Microbial Molecular Epidemiology: An Overview

Article *in* Methods in Molecular Biology · February 2009

DOI: 10.1007/978-1-60327-999-4_1 · Source: PubMed

1 author:

Michel Tibayrenc
Institute of Research for Development
490 PUBLICATIONS **14,254** CITATIONS

SEE PROFILE

Chapter 1

Microbial Molecular Epidemiology: An Overview

Michel Tibayrenc

Abstract

In this introductory chapter, I stress one more time the urgency to better connect molecular epidemiology and evolutionary biology. I show how much population genetics and phylogenetic analyses can confer a considerable added value to all attempts to characterize strains and species of pathogens. The problems dealing with the mere definition of basic concepts, such as species, subspecies, or strains, are briefly summarized. Last, I show the important contribution of molecular epidemiology to our knowledge of the basic biology of pathogens and insist on the necessity not to separate the studies dealing with pathogens from those that concern the hosts and the vectors, in the case of vector-borne diseases.

Key words: Cladistics; molecular marker; phylogenetic analysis; population genetics; species; strain typing.

1. Introduction

This introductory chapter is definitely not a comprehensive survey of what molecular epidemiology is today. It instead aims at putting the field into perspective, with its promises fulfilled or let down, its practical implications in terms of public health, its unsolved challenges, and its future potential with the burgeoning of advanced technologies. For more complete overviews of the field, refer to recent reviews *(1*, *2*). The present text is something of a political claim. Other authors of this book may not share the same views.

The field covered by this book is undeniably a topical one: A Medline search with the key words "molecular epidemiology" produces more than 5,000 references. For the sole year 2007,

D.A. Caugant (ed.), *Molecular Epidemiology of Microorganisms*, Methods in Molecular Biology, Vol. 551
DOI: 10.1007/978-1-60327-999-4_1,

the number approaches 1,000. Of these references, roughly 10% cover a different field, which should instead be called genetic susceptibility to diseases. The rest are related to the very theme of this book, which I try to define below. This definition reflects my own views and again may be not shared by the other authors of the book.

1.1. An Attempt to Define Molecular Epidemiology

The Centers for Disease Control and Prevention in Atlanta, Georgia, or more exactly its branch that specializes in transmissible diseases, the National Center for Infectious Diseases, can be considered the mecca of molecular epidemiology. In 1994, this institute issued the following definition of microbial molecular epidemiology: "the various biochemical and molecular techniques used to type and subtype pathogens" ***(3)***. This definition is strictly a technology-based one. As developed in this chapter, I feel it is indispensable to broaden and enrich this definition. First, technology is not enough to characterize pathogens, and its exclusive use could prove to be grossly misleading. The use of evolutionary concepts makes molecular epidemiology considerably more efficient and makes it possible to gather precious knowledge of the basic biology of the organisms under study. Second, identifying pathogens is too narrow a goal for molecular epidemiology. The so-called downstream studies ***(4)*** aim at evaluating the impact of the genetic diversity of pathogens on their relevant medical properties (pathogenicity, antigenic diversity, and drug and antibiotic resistance). These reflections have led me to propose a broad definition of molecular epidemiology ***(1)***: (1) the definition, identification, and tracking of relevant pathogen species, subspecies, strains, clones, and genes by means of molecular technology and evolutionary biology; and (2) the evaluation of the impact of a pathogen's genetic diversity on its relevant medical properties.

1.2. Increasing Importance of the Field and Advanced Technologies

The field of molecular epidemiology has experienced a rapid growth year after year, from fewer than ten references in Medline before 1981 to close to 1,000 for the sole year 2007. It is striking, when doing a retrospective search, to see techniques, such as multilocus enzyme electrophoresis (MLEE), that have been earlier considered gold standards vanish in favor of the new stars: microarrays, real-time polymerase chain reaction (PCR), and especially multilocus sequence typing (MLST). There is undoubtedly something of a fad here. The treasured old techniques did not prove to be unworthy, and they still deserve recognition for certain of their uses (*see* my chapter on MLEE in this book). Moreover, the new stars, although they are very powerful, are by no means panaceas.

It is not my purpose here to denigrate the new technologies. They have undoubtedly contributed considerably to the progress

made in the field. For example, MLST is incomparable in finely dissecting the impact of recombination in microbes (*see* **Subheading 2.5**). I can only repeat here what I have said many times: There are no good and bad techniques; there are only techniques that are better designed to answer given questions. Still the fact remains that, all things being equal, a paper that relies on the hottest technique in fashion will be more easily published than another one based on MLEE or restriction fragment length polymorphism (RFLP).

