UC Irvine
UC Irvine Previously Published Works

Title
Reproductive clonality in protozoan pathogens—truth or artifact? A comment on Ramírez and Llewellyn

Permalink
https://escholarship.org/uc/item/8jn2k9vz

Journal
Molecular Ecology, 24(23)

ISSN
0962-1083

Authors
Tibayrenc, Michel
Ayala, Francisco J

Publication Date
2015-12-01

DOI
10.1111/mec.13443

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 5778–5781

NEWS AND VIEWS

COMMENT

Reproductive clonality in protozoan pathogens—truth or artifact? A comment on Ramírez and Llewellyn

MICHEL TIBAYRENC* and FRANCISCO J. AYALA†

*Maladies Infectieuses et Vecteurs Ecologie, Génétique, Evolution et Contrôle, MIVEGEC (IRD 224-CNRS 5290-UM1-UM2), IRD Center, BP 64501, 34394 Montpellier Cedex 5, France; †Department of Ecology and Evolutionary Biology, University of California, Irvine, 321 Steinhaus Hall, Irvine, CA 92697-2525, USA

The predominant clonal evolution (PCE) model of micropathogens proposed by us has been challenged by a recent paper in *Molecular Ecology*. We review the main tenets of our model and show that the criticisms raised by the paper's authors are based on papers that are either misunderstood or misquoted. We argue that the PCE model and its recent developments (in particular the 'Russian doll model' dealing with micro-clonal evolution) are supported in most cases when adequate data are available.

Keywords: bacteria, breeding systems, evolution of sex, fungus, linkage disequilibrium, molecular evolution, parasite, recombination, virus

Received 9 January 2015; revision received 7 July 2015; accepted 20 July 2015

We have read with great interest the News and Views paper by Ramírez & Llewellyn(2014): 'Reproductive clonality in protozoan pathogens—truth or artifact?', recently published in *Molecular Ecology* (2014) **23**, 4195-4202.

This comment challenges key points of the predominant clonal evolution model (PCE), recently explained by us in a 'spate of recent refinement and re-iteration (Tibayrenc & Ayala 2012, 2013, 2014a,b)' (Ramírez & Llewellyn 2014, p. 4194). Such a spate is far from being a luxury, given that many points of the PCE model still are either misunderstood or mispresented.

Does it make 'little sense' to explore PCE at the level of the entire species?

Predominant clonal evolution is defined as strongly restricted recombination, not absence of recombination. This definition is widely accepted, not only by us, but also by many authors working on parasitic protozoa, fungi, bacteria and viruses (see Tibayrenc & Ayala 2012). This means that evidencing occasional bouts of recombination or hybridization (many cases cited by the authors) does not falsify the PCE model. More specifically, we propose that PCE is verified when the species considered has reached a 'clonality threshold' beyond which clonal evolution overcomes the effects of recombination in the long run. This concept of a clonality threshold is not based on precise quantification of recombination events, but rather, on the empirical observation that the main PCE features (LD, near-clades, deep phylogenies) appear more and more clearly as additional relevant data are considered, according to the 'congruence principle' (Avise 2004). The increase in the phylogenetic signal can be tested by classical linkage disequilibrium (LD) tests and phylogenetic analyses. Moreover, as repeated many times, we have never stated that recombination is 'inconsequential' (Ramírez & Llewellyn 2014). On the contrary, we think that its impact is crucial, but on an evolutionary scale.

