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

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

[1]Department of Parasitology, Faculty of Sciences, Charles University, Prague, Czech Republic
[2]Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Prague, Czech Republic

Corresponding author: Jaroslav Flegr, Department of Parasitology, Faculty of Sciences, Charles University, Viničná 7, 128 44 Prague 2, Czech Republic. Fax (4202) 24919704; Tel. (4202) 2195 3289; e-mail flegr@cesnet.cz

Running title: RAPD analysis of trichomonads using FreeTree

Keywords: Fingerprinting, software, *Tritrichomonas, Trichomonas, Tetratrichomonas.*
Abbreviations: AP-PCR, arbitrarily primed PCR; OTU, operational taxonomic unit; RAPD, random amplified polymorphic DNA; SSCP, single-strand conformational polymorphism; TYM, trypton, yeast extract, maltose medium.

SUMMARY

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), restriction fragment length polymorphism (RFLP) and allozyme data is presented. In contrast to other similar software, the program FreeTree (available at www.natur.cuni.cz/~flegr/programs/freetree) can also assess the robustness of the tree topology by bootstrap, jackknife or OTU-jackknife analysis. Moreover, the program can be also used for an analysis of data obtained in several independent experiments performed with nonidentical 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 by the FreeTree the high bootstrap values and short terminal branches for *Tritrichomonas foetus/suis* 14-strains-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 *Trichomonas vaginalis* 20-strains-branch suggested more ancient radiation of this species and possible existence of genetic recombination (sexual reproduction) in this human pathogen. The low bootstrap values and a star-like topology of the whole Trichomonadidae tree confirm that RAPD method is not suitable for phylogenetic analysis of protozoa on the level of higher taxa. We suggest that the repeated bootstrap analysis should be an obligatory part of any RAPD study. It makes possible to assess the reliability of the obtained tree and to adjust the amount of collected data (number of random primers) to the amount of phylogenetic signals in the RAPD data of the analysed taxon. The FreeTree program makes such analysis possible.

The advent of molecular taxonomy techniques offered a solution for many problems, which were out of reach of classical taxonomy methods and approaches. Currently, the methods of construction of phylogenetic trees on the basis of 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 multi-locus methods. The result of single-locus methods such as DNA sequencing, microsatellite analysis and single-strand conformational polymorphism (SSCP) analysis is a so-called gene tree, the topology of 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 the study or the genealogical tree of the individuals in the studied population. Under favourable conditions (long intervals between speciation events, 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 the multi-locus methods like DNA hybridization, random amplified polymorphic DNA (RAPD) and restriction fragment length polymorphism (RFLP). These methods provide the species tree based on phylogenetically relevant information contained in many loci or (in the ideal case) 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 for routine application of multi-locus methods. For single-locus methods, a broad collection of application software exists for all stages of the data analysis including the programs for automatic input of data, construction of dendrograms and statistical analysis of the reliability of the results. However, for the multi-locus methods, such programs are scarce and for some steps of the analysis are even not available. As far as we know, no program exists for the construction of phylogenetic/genealogic trees on the basis of RFLP and RAPD data that can also perform the bootstrap or jackknife analysis of the robustness of the tree topology. While all trees constructed on the basis of DNA sequencing in the scientific literature presently contain the information about the robustness of the tree topology (mostly the bootstrap values for internal branches of the tree), this fundamental information is never provided for the trees constructed on the basis of multi-locus methods. At the same time the information about robustness of the tree is often critical for biological interpretation of the data.

Here we present a description of our program FreeTree and demonstrate its performance by analysing forty two (42) strains of 10 species of trichomonads including 6 strains of *Tritrichomonas suis*, 8 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 cryoprotected with 5% dimethyl sulfoxide and grown in Diamond's trypton, 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 pH 7.2 for all other organisms. Trichomonads from mammals and birds were maintained at 37° C, those from amphibians and reptiles at 26° C. The last three transfers before harvesting were grown in TYM medium without agar. Cells were harvested in the late logarithmic phase at an approximate density of 1-3 x 10^6 cells ml^{-1}. Nucleic acids were isolated using a modified guanidium chloride method (Pramanick *et al.*, 1975) and analysed using 18 primers (OPA 8, 9, 11,12, 14, 15, 19, OPF 1, 3-6, 9, 10, 12, 14, 16; Operon Technologies, Inc., USA) 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 phylogenetic tree and the bootstrap analysis (Nei-Li distances, neighbor-joining tree-construction method, 250 resampled data sets).