1.3. What Is Molecular Epidemiology Good for?

In the heroic times of molecular epidemiology (late 1970s), hopes were high that it would become a routine diagnostic tool like enzyme-linked immunosorbent assay and indirect immunofluorescence. This happened only partially. The practical contribution to daily patient care remains limited and mainly consists of species identification using PCR techniques, which is still limited to specialized laboratories. Where strain typing (i.e., characterization at the subspecific level) is concerned, it is not used as a routine analysis.

It can be said that in the present state of the art, molecular epidemiology is more a research tool than a significant contribution to routine clinical medicine *(5)*. In this perspective, many papers apply the current state of knowledge to epidemiological surveys. Many tools are quite standardized and can be successfully applied to various situations. Spoligotyping for the identification of *Mycobacterium tuberculosis* strains is a typical example. Since it was designed more than 10 yr ago *(6)*, hundreds of papers using this technique have been published. Each of them now has a limited added value, restricted to the analysis of specific, local situations. At the other end of this scale, articles developing the most advanced research are paving the way for the molecular epidemiology of tomorrow (*see* **refs**. *7–11*, among others).

1.4. The Distressing and Persistent Gap Between Molecular Epidemiologists and Evolutionists/Population Biologists

I can say that my entire career has been devoted to spreading propaganda in favor of uncompartmentalizing molecular epidemiology and population genetics/evolution. To a large extent, this has proved to be a failure. A recent article again focusing on this need *(1)* amounted to preaching in the desert and is among my least-cited articles. Many evolutionists are attracted to the fascinating models offered by transmissible diseases and coevolution between hosts, pathogens, and vectors. This is the case for the authors of the cited masterpiece papers. However, as a rule, they adhere to a vision of evolutionists, could have very speculative approaches (which is welcome in basic research), and sometimes do not heed the potential applicability of their research in terms of public health. This makes most of these papers simply unreadable for clinicians, public health managers, and even scientists involved in applied research. Notable exceptions can be found in the recent literature. Some evolutionists and phylogeneticians do their best to make themselves accessible to nonspecialists *(9,12)*.

On the other hand, many contributions related to molecular epidemiology and strain typing do not say a word about the possible contribution of evolutionary biology to this discipline. This is true even for very recent papers, some of them published in high-impact journals, supposedly the state-of-the-art in the field *(13–15)*. These articles, although they contain extremely valuable information and may propose innovative concepts, entirely miss an evolutionary interpretation of the data. Bacterial populations are simply considered a set of eternal clones with no recombination among them, which is a glaring mistake for many bacterial species, if not all. Hybrid papers that underline the contribution of evolutionary studies to molecular epidemiology, and remain accessible to nonevolutionists, are the exception rather than the rule *(16)*.

2. The Targets of Molecular Epidemiology: Relevant Species, Subspecies, Strains, Clones, and Genes

The first, basic goal of molecular epidemiology is to identity, characterize, and follow those entities (units of analysis) that are relevant to the clinician and the epidemiologist. This again emphasizes the crucial role of evolutionary biology since these entities are extremely difficult or impossible to characterize and even to define without the help of the concepts from this discipline.

2.1. Species

The concept of species is a typical example of how difficult it is to define and delimit the units of analysis for molecular epidemiology. This has been discussed at length in another article *(17)*, and I only review the many challenges raised by the problem.

Intuitively, pathogen species look like solid entities that should be easy to characterize and follow. However, an entity that is not clearly defined is like a vanishing mirage. A personal anecdote illustrates how misleading it can be to adhere to the unfounded belief that species made official with a Latin name are engraved in stone. Years ago my laboratory was asked to determine the species of a *Leishmania* strain from Latin America (*Leishmania* are the kinetoplastid parasitic protozoa responsible for leishmanioses). Using MLEE and comparison with a set of reference strains, we identified the strain as *Leishmania panamensis*. The colleague who sent the strain responded that the identification was glaringly wrong. He had a counteranalysis done by another laboratory, which identified the strain as another species, *Leishmania guyanensis*. Puzzled by these contradictory results, we performed a broad survey of many strains of both species. The conclusion was crystal clear: If a blind approach was used, by MLEE analysis *L. panamensis* and *L. guyanensis* strains showed

no differences. In other words, these two supposedly separate species had been described first on geographical (phenotypical) grounds, but from a phylogenetic point of view, they could not be distinguished from each other.