One of the most important consequences of PCE in a given species is the existence, due to restricted recombination, of discrete genetic subdivisions that are extremely stable in space and time. We have proposed to call these entities: 'near-clades', to replace the many terms found in the literature (assemblages, clusters, genoclouds, lineages, etc.). A near-clade is a phylogenetic line that is clouded by occasional recombination/hybridization. However, once the species has crossed the 'clonality threshold', the discreteness of the near-clades becomes clearer and clearer when more relevant data are considered. The term 'clade' is not appropriate, as some recombination goes on among the near-clades in most, if not all, micropathogen species surveyed by us. This feature of the model has considerable evolutionary and epidemiological consequences. As a matter of fact, the near-clades, due to their stability and discreteness, constitute appropriate targets for molecular epidemiology (strain typing), and tend to exhibit different phenotypic characteristics, including with respect to epidemiology and clinical features. The authors aim to invalidate this remarkable finding, by stating that 'it makes little sense to address each parasite species (or genus) as a whole', because these 'genetic subdivisions act as reproductive barriers'. The authors omit to mention that evidencing the genetic subdivisions (that is to say: the near-clades) and showing that they act as reproductive barriers has been possible only through the PCE approach, and constitutes one of its main achievements. This way of presenting the affair amounts to making believe that the main conclusion of the PCE approach was self-evident, which is quite untrue: it is the result of long-term research

Correspondence: Michel Tibayrenc, Fax: + 33 4 67 41 62 99; E-mail: Michel.Tibayrenc@ird.fr

© 2015 John Wiley & Sons Ltd

of more than 30 years, started at a time when pathogen population genetics was almost nonexistent. This statement of the authors is inappropriate, because the job still has to be completed or perfected for many pathogen species, and only the PCE approach will allow to do so. Strangely, the authors consider that the PCE approach at the level of the whole species makes little sense. However, in the case of *Toxoplasma gondii*, they accept the hypothesis of PCE based on data that consider the whole species (Sibley *et al.* 2009).

Within-near-clade PCE: the 'Russian doll' (RD) model

This model (Tibayrenc & Ayala 2013) aims at exploring the level of genetic exchange, not among the near-clades, but rather, within each of them. As a matter of fact, many authors consider that there is more recombination within the near-clades than between them. The RD model takes as null hypothesis that within the near-clades, recombination is potentially random (not just that occasional recombination can occur), because the expectations for random recombination are well known (linkage equilibrium). The working hypothesis is that the near-clades each exhibit a miniature picture of the whole species, with the two main features of the PCE model, namely (i) strong LD, or non-random association of genotypes occurring at different loci, and (ii) lesser near-clades. The RD approach simply consists in testing the PCE model within each of the near-clades. Ramírez & Llewellyn attempt to invalidate RD by favouring the null hypothesis. However, the examples they advance are either weak or misleading. Some of them are even counterexamples of what the authors try to demonstrate.

In the case of *Giardia*, the only case cited by the authors to falsify RD (Cooper *et al.* 2007) is not based on a population genetics approach, but rather on the sequencing of a few genes. The conclusions of this article are based on observed discrepancies between the phylogenetic trees based on three loci for seven strains. This is a possible indication that some genetic exchange may go on (although other explanations are possible), but says nothing about the frequency of recombination. Moreover, with such a limited set of data, the risk of statistical type II error (impossibility to reject the null hypothesis of random recombination, not because it is proven, but because the test lacks power) is high. Lastly, this study does not concern a whole *Giardia* 'assemblage' (= near-clade), but rather, the subassemblage A2 of the assemblage A. The mere existence of this subassemblage, which shows that assemblage A is subdivided into discrete units, is in itself evidence supporting RD.

Two of the examples cited about *Leishmania* (Rougeron *et al.* 2009, 2011) do not concern intra-near-clade recombination but rather, the whole species. They have therefore nothing to do with the RD controversy. Moreover, the 'widespread genetic exchange' that the examples are supposed to evidence is actually selfing/inbreeding, which is considered by us and many, if not most, authors working on pathogens population genetics as a particular case of PCE. Lastly, the tests used to infer selfing/inbreeding in *Leishmania* are based on the hypothesis that these parasites are diploid, while convergent results rather favour that *Leishmania* exhibits widespread aneuploidy (Sterkers *et al.* 2014). This bias has been underlined by us (Tibayrenc & Ayala 2012, 2013), as well as by Ramírez *et al.* (2012a). Another study (Rogers *et al.* 2014), repeatedly cited by Ramírez & Llewellyn, is actually an illustrative counterexample of what they try to demonstrate, and a talking illustration for RD. By high-resolution sequencing, Rogers *et al.* (2014) evidence a monophyletic subdivision (near-clade) within *Leishmania donovani/infantum*. They do find indication of hybridization within this near-clade. However, what Ramírez & Llewellyn (2014) omit to mention is that Rogers *et al.* (2014) clearly state that 'extensive LD between chromosomes is consistent with rare recombination and supports clonality as the primary reproductive mode' and that there is 'mainly clonal reproduction in the parasite population'. The data presented do not make it possible to know whether there are lesser near-clades within this monophyletic subdivision. However, they do show that clonality is the dominant mode of reproduction within this near-clade, which supports a RD pattern.

