomonas foetus* and *Tritrichomonas suis*, however, formed a common branch in which the representatives of both species were intermixed. This branch again had a bootstrap value of 100%. These results are in agreement with the results of other authors (Hammond & Leidl, 1957; Felleisen, 1998), who also suggest that *Tritrichomonas suis*, a commensal that lives in the intestines of pigs, and *Tritrichomonas foetus*, an important pathogen that lives in the urogenital systems of cattle, are in fact the same species. Our results demonstrate that the strains of this species are clearly separated from the sister taxon, i.e. from the morphologically similar species *Tritrichomonas mobilensis*. The bootstrap values of internal branches of the *Tritrichomonas foetus/suis* subtree were very high, not only in comparison with bootstrap values of major branches of the whole trichomonad tree (i.e. the branches of particular species and genera), but also in comparison with internal branches of the *Trichomonas vaginalis* subtree. This suggested that for *Trichomonas vaginalis* the phylogenetic signal-to-noise ratio was too low to allow a reliably resolved tree to be obtained. Such a difference between *Trichomonas vaginalis* and *Tritrichomonas foetus/suis* could be caused either by a faster molecular 'clock' in *Trichomonas vaginalis* or by an earlier time of radiation of its strains. Theoretically, the lower bootstrap values of the *Trichomonas vaginalis* branch could be also caused by the existence of a flow of genetic information between different strains of *Trichomonas vaginalis*, i.e. by the existence of cryptic sexual processes in this species. Only indirect evidence for sexual processes in trichomonads exist (Kulda *et al*., 1986), but some results of karyological studies suggest the possible existence of meiosis in a fraction of *Trichomonas vaginalis* cells in *in vitro* cultures (Drmota & Král, 1997). Even a low frequency of sexual processes, and therefore a low intensity of genetic flow between particular strains, could result in the decrease in bootstrap values of phylograms constructed on the basis of data from a multilocus method.

Our results confirm that the applicability of the RAPD technique in taxonomic or epidemiological research critically depends on the nature of the particular species or taxon under study. In some cases, we can obtain a reasonably robust tree from a moderate amount of RAPD data at the level of genus or higher; in other cases, even the robustness of an intraspecies tree can be rather poor. We suggest that the repeated bootstrap analyses should be an obligatory part of any RAPD study. It makes it possible to assess the reliability of the tree obtained and to adjust the amount of collected data (the number of random primers) to the amount of phylogenetic signals in the RAPD data for the taxon being analysed. The FreeTree program makes such analyses possible.

Acknowledgements

We thank Jaroslav Kulda for providing the biological material, and Tomáš Pavlíček for his help with programming. The work was supported by grants GAUK 107/1998 and MSMT VS96142.

References

Culberson, D. E., Pindak, F. F., Gardner, W. A. & Honigberg, B. M. (1986). *Tritrichomonas mobilensis* n. sp. (Zoomastigophorea: Trichomonadida) from the Bolivian squirrel monkey *Saimiri boliviensis boliviensis*. *J Protozool* **33**, 301–304.

Diamond, L. S. (1982). A new liquid medium for axenic cultivation of *Entamoeba histolytica* and other lumen-dwelling protozoa. *J Parasitol* **68**, 958–959.

Drmota, T. & Král, J. (1997). Karyotype of *Trichomonas vaginalis*. *Eur J Protistol* **33**, 131–135.

Felleisen, R. S. J. (1998). Comparative genetic analysis of tritrichomonadid protozoa by Random Amplified Polymorphic DNA (RAPD) technique. *Parasitol Res* **84**, 153–156.

Forsgren, A. & Forssman, L. (1979). Metronidazole resistant *Trichomonas vaginalis*. *Br J Vener Dis* **55**, 351–353.

Hammond, D. & Leidl, W. (1957). Experimental infections of the genital tract of swine with *Tritrichomonas foetus* and *Trichomonas* species from the cecum or feces of swine. *Am J Vet Res* **18**, 461–465.

Kulda, J. & Honigberg, B. M. (1969). Behavior and pathogenicity of *Tritrichomonas foetus* in chick liver cell cultures. *J Protozool* **16**, 479–495.

Kulda, J., Suchánková, E. & Svoboda, S. (1974). Studies on pathogenicity of *Tetratrichomonas gallinarum* in mice and turkey poults. *Acta Vet Brno* **43**, 53–64.

Kulda, J., Vojtěchovská, M., Tachezy, J., Demeš, P. & Kunzová, E. (1982). Metronidazole resistance of *Trichomonas vaginalis* as a cause of treatment failure in trichomoniasis. *Br J Vener Dis* **58**, 394–399.

Kulda, J., Nohýnková, E. & Ludvík, J. (1986). Basic structure and function of the trichomonad cell. *Acta Univ Carol Biol* **30**, 181–198.

Kulda, J., Poislová, M., Suchan, P. & Tachezy, J. (1999). Iron enhancement of experimental infection of mice by *Tritrichomonas foetus*. *Parasitol Res* **85**, 692–699.

Lossick, J. G., Müller, M. & Gorrell, T. E. (1986). In vitro drug susceptibility and doses of metronidazole required for cure in cases of refractory vaginal trichomoniasis. *J Infect Dis* **153**, 948–955.

Meingassner, J. G. & Thurner, J. (1979). Strain of *Trichomonas vaginalis* resistant to metronidazole and other 5-metronidazole. *Antimicrob Agents Chemother* **15**, 254–257.

Neigel, J. E. & Avise, A. C. (1986). Phylogenetic relationship of mitochondrial DNA under various models of speciation. In *Evolutionary Processes and Theory*, pp. 515–534. Edited by S. Karlin & E. Nevo. New York: Academic Press.

Pavlíček, A., Hrdá, Š. & Flegr, J. (1999). FreeTree – freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness. Application in the RAPD analysis of genus *Frenkelia*. *Folia Biol* (*Prague*) **45**, 97–99.

Pramanick, D. J., Forstová, J. & Pivec, L. (1975). 4 M Guanidine hydrochloride applied to the isolation of RNA from different sources. *FEBS Lett* **62**, 488–490.

Reardon, L. V., Ashburn, L. L. & Jacobs, L. (1961). Differences in strains of *Trichomonas vaginalis* as revealed by intraperitoneal injections into mice. *J Parasitol* **47**, 527–532.

Silva Filho, F. C., Elias, C. A. & Desouza, W. (1986). Further studies on the surface charge of various strains of *Trichomonas vaginalis* and *Tritrichomonas foetus*. *Cell Biophys* **8**, 161–176.

Takahata, N. & Nei, M. (1985). Gene genealogy and variance of interpopulational nucleotide difference. *Genetics* **110**, 325–344.

Vaňáčová, S., Tachezy, J., Kulda, J. & Flegr, J. (1997). Characterization of trichomonad species and strains by PCR fingerprinting. *J Eukaryot Microbiol* **44**, 545–552.