For so-called higher organisms, the species concept is already a headache, although species of mammals, birds, and insects do exist and are confirmed by recurrent observations. If pathogen species are concerned, the definition of *species* is a "mission impossible," as confirmed by the abundant literature devoted to it. This led some evolutionists to consider that a definition of the species was hopeless and useless, except in birds ***(18)***. However, scientists working in applied research, clinicians, health professionals, and decision makers cannot accept such an extreme and puristic view: It is an obligation to define the targets of medicine and control measures. Malaria is *not* caused by *Escherichia coli*, and *Leishmania* parasites are *not* transmitted by tsetse flies. Thousands of species are described and used in the world of pathogens. When designing molecular epidemiology tools to try to characterize them, it is crucial to know which upstream concept has been used to define these species. Many microbial species have been defined on epidemiological or medical bases; this is a special case of the phenotypic species concept, according to which species are defined on phenotypic characteristics. For example, *Leishmania infantum* is the causative agent of infant leishmaniosis in the Mediterranean basin, *M. tuberculosis* is the agent of tuberculosis, and so on. When targeting such species with molecular epidemiology, it is necessary to verify that they correspond, at least to some extent, to discrete collections of genetically related genotypes. If this is not obtained, as in *L. panamensis* and *L. guyanensis*, for example, it will be impossible to characterize such species as a whole and to distinguish them from other species. An extreme and classical example is *Shigella* bacteria, which have been assigned the rank of a specific genus by clinicians due to their striking pathogenic properties (they cause severe dysentery). Yet a phylogenetic analysis reveals that *Shigella* are merely a bunch of *E. coli* clones, which are not even monophyletic (they do not constitute a specific, unique evolutionary line). Characterizing all *Shigella* as a discrete genetic entity is therefore hopeless. The only means to specifically track *Shigella* strains is to characterize those pathogenicity genes that make them so virulent.

To handle the species concept for pathogens, the operational view I have defended ***(17)*** states that (i) the world of pathogens in a genetic view is not level and undifferentiated. It has clear discontinuities, even if their borders are not always sharply defined. It is therefore desirable to use the phylogenetic species concept to describe pathogen species ***(19)***, but using a very flexible approach, since the genetic discontinuities that exist in the pathogen world many times do not correspond to sharply defined phylums

(see the discussion of the concept of discrete typing unit). ***(ii)*** One should by all means try *not* to describe new species. In the case of pathogens, it is clear that the definition of a species is really a matter of convenience. One describes species when it is relevant for applied research, clinical practice, and health policy, not when it gives the opportunity to publish a new paper. Let us stop the species inflation.

2.2. Subspecies

Subspecies are subdivisions of a given species that are given a triname. For example, the zebu, long considered a species that was different from the European ox (*Bos indicus* vs. *Bos taurus*), has been made a simple subspecies of the latter (*Bos taurus indicus*), which is logical since absolutely no mating barriers exist between the two formerly described species. In so-called higher organisms, subspecies are defined as geographic morphological variants of a given species. They do exist, as shown by recurrent observations ***(2)***. If pathogens are concerned, nothing clear emerges. One can say that scientists describe subspecies on the same grounds as species (phenotypic or phylogenetic criteria or both) when they dare not describe a species. A pathogen subspecies is something like a timid species—not a very operational concept. It would be wise to drop this practice with pathogens. Either the entity deserves to be defined and is given the rank of species or it does not.

2.3. Strains

The term *strain* is one of the most widely used and the most confusing in the literature dealing with pathogens. In laboratory jargon, a strain is no more than the collection of parasites you handle in Petri dishes or culture flasks. The right term here should be *stock*. Specialists (myself included) often speak about a reference strain, which is a cell line isolated from a given host at a given time in a given place. The correct name should be *isolate*. If molecular epidemiology is concerned, people seek to characterize strains with molecular tools. In this case, one refers to multilocus genotypes, which immediately opens two closely related Pandora's boxes: how to delimitate multilocus genotypes and the problems about defining the notion of clonality.

2.4. Clones, Clonal, Clonality

There is great confusion in the use of the terms *clones*, *clonal*, and *clonality*. When speaking about a clonal species, many authors actually refer to a species whose genetic diversity is either weak or null ***(14)***. Many sexual species have very low genetic variability, while some species with no genetic recombination are genetically extremely diversified. This has nothing to do with the mode of reproduction. Rather, a species whose genetic diversity is very low is simply assumed to have a recent common ancestor, whatever its mating system. Other authors limit the term *clonal* to only mitotic propagation. My articles dealing with clonality have been

frequently misunderstood because of this confusion. A clonal species should instead refer to a species in which descendants are genetically identical to the parent. This gives a genetic definition to the clone. Many cases of uniparental reproduction produce genetic clones, not only mitotic propagation, but also apomictic parthenogenesis, gynogenesis in some fish species, and self-fertilization in haploid organisms. Extreme cases of homogamy will also lead to the production of genetic clones since only those cells that are genetically identical or extremely similar will mate together *(4)*.