The studies cited concerning *Trypanosoma cruzi* are not strong evidence against RD. The first one (Ocaña-Mayorga *et al.* 2010) is consistent with recombination in a small subpopulation of the TCI near-clade in Ecuador. However, as we have shown (Tibayrenc & Ayala 2013), (i) this 'domestic' population is poorly defined (33% of selvatic strains are included in it, while the 'selvatic' clade comprises 12% of domestic strains); (ii) the risk of type II error is high, due to very limited sampling (only 12 truly 'domestic' strains). Additionally, the 'selvatic' clade shows strong LD and clonality, which favours the RD hypothesis. Observing a lack of congruence between mitochondrial and nuclear phylogenies (Ramírez *et al.* 2012a) indicates at best limited introgression, as it can be observed even between different biological species. de Paula Baptista *et al.* (2014) have postulated 'substantial recombination' in Brazilian strains of the TCII near-clade. However, this study relies on limited sampling and is subject to type II error. There is a strong contradiction between the hypothesis of 'substantial recombination' (with lack of LD) and the evidence for clear structuration of these populations, even in sympatry, as shown by STRUCTURE analysis. A recently published study deals with TCI *T. cruzi* populations isolated from selvatic triatomine bugs in Bolivia (Barnabé *et al.* 2013). The data are compatible with frequent genetic exchange, however, as the authors assert their data are prone to type II error and should be corroborated by more extended sampling (Barnabé *et al.* 2013).

While focusing on these disputable examples consistent with the null hypothesis, Ramírez & Llewellyn omit to cite the many cases presented by us, which strongly support the RD model, including in *T. cruzi*, *Leishmania*, *Toxoplasma*, *Giardia* and *Cryptococcus* (Tibayrenc & Ayala 2013, 2014a,b). This is all the more astonishing, since some of the best examples of RD patterns were found in the authors' own

© 2015 John Wiley & Sons Ltd

data. As a matter of fact, convergent results show that, within *T. cruzi*, the near-clade TCI is subdivided into several smaller near-clades that are widespread and stable in space and time, including in sympatry (Guhl & Ramírez 2011). These 'mini-near-clades' have been corroborated by several different molecular markers, including the mini-exon gene, the cytochrome B gene, multilocus PCR-RFLP and SSUrDNA (Ramírez *et al.* 2011, 2012b,c, 2013). Contrary to what the authors state, within Colombian TCI strains, there is LD. As a matter of fact, the p values for the LD test and the Ia index of association (another LD test) are 4×10^{-4} and 0.037, respectively (Ramírez *et al.* 2013). As noted by Tomasini *et al.* (2014), this is evidence for LD, and not for the opposite, as wrongly concluded by the authors. Similarly, within TCI selvatic strains, strong LD evidences 'widespread clonality, infrequent recombination' (Llewellyn *et al.* 2011). Lastly, strong evidence for a RD pattern within Argentinian TCI strains has been recently published (Tomasini *et al.* 2014).

We do not state that the RD pattern is presently verified in all pathogens. But we do argue that it has been strongly supported by many data, especially in *T. cruzi*, *Leishmania infantum*, *Cryptococcus neoformans*, *Giardia* and *Toxoplasma gondii*, in each case when genetic and population samplings are enough and adequate.

Conclusion

We argue that the PCE hypothesis is not just an artefact, and that it is the most parsimonious model for many, if not most pathogens, including viruses, bacteria, parasitic protozoa and fungi. Our conclusions are based on a large set of published data, dealing with bacteria (48 species), yeasts and fungi (nine species), parasitic protozoa (21 species) and viruses (11 species or categories). These data are easily accessible to any reader wanting to verify our proposals.

If looking for recombination in pathogens amounts to 'searching for a needle in a haystack' (Ramírez & Llewellyn 2014), it obviously means that recombination is severely restricted; in other words, that the PCE model is supported. When it is not, like in some populations of *Plasmodium falciparum*, or in *Helicobacter pylori*, one has no trouble finding clear indications for recombination without searching in a haystack. Until now, the use of more resolutive markers has reinforced the PCE and RD hypotheses. As an example, near-clades and RDs in *Toxoplasma gondii* are better evidenced by a set of highly resolutive markers (Su *et al.* 2012) than by more conventional tools such as multilocus sequence typing. Similarly, whole-genome sequencing in *Neisseria meningitidis* uncovers the existence of deep phylogenies that were not apparent with multilocus sequence typing (Budroni *et al.* 2011). As a matter of fact, contrary to what Ramírez & Llewellyn (2014) state, lack of resolution should favour the null hypothesis (random recombination) rather than PCE, even when highly resolutive markers make it possible to find 'the needle in the haystack'.

Concerning the amount and frequency of recombination, we have called for discarding vague, subjective assertions, such as 'widespread genetic exchange', 'more frequent than expected' (Ramírez & Llewellyn 2014), which amount to a conflict between the supporters of the half full and the half empty bottle. We have proposed, rather, to use a clear-cut PCE criterion relying on the notion of a 'clonality threshold' (see 'Does it make 'little sense' to explore PCE at the level of the entire species?'). If, within a given species, a growing phylogenetic signal is observed as more relevant data are obtained, it means that restricted recombination efficiently counters the effects of genetic exchange, and that the near-clades that subdivide the species will tend to diverge more and more. This PCE-specific signal should be looked for, first at the level of the entire species. Only once the near-clades have been clearly evidenced by this approach, and are fully recognized by most authors (which is still the case for a limited number of species only), RD patterns should be looked for within them, but only in a second round.

References

Avise JC (2004) *Molecular Markers, Natural History and Evolution*, 2nd edn. Chapman & Hall, New York, London.

Barnabé C, Buitrago R, Brémond P *et al.* (2013) Putative Panmixia in restricted populations of *Trypanosoma cruzi* isolated from wild *Triatoma infestans* in Bolivia. *PLoS ONE*, **8**, e82269.

Budroni S, Siena E, Dunning Hotopp JC *et al.* (2011) *Neisseria meningitidis* is structured in clades associated with restriction modification systems that modulate homologous recombination. *Proceedings of the National Academy of Sciences, USA*, **108**, 4494–4499.

Cooper MA, Adam RD, Worobey M, Sterling CR (2007) Population genetics provides evidence for recombination in *Giardia*. *Current Biology*, **17**, 1984–1988.

Guhl F, Ramírez JD (2011) *Trypanosoma cruzi* I diversity: towards the need of genetic subdivision? *Acta Tropica*, **119**, 1–4.

Llewellyn MS, Rivett-Carnac JB, Fitzpatrick S *et al.* (2011) Extraordinary *Trypanosoma cruzi* diversity within single mammalian reservoir hosts implies a mechanism of diversifying selection. *International Journal for Parasitology*, **41**, 609–614.

Ocaña-Mayorga S, Llewellyn MS, Costales JA, Miles MA, Grijalva M (2010) Sex, subdivision, and domestic dispersal of *Trypanosoma cruzi* lineage I in southern Ecuador. *Plos Neglected Tropical Diseases*, **4**, e915.

de Paula Baptista R, Alchaar D'Ávila D, Segatto M *et al.* (2014) Evidence of substantial recombination among *Trypanosoma cruzi* II strains from Minas Gerais. *Infection, Genetics and Evolution*, **22**, 183–191.

Ramírez JD, Llewellyn JD (2014) Reproductive clonality in protozoan pathogens—truth or artefact? *Molecular Ecology*, **23**, 4195–4202.

Ramírez JD, Duque MC, Guhl F (2011) Phylogenetic reconstruction based on Cytochrome b (Cytb) gene sequences reveals distinct genotypes within Colombian *Trypanosoma cruzi* I populations. *Acta Tropica*, **119**, 61–65.

Ramírez JD, Guhl F, Messager LA *et al.* (2012a) Contemporary cryptic sexuality in *Trypanosoma cruzi*. *Molecular Ecology*, **17**, 4216–4226.

Ramírez JD, Duque MC, Montilla M, Cucunubá ZM, Guhl F (2012b) Multilocus PCR-RFLP profiling in *Trypanosoma cruzi* I highlights an intraspecific genetic variation pattern. *Infection, Genetics and Evolution*, **12**, 1743–1750.

© 2015 John Wiley & Sons Ltd

Ramírez JD, Duque MC, Montilla M, Cucunubá ZM, Guhl F (2012c) Natural and emergent *Trypanosoma cruzi* I genotypes revealed by 3 mitochondrial (Cytb) and nuclear (SSU rDNA) genetic markers. *Experimental Parasitology*, **132**, 487–494.

Ramírez JD, Tapia-Calle G, Guhl F (2013) Genetic structure of *Trypanosoma cruzi* in Colombia revealed by a high-throughput nuclear multilocus sequence typing (nMLST) approach. *BMC Genetics*, **14**, 96.

Rogers MB, Downing T, Smith BA *et al.* (2014) Genomic confirmation of hybridisation and recent inbreeding in a vector-isolated *Leishmania* population. *Plos Genetics*, **10**, e1004092.

Rougeron V, De Meeûs T, Hide M *et al.* (2009) Extreme inbreeding in *Leishmania braziliensis*. *Proceedings of the National Academy of Sciences, USA*, **106**, 10224–10229.

Rougeron V, Bañuls AL, Carme B *et al.* (2011) Reproductive strategies and population structure in *Leishmania*: substantial amount of sex in *Leishmania Viannia guyanensis*. *Molecular Ecology*, **20**, 3116–3127.

Sibley LD, Khan A, Ajioka JW, Rosenthal B (2009) Genetic diversity of *Toxoplasma gondii* in animals and humans. *Philosophical Transactions of the Royal Society London Series B, Biological Sciences*, **364**, 2749–2761.

Sterkers Y, Crobu L, Lachaud L, Pagès M, Bastien P (2014) Parasexuality and mosaic aneuploidy in *Leishmania*: alternative genetics. *Trends in Parasitology*, **30**, 429–435.

Su C, Khan A, Zhou P *et al.* (2012) Globally diverse *Toxoplasma gondii* isolates comprise six major clades originating from a small number of distinct ancestral lineages. *Proceedings of the National Academy of Sciences, USA*, **109**, 5844–5849.

Tibayrenc M, Ayala FJ (2012) Reproductive clonality of pathogens: a perspective on pathogenic viruses, bacteria, fungi, and parasitic protozoa. *Proceedings of the National Academy of Sciences, USA*, **109**, E3305–E3313.

Tibayrenc M, Ayala FJ (2013) How clonal are *Trypanosoma* and *Leishmania*? *Trends in Parasitology*, **29**, 264–269.

Tibayrenc M, Ayala FJ (2014a) New insights into Clonality and Panmixia in *Plasmodium* and *Toxoplasma*. *Advances in Parasitology*, **84**, 253–268.

Tibayrenc M, Ayala FJ (2014b) *Cryptosporidium*, *Giardia*, *Cryptococcus*, *Pneumocystis* genetic variability: cryptic biological species or clonal near-clades? *Plos Pathogens*, **10**, e1003908.

Tomasini N, Lauthier JJ, Monje Rumi MM *et al.* (2014) Preponderant clonal evolution of *Trypanosoma cruzi* I from Argentinean Chaco revealed by multilocus sequence typing (MLST). *Infection, Genetics and Evolution*, **27**, 348–354.

M.T. and F.J.A. both elaborated the concepts. M.T. wrote the text. F.J.A. edited it (if Francisco agrees with this).

---

doi: 10.1111/mec.13443

© 2015 John Wiley & Sons Ltd