The FreeTree program

The program FreeTree was originally designed for the analysis of results of DNA fingerprinting methods (RFLP, RAPD, arbitrarily primed PCR (AP-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 UPGMA or neighbor-joining method) and computes bootstrap, jackknife or OTU-jackknife values for internal branches of the tree. The program FreeTree can be used also for an analysis of the data obtained in two or more independent experiments performed with nonidentical subsets of taxa and also for the construction of trees on the basis of frequency data (e.g. results of isoenzyme analysis). With the frequency data, however, the program cannot test the reliability of the trees. FreeTree is the Windows 95/98/NT program. The bootstrap analysis of large matrices can take a relatively long time on slow computers. The program is available as an autoextractive archive containing the installation files of FreeTree, manual in MS Word format and a sample of the input file at www.natur.cuni.cz/~flegr/programs/freetree.

Analysis of trichomonads

On the phylogenetic tree constructed by our program all strains of *Trichomonas vaginalis* formed one distinct branch with bootstrap value 100%. On the other hand the strains of *Tritrichomonas foetus* and *Tritrichomonas suis* formed a common branch in which the representatives of both species were intermixed. This branch had again the bootstrap value 100%. These results are in agreement with the results of other authors (Hammond & Leidl, 1957; Felleisen, 1998) who also suggest that *Tritrichomonas suis*, the commensal living in the intestine of pig, and *Tritrichomonas foetus*, the important pathogen living in the urogenital system 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 morphologically similar species *Tritrichomonas mobilensis*. The bootstrap values of internal branches of *Tritrichomonas foetus/suis* subtree were very high in comparison with not only bootstrap values of major branches of the whole trichomonads 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 obtain a reliably resolved tree. Such difference between *Trichomonas vaginalis* and *Tritrichomonas foetus/suis* could be caused either by a higher rate of molecular clock in *Trichomonas vaginalis* or by an earlier time of radiation of its strains. Theoretically, the lower bootstrap values of *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 evidences for sexual processes in trichomonads exist (Kulda *et al.*, 1987), however, 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 of bootstrap values of phylograms constructed on the basis of data from a multi-locus method.

Our results confirm that applicability of RAPD technique in taxonomical or epidemiological research critically depends on the nature of the particular species or taxon under the study. In some cases we can obtain a reasonably robust tree from a moderate amount of RAPD data on the level of genus or higher, in other cases even the robustness of intraspecies tree can be rather poor. We suggest that the repeated bootstrap analysis should be an obligatory part of any RAPD study. It makes possible to assess the reliability of the

obtained tree and to adjust the amount of collected data (number of random primers) to the amount of phylogenetic signals in the RAPD data of the analysed taxon. The FreeTree program makes such analysis possible.

ACKNOWLEDGEMENT

We thank to Jaroslav Kulda for providing the biological material and to Tomáš Pavlíček for his help with programming. The work was supported by grant 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 Protist 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., Vojtěchovský, 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.

Page, R. D. M. (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. *Comp Applicat Biosc* 12, 357-358.

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. Aplication in the RAPD analysis of genus *Frenkelia*. Folia Biologica (Praha) 45, 97-99.

Pramanick, D. J., Forstová, J. & Pivec, L. (1975). 4 M Guanidine hydrochloride applied to the isolation 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.

Species	Strain	Host	Origin	Isolation
Trichomonas vaginalis	TV 10-02	*Homo sapiens*, vagina	Prague, Czech Rep.	Kulda, 1973 [1)]
	TV 73-87	*Homo sapiens*, ditto	Prague, Czech Rep.	Kulda, 1973 [2)]
	TV 71-96	*Homo sapiens*, ditto	Prague, Czech Rep.	Kulda, 1973 [2)]
	TV 79-49	*Homo sapiens*, ditto	Prague, Czech Rep.	Kulda, 1973 [2)]
	TV 7-37	*Homo sapiens*, ditto	Prague, Czech Rep.	Kulda, 1973 [1)]
	TV 85-08	*Homo sapiens*, ditto	Prague, Czech Rep.	Kulda, 1973 [2)]
	TV 14-85	*Homo sapiens*, ditto	Prague, Czech Rep.	Kulda, 1973 [1)]
	TV 67-77	*Homo sapiens*, ditto	Prague, Czech Rep.	Kulda, 1973 [1)]
	TV 17-48	*Homo sapiens*, ditto	Prague, Czech Rep.	Kulda, 1973 [1)]
	HL-4MT	*Homo sapiens*, ditto	Liberec, Czech Rep.	Těmín, 1986 [1)]
	FF28	*Homo sapiens*, ditto	Bratislava, Slovakia	Demeš, 1987 [2)]
	C:1-NIH ATCC 3001	*Homo sapiens*, ditto	Washington, D.C., USA	Jacobs, 1956 [3)]
	JH31A ATCC30236	*Homo sapiens*, ditto	Baltimore, USA	Hollander, 1963 [1)]
	CP1	*Homo sapiens*, ditto	Peking, China	Tachezy, 1987 [2)]
	JTCRYO	*Homo sapiens*, ditto	Rio de Janeiro, Brazil	Silva Filho, 1982 [4)]
	CDC-85 ATCC 50143	*Homo sapiens*, ditto	Columbus, USA	Lossick, 1980 [5)]
	RU357 ATCC 50139	*Homo sapiens*, ditto	Pennsylvania, USA	Sondheimer, 1982 [5)]
	TALL-MT	*Homo sapiens*, ditto	Tallin, Estonia	Tompel, 1987 [2)]
	BO	*Homo sapiens*, ditto	Gothenburg, Sweden	Forsgren, 1978 [6)]
	IR78	*Homo sapiens*, ditto	Vienna, Austria	Meingassner, 1978 [7)]
Trichomonas gallinae	TGK	*Columba livia, f.dom.*, crop	Prague, Czech Rep.	Tachezy, 1994 [2)]
Tritrichomonas foetus	LUB	*Bos taurus*, rectum	Lublin, Poland	Stepkowski, 1965 [8)]
	KVC-1	*Bos taurus*, prepucium	Žalmanov, Czech Rep.	Lípová, 1962 [9)]
	LIL-1	*Bos taurus*, ditto	Lublin, Poland	Stepkowski, 1970
	CO-1	*Bos taurus*, ditto	Colorado, USA	Kulda, 1967 [9)]
	CB-1	*Bos taurus*, ditto	Itaca, USA	Kulda, 1967 [9)]
	UTO	*Bos taurus*, ditto	Utah, USA	Mc Loughin, 1967 [9)]
	130	*Bos taurus*, ditto	San Cristobal, Cuba	Kulda, 1966
	B93	*Bos taurus*, ditto	Bayamo, Cuba	Kulda, 1966
Tritrichomonas suis	SU-H3B	*Sus scrofa*, caecum	Halle, Germany	Kulda, 1988
	PC-9	*Sus scrofa*, caecum	Prague, Czech Rep.	Kulda, 1964
	RND	Sus scrofa, nasal cavity	Doksany, Czech Rep.	Kadlec, 1974
	C19F ATCC 30169	*Sus scrofa*, caecum	Logan, USA	Hibler, 1959
	11S ATCC 30168	Sus scrofa, stomac	Ames, USA	Buttrey, 1956
	1N ATCC 30167	*Sus scrofa*, nasal cavity	Ames, USA	Buttrey, 1956
Tritrichomonas nonconforma	R114	*Anolis bartschii*, cloaca	San Vincente, Cuba	Kulda, 1965 [2)]
Tritrichomonas augusta	T37 ATCC 30077	*Bufo sp.*, cloaca	Still water, USA	Twohy, 1959
Tritrichomonas mobilensis	M 776 ATCC 50116	*Saimiri boliviensis*, intestine	Bolivia, iz. Mobile, USA	Pinďák , 1985 [10)]
Pentatrichomonas hominis	HOM V-3	*Homo sapiens*, feces	Da-Nang, Vietnam	Tolarová, 1988 [2)]
Trichomitus batrachorum	BUB	*Bufo bufo*, cloaca	Veselí nad Lužnicí, Czech Rep.	Kulda, 1983 [2)]
Hypotrichomonas acosta	L3 ATCC 30069	*Drymarchon corais*, cloaca	Southern hemisphere, iz. USA	Honigberg, 1948
Tetratrichomonas gallinarum	M3	*Meleagris gallopavo*, caecum	Uhlířské Janovice, Czech Rep.	Suchánková, Kulda, 1970 [11)]

Table 1: List of trichomonads strains

[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)] Meingasssner & Thurner (1979), [8)] Kulda *et al.* (1999), [9)] Kulda & Honigberg (1969), [10)] Culberson & Pindak (1986), [11)] Kulda *et al.* (1979).

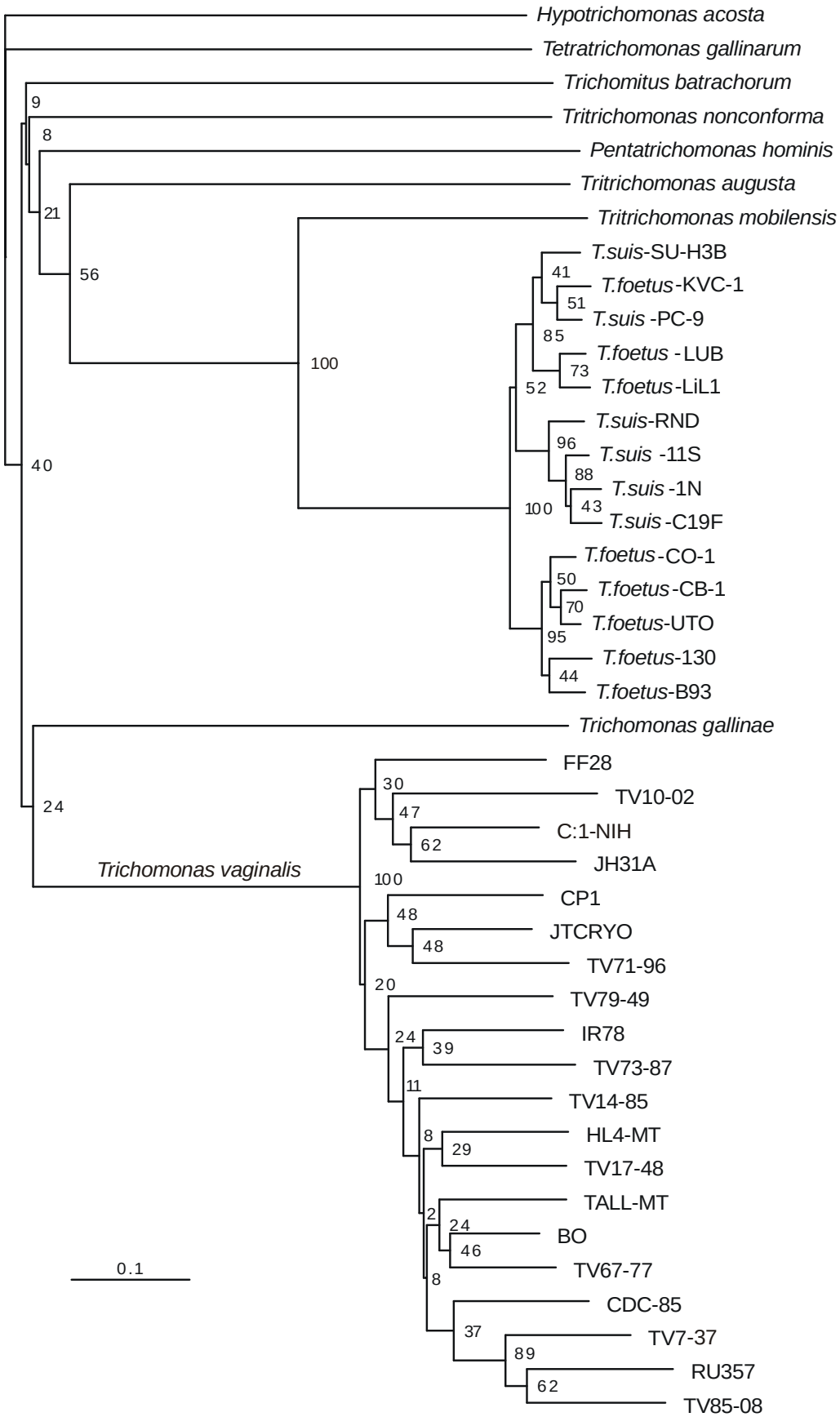

Fig. 1: Phylogenetic tree of trichomonads constructed on the basis of 731 RAPD traits by neighbor-joining method.