Even if clonality is properly defined in this way, in population genetic terms, it does not mean that our trouble is over. First comes the problem of properly characterizing clones. Let us imagine a species that is perfectly clonal, that is, in which gene exchange is totally absent. As we discussed in this **Section 2.4**, this probably does not exist in the world of pathogens. But even if it did, let us characterize clones of this purely clonal species with one of the stars of the fashionable techniques available today, MLST *(11)*. The strains that share identical MLST alleles are referred to as *sequence types*. Can they be considered clones? In other words, are they really genetically homogeneous? The answer is no. The promoters of the MLST approach themselves soon discovered (M. Achtman, personal communication) that RFLP based on a few antigen genes added to MLST considerably improved its resolution power. In other words, the clones identified with MLST are genetically heterogeneous. This is true for any technique. The concept of *clonet* was forged to overcome this problem (*see* **Subheading 3**.).

2.5. Not-So-Clones and Not-So-Clades

The first expression, not-so-clones, is a joke by my witty friend B. Levin; the second, not-so-clades, is a plagiarism of my own. After the successful clone concept was born *(20)*, it was soon evidenced that many bacterial species did not amount to a mere collection of eternal clones *(21)*. Actually, most pathogen species are capable of both clonal propagation and genetic exchange *(7)*. The contribution of each varies between species and, more surprisingly, may vary within the same species between different populations, transmission cycles, and ecosystems *(22)*. These considerations are not relevant only to evolutionists. They have considerable implications for molecular epidemiology/strain typing. Although genetic exchange in pathogens may have various faces, such as conjugation; transformation; transfection in bacteria; meiotic recombination in *Trypanosoma brucei*, the agent of sleeping sickness *(23)*; and nonmeiotic hybridization in *Trypanosoma cruzi*, the agent of Chagas' disease *(10)*, its consequences on population structure are similar. When genetic exchange is frequent, multilocus genotypes are ephemeral, and clades no longer deserve the name since different genetic lineages are only imperfectly separated from each other.

3. Concepts that Proved to be Only Partially Successful: Discrete Typing Units, Tags, and Clonets

Let us summarize the headache: Clades in pathogen species mate with each other from time to time. Even worse, some clades have two ancestors instead of one, as is the case for the remarkable hybrid genotypes recorded in *T. cruzi* ***(10)***. Wisely, Hall and Barlow ***(12)*** called for great caution when performing phylogenetic analysis in those species for which genetic exchange is frequent. Indeed, careless use of such analyses could prove to be grossly misleading. Clades are evolutionary lineages that have one ancestor and are genetically isolated from each other ***(24)***. With this clean definition, for example, even in the case of a species such as *T. cruzi*, in which clonal evolution is preponderant, the genetic subdivisions do not deserve the term of clade ***(22)***. Still in many pathogen species, it is clear that genetic variation is not evenly distributed, and that unambiguous, stable subdivisions are apparent. Such units may be characterizable for epidemiological tracking and may have different relevant properties in terms of ecological distribution, pathogenicity, and so on. Should one renounce the attempt to describe these subdivisions only because the concept of clade is ineffective? Other terms have been proposed but are not satisfactory: "Cluster" is only a visual description of subdivisions within a dendrogram. It is as informative as saying that a cake is divided into slices. "Line" and "lineage" are utterly vague. Mammals are a lineage. The Bourbon kings of France are a lineage as well.

I have proposed the term *discrete typing unit* (DTU) to refer to these stable genetic subdivisions within pathogen species: collections of genotypes that are genetically more similar to each other than to any other collections of genotypes, that appear to persist in space and time, and that can be characterized in common by specific markers, or tags ***(25)***. Although the proposal was well accepted in several congresses, it has proved to be particularly successful only among scientists working on *T. cruzi*. The rest of the literature struggles with a tangle of vague concepts: not-so-clades, lineages, and clusters. This is distressing since DTUs constitute a highly reliable target for molecular epidemiology, and tags are identified and designed with the very goal of characterizing the DTUs specifically.

The clonet ***(26)*** is another partial success story. A *clonet* is a set of genotypes that appear to be identical with a given set of genetic markers in a clonal species. As explained, even in those species that are highly clonal, a clone characterized by a given set of genetic markers may not be a real clone but more probably a family of genetically related clones. The nuance is considerable. Clones characterized with a given technique, for example, a few

primers for randomly amplified polymorