UC Irvine
UC Irvine Previously Published Works

Title
Towards a population genetics of microorganisms: The clonal theory of parasitic protozoa

Permalink

Journal

ISSN

Authors
Tibayrenc, M
Ayala, FJ

Publication Date

DOI

Copyright Information

Peer reviewed

─────────────── Viewpoints ───────────────

Towards a Population Genetics of Micro-organisms: The Clonal Theory of Parasitic Protozoa

M. Tibayrenc and F.J. Ayala

Over the past 15 years, molecular investigations, including the study of isozymes and DNA markers, have provided much information on the genetic variation, population structure, breeding system and other population characteristics of parasitic protozoa. For some parasitic protozoa, but not for others, the evidence indicates that their reproduction is prevailingly clonal. In this article, Michel Tibayrenc and Francisco Ayala propose that the issue of whether the predominant mode of reproduction of a given micro-organism is clonal or sexual can only be settled by population genetics information, and they summarize evidence favoring a clonal population structure for a number of parasitic protozoa.

The main feature determining the population structure of an organism is the mating system of the species. What matters, of course, is the prevailing mode of reproduction in natural populations. Thus, successful interbreeding experiments in the laboratory show that the potential for sexual reproduction exists but that it does not occur in nature. Even the occasional observation of sexual reproduction in nature cannot settle the issue of how prevalent sexual reproduction is in natural populations. We submit that the mode of reproduction in parasitic protozoa and other microbial organisms can be ascertained only by population genetic considerations of the distribution of genotypes (especially multilocus genotypes) in populations.

We have previously used population genetic methods to determine the population structure of a number of parasitic protozoan species[1] and have more recently extended these methods to other protozoan species and some yeasts and other fungi[2]. There is evidence of clonal propagation for many species, although the strength of the evidence varies (in part at least) because the strength of the tests depends on the number of genetic markers and the size of the population samples (see Table 1). The evidence for a clonal population structure is often overwhelming (eg. many *Trypanosoma* and *Leishmania*); in others it remains a working hypothesis that can readily be corroborated or falsified by

gathering additional evidence based on larger samples and more genetic markers. Whether or not parasitic protozoa reproduce clonally is of considerable medical importance (see below). Moreover, given recent claims that sexuality and even panmixia are common modes of reproduction for parasitic protozoa[3], the matter deserves attention.

Sexual reproduction

The case of *Plasmodium falciparum* deserves special attention. The hypothesis of uniparental propagation would seem particularly iconoclastic in this case, since an obligatory sexual phase is thought to be required in the life cycle of the parasite. Our analysis of a very limited data set (showing, in one case, strong linkage disequilibrium [see Box 1] in a multilocus allozyme survey of 17 individuals and, in two other studies, the presence in widely separated localities of identical genotypes as evidenced by two-dimensional gel electrophoresis or RFLPs) suggests that gene flow might be biologically restricted, at least in some cases[1,2]. There is, of course, the possibility that these results are aberrant as a consequence of small sample sizes or of accidental laboratory mixups, the more so since some studies favor a sexual population structure for *P. falciparum* (Refs 4, 5 and D. Walliker, this issue). But the matter cannot be settled just by pointing out that the life cycle of *P. falciparum* includes a sexual phase. There are many organisms, including higher metazoa, in which meiosis (and, in some cases, fertilization) occurs but without fusion of gametes or genetic recombination. In automictic parthenogenesis, meiosis and recombination take place but fusion occurs between two haploid nuclei from the same individual. In apomictic parthenogenesis, meiosis in the eggs typically involves only one maturation division, which is equational and yields eventually zygotes that are genetic clones of the parental genotype. Self-fertilization, as it occurs in many annual plants, promotes homozygosis; self-fertilization in homozygous individuals yields genetic clones of the parent. Other mechanisms, such as gynogenesis, also yield progenies that are genetic clones of one parent even though meiosis and fertilization occur. It may very well be the case (and, *a priori*, it seems most likely) that genetic recombination typically, or at least often, occurs in *P. falciparum*. But the matter deserves to be settled

Michel Tibayrenc is at the Laboratory of Genetics for Parasites and Vectors, ORSTOM, BP 5045, 34032 Montpellier Cedex, France and Francisco J. Ayala is at the Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92717, USA.

Table I. Organisms surveyed and rank based on the strength of evidence for clonal reproduction

Organism	Rank[a]	Criteria supporting clonality[b]
Fungi		
Candida albicans	i	none
Candida tropicalis/ paratropicalis	ii	dI,d2,e,f
Cryptococcus neoformans B + C serotypes	ii	e,f
Cryptococcus neoformans A + D serotypes	ii	f
Cryptococcus neoformans all serotypes	ii	dI,d2,e,f
Saccharomyces cerevisiae	ii	f
Protozoa		
Entamoeba histolytica	iii	dI,d2,e,f
Giardia sp.	iii	dI,d2,e,f,g
Leishmania braziliensis guyanensis	iv	dI,d2,e,f
Leishmania infantum	iv	dI,d2,e,f
Leishmania tropica	iv	a,dI,d2,e,f
Leishmania major	iv	dI,d2,e,f
Leishmania Old World as a whole	iv	dI,d2,e,f
Leishmania sp.	iii	g
Naegleria australiensis	ii	a,d
Naegleria fowleri	ii	a
Naegleria gruberi	ii	a
Plasmodium falciparum	ii	dI,d2,e,f,d
Toxoplasma gondii	ii	dI,d2,f
Trichomonas foetus	ii	d
Trichomonas vaginalis	ii	d
Trypanosoma brucei s.l.		
West Africa	iv	dI,d2,e,f
East Africa	iv	dI,d2,e,f
East Africa (wild)	ii	e,f
Liberia	iv	dI,d2,e,f
Busoga, Uganda	iv	dI,d2,e,f
Lambwe Valley, Kenya	iv	dI,d2,e,f
Lambwe Valley (nonhuman stocks)	iv	dI,d2,e,f
Ivory Coast	iv	dI,e,f
Ivory Coast (nonhuman stocks)	iv	dI,f
Trypanosoma congolense	iii	a,dI,d2,e,f
Trypanosoma cruzi	iv	a,b,c,d,f,g
Trypanosoma vivax	iv	dI,d2,e,f

[a] Rank i indicates that the available data do not show clonality; ii, clonality is only a working hypothesis because the supporting evidence comes from small sample numbers; iii, there is evidence for clonality but the limited number of markers prevents equating the strains with actual clones; iv, clonal population structure is well ascertained.

[b] Criteria a, b, and d are qualitative; all others are based on statistical tests significant at the 0.05 level, or at the 0.01 level in the case of criterion f (see Ref. 2). These criteria are identified in Table 2.

and can simply be resolved by obtaining population genetic information, as called for by the criteria listed in Table 2. Such studies are well worth undertaking whatever the outcome, given that so little is known about this parasite's population structure.

Another puzzling case is *Candida albicans*, a yeast for which no sexual stage is known. The one available data set that we analysed[2] suggests a genotype distribution consistent with panmixia. It should, of course, be realized that the failure to reject a null hypothesis does not establish its truth, particularly when the statistical power of the test is low owing to small sample sizes. More extensive samples are required before the prevailing mode of reproduction of *C. albicans* is settled.

The two relevant consequences of sexual reproduction are the segregation of alleles at a locus and their recombination between loci. Evidence that these two processes are rare or absent amounts to evidence that sexual reproduction is rare or absent in nature. Table 2 lists the sources of evidence that should be sought to ascertain whether segregation and recombination are absent, as we have proposed elsewhere[1]. These are the criteria that we used in proposing a clonal population structure for *T. cruzi*[6-9]. The criteria expand the two classical statistics used in population genetics to ascertain random mating and free genetic recombination: Hardy-Weinberg equilibrium frequencies and linkage disequilibrium (see Box 1). They should not be considered as redundant. The various criteria should all be tested whenever possible, although some of them may be either preferable or the only ones possible in certain circumstances. For example, when particular alleles cannot be discerned, or the organism is haploid, or the ploidy is unknown, only the recombinant tests are applicable.

The null hypothesis underlying the tests listed in Table 2 is panmixia. No real organism is perfectly panmictic, which raises the issue of how to distinguish absence of sexual reproduction from the effects of population subdivision, ie. deviations from random mating among populations with different allelic frequencies. The concern is particularly vexing when samples obtained over wide geographic ranges are examined. However, the matter can be discerned because the deviations from the expected random-mating pattern are different for geographic subdivisions and for clonal reproduction[2]. First, deviations from Hardy-Weinberg due to population subdivision entail a deficiency of heterozygotes (Wahlund effect), whereas clonal reproduction is often associated with an excess of heterozygotes (fixed heterozygosity, in the extreme case). Moreover, when segregation or recombination tests yield some genotypes in excess of the expected frequencies, particular genotypes will occur in excess only in particular localities in the case of geographic subdivision, whereas they may be ubiquitous in the case of clonal propagation.

Geographic subdivisions

In principle, it is always possible to overcome the consequences of geographic subdivision on genotypic frequencies by limiting any study to samples that are extremely confined in space and time; in practice this is not always possible, and has the undesirable consequence that biologically meaningful patterns may be missed. An example of such a pattern is shown in Fig. 1, with an over-represented

---

Box 1. Some Genetic Terms Used in This Paper

● **Clonal reproduction:** cellular reproduction involving mitosis, with the consequence that the daughter cells are genetically identical to one another. Modes of reproduction that yield individuals genetically identical to one another have the same population genetic consequences as clonal reproduction (even though meiosis, without genetic recombination, may take place). A species is clonal when the progeny is genetically identical to the reproducing individual.

● **Clonet:** a term coined by the authors to designate, in a clonal species, all the isolates that appear to be genetically identical to one another *on the basis of a particular set of markers.*

● **Fixed heterozygosity:** when all individuals sampled are heterozygous (at one or more loci); this is inconsistent with meiotic segregation and, hence, an indication of clonal propagation.

● **Hardy-Weinberg equilibrium:** the genotype frequencies (given by the square expansion) expected when mating is random. If the frequencies of two alleles, a and b, are p and q ($p+q=1$), the equilibrium frequencies are p^2 (homozygotes aa), q^2 (homozygotes bb) and $2pq$ (heterozygotes ab).

● **Linkage disequilibrium:** nonrandom association between alleles or genotypes at different loci. With linkage equilibrium, the expected frequency of a multilocus genotype is the product of the frequencies of the single-locus genotypes. When there is disequilibrium, the presence of a particular genotype at a polymorphic locus makes it more likely that certain particular genotypes will occur at other polymorphic loci. Thus, in *T. cruzi*, knowing the genotype at the *Gpi* locus makes it possible to predict with high probability the genotype that a given stock will have at any of the other 14 loci (see Ref. 9). Linkage disequilibrium may arise in outbreeding sexually reproducing organisms as a consequence of natural selection and random drift, but it occurs in outcrossing organisms only in low levels if at all. On the contrary, in populations of *T. cruzi*, the observed level of linkage disequilibrium approaches the theoretical maximum, which is strong evidence against interbreeding.

● **Panmixia:** mating between individuals (and, therefore, association between alleles at a locus) occurs at random.

● **RFLP:** restriction fragment length polymorphism, which is exhibited after digestion of DNA with a restriction enzyme.

---

and widely distributed genotype. Such a pattern may provide definite evidence of clonal propagation and also guidance for formulating research strategies concerning, for example, the development of drugs and vaccines.

In addition to geographic subdivision, natural selection is a factor that may result in genotypic distributions different from those expected from panmixia. Natural selection could indeed substantially modify allelic frequencies over the generations and yield genotypic frequencies that are nontrivially different from the panmictic expectations at given loci. However, the multilocus genotypic frequencies frequently observed in our analyses could hardly be accounted for in a parsimonious manner by natural selection since the number of *ad hoc* explanations required to account for missing genotypes increases geometrically with the number of loci (because the number of possible genotypes at n loci is the product of the number of genotypes at each locus[1]). Thus, whereas natural selection may (and probably does) explain some of the peculiarities of the genotypic distributions analysed, it could hardly account for all of them.

The weight of the evidence summarized in Table 1 favors the hypothesis that clonal propagation is a major evolutionary feature of most of the organisms listed, as we have reported[2]. But two aspects of the model we propose need to be kept in mind. First, separate samples that appear to be genetically identical on the evidence of a set of markers should not be thought to be genetically completely identical or actual clones of one another. They should rather be seen as families made up of closely related clones; as more and more genetic markers are examined, it would seem likely that more and more independent

samples will be shown to be genetically somewhat different from one another. We have recently proposed[10] that the term 'clonet' be used, in the case of clonal species, to encompass all the samples in a family that appear to be genetically identical to one another on the basis of a particular set of genetic markers. For example, when the markers used are isozyme loci, the 'zymodemes' (sets of isolates that share a given isozyme profile) can be equated to clonets, provided that clonality has been fully ascertained by population genetics analysis in the species considered. Fig. 1 shows the distribution in South America of a particular *T. cruzi* clonet, identified on the basis of 15 isozyme loci analysed by a population genetic approach[8,9].

The second point to bear in mind is that the clonal model does not imply that genetic recombination is totally absent, but rather that recombi-

Table 2. Population genetic criteria of clonal reproduction

Label[a]	Description
	Segregation (single-locus data)
a	Fixed heterozygosity
b	Absence of segregating genotypes
c	Deviation from Hardy-Weinberg equilibrium
	Recombination (multi-locus data)
d	Over-represented, widespread identical genotypes. Statistical tests: combinatorial (d1), Monte Carlo (d2)
e	Absence of recombinant genotypes
f	Linkage disequilibrium
g	Correlation between independent sets of genetic markers

[a] Criteria a and b are qualitative; all others are based on statistical tests, except d, which can be used qualitatively or statistically (d1 and d2) (see Refs 1,2).

nation is rare even on an evolutionary scale, so that the population genetic consequences of a clonal mode of reproduction persist. Our model is compatible with successful experiments testing for genetic recombination in the laboratory (for review, see Ref. 3). Moreover, the 'dose' of sex that occurs in nature might vary from one species to another and even within a species. As an example, some preliminary results (M. Tibayrenc, unpublished) suggest that *T. brucei* genotype diversity is higher in wild mammals or tsetse flies than in humans or cattle. This could indicate either that recombination is more frequent or simply that clonal diversity is higher in wild mammals or tsetse flies than in domesticated animals and humans.

For underscoring the value of the population genetics approach that we have proposed, as well as for checking its level of resolution, we have investigated a counterexample from human populations[2]. The sample involves 1600 individuals from South America, Europe, Africa and French Polynesia. At a level of resolution that is comparable to our parasitic examples (with respect to sample size, genetic labelling and geographical range), the tests yield results that are quite different from the ones obtained for parasites. In spite of the widely separate human populations sampled, all our tests (Table 2) support a quasi-panmictic model of population structure[2].

Implications of clonality

The clonal model that we propose has significant implications concerning such matters as the development of curative drugs and vaccines, epidemiological surveys and even the diagnosis and treatment of patients. The possibility of identifying, within a given species, particular genetic make-ups that might be studied separately is inversely proportional to the extent of sexual reproduction. Consider a potentially panmictic population model (such as has been proposed for *P. falciparum*[11]); there is little purpose in identifying and studying particular genotypes since individual genetic make-ups are reshuffled every generation. The useful biological entity is the population or the species. Consider now the clonal model; distinctive genetic make-ups ('clonets') exist that persist as wholes through the generations (except for the incorporation of mutations). The distinctive characteristics of a particular clonal lineage (or set of closely related lineages) are of the utmost potential importance in understanding human disease, particularly when different sets of clones are genetically very heterogeneous as a consequence of long separate evolution; such is, for example, the case of *T. cruzi*[6–9].

The population genetic perspective that we are promoting also provides a much-needed foundation for a new taxonomy relevant to clonal parasitic protozoa (and other clonal micro-organisms). Evidence that different isolates are genetically heterogeneous has led, for example in *Leishmania*, to the

Fig. 1. Geographic distribution of clonet 39 of Trypanosoma cruzi. This clonet is heterozygous at three of the 15 characterized isozyme loci (Ref. 9) and it has been recorded from all localities shown, extending through Chile, Bolivia and Southern Brazil. Widespread geographic distribution of a multilocus genotype evinces clonal propagation (see also Ref. 12).

multiplication of species or subspecies names. The process may approach chaos as the number of genetically characterized samples becomes large. It is more parsimonious and informative to identify separate clonal lineages or families of related lineages (or clonets) with a simple code or numerical labelling[10] than to clutter the field with binomial designations.

The clonal model that we propose needs much elaboration and refinement. Issues that need investigation include the following: What level of genetic resolution is called for in practice in order to differentiate clonal lineages that deserve separate investigation or classification? Is the required level of genetic resolution different between species? How can the consequences of residual sexuality in the long-term evolution of the clones be evaluated? Conversely, how can one evaluate the amount of residual sexuality from the population genetic data? The elaboration of the clonal model will largely benefit from comparisons between different sorts of clonal organisms, such as parasitic protozoa with bacteria. The propagation of many bacterial species is known to be prevailingly clonal[12], yet limited genetic recombination has a non-negligible effect on

genetic fine structure[13]. A synthesis of population genetic investigations of procaryotes and unicellular eucaryotes might yield a population genetics of micro-organisms with substantial pay-offs in fundamental biology as well as in applied research.

References

1 Tibayrenc, M. *et al.* (1990) *Proc. Natl Acad. Sci. USA* 87, 2414–2418

2 Tibayrenc, M. *et al.* (1991) *Proc. Natl Acad. Sci. USA* 88, 5129–5133

3 Baker, J.R. (1989) *Int. J. Parasitol.* 19, 465–472

4 Carter, R. and Voller, A. (1975) *Trans. R. Soc. Trop. Med. Hyg.* 69, 371–376

5 Conway, D.J. and McBride, J.S. (1991) *Parasitology* 103, 7–16

6 Tibayrenc, M. *et al.* (1981) *C. R. Acad. Sci. Paris* 293, 207–209

7 Tibayrenc, M. *et al.* (1984) *Trans. R. Soc. Trop. Med. Hyg.* 78, 519–525

8 Tibayrenc, M. *et al.* (1986) *Proc. Natl Acad. Sci. USA* 83, 115–119

9 Tibayrenc, M. and Ayala, F.J. (1988) *Evolution* 42, 277–292

10 Tibayrenc, M., Kjellberg, F. and Ayala, F.J. *BioScience* (in press)

11 Walliker, D. (1985) *Ann. Soc. Belg. Med. Trop.* 65, 60–77

12 Selander, R.K. *et al.* (1987) *Microb. Pathog.* 3, 1–7

13 Maynard Smith, J. *et al.* (1991) *Nature* 349, 29–31

Malaria Parasites: Randomly Interbreeding or 'Clonal' Populations?

D. Walliker

The 'clonality' hypothesis proposed by Michel Tibayrenc and his colleagues[1] has stimulated a long-overdue debate on the genetic structure of populations of protozoan parasites. A critical aspect of the hypothesis is the role of a sexual phase in the life cycle of these organisms. In the malaria parasite, Plasmodium, the existence of a sexual phase is unquestioned and is, indeed, a compulsory part of the cycle in the mosquito host. For this parasite, therefore, the principal question to be addressed, here by David Walliker, is whether populations of this parasite in nature are in a state of random mating (panmixia) or whether they comprise a limited number of clones which only occasionally undergo crossmating.

Intuitively, one can envisage situations in which a state of clonality in *Plasmodium* could occur. It is certainly possible that a single clone could give rise to an outbreak of malaria in a region where the disease occurs infrequently. Mosquito transmission of such a clone would not generate recombinant forms, since the parasite is haploid, and so the same genotype would be found in all infected people. On the other hand, in regions of highly endemic malaria with intense exposure to infection, mixed infections with more than one genotype are likely to be common, allowing frequent crossing between different gametes. In this instance, a state of panmixia would exist. Most work to examine this question has been carried out on parasites in areas of high endemicity and the results support the view that panmixia is the norm in such conditions.

Genetic events in the parasite life cycle

Plasmodium undergoes an entirely haploid cycle in its vertebrate host, the chromosome number

David Walliker is at the Institute of Cell, Animal and Population Biology, Genetics Building, West Mains Road, Edinburgh EH9 3JN, UK.

being 14 (Ref. 2). Cloned blood forms produce both male and female gametes. The zygote (ookinete), produced by fertilization of gametes in the mosquito, is the only diploid stage. Meiosis occurs within a few hours of zygote formation[3], resulting in the eventual formation of haploid sporozoites.

In the life cycle of higher organisms, meiosis is the most important stage for the generation of novel genotypes. If the zygote derives from a mating between unlike genotypes, ie. is heterozygous at many loci, recombination during meiosis will result in numerous progeny with novel combinations of alleles of the genes involved. Such recombination comes about by classical independent assortment of genes on different chromosomes and by crossing-over events between linked genes. However, if the zygote is homozygous at all loci, as would occur following mating between identical gametes, then recombination events at meiosis would not normally be expected to have any detectable genetic consequences. Some of these points are illustrated in Fig. 1.

Crossing experiments in the laboratory

Information on genetic recombination in *Plasmodium* has come from numerous crossing experiments in which deliberate mixtures of gametocytes of two clones are fed to mosquitoes. Assuming that male and female gametes are produced in equal numbers and undergo random mating, then an equal number of self- and cross-fertilization events will occur, and thus 50% of the resulting zygotes can be expected to be hybrids (Fig. 1); recent studies[4] using the polymerase chain reaction to examine single oocysts in mosquitoes infected with such mixtures indicate that this is indeed the case. The resulting sporozoites are then used to infect a vertebrate host and the subsequently developing parasites are examined for the presence of recombinant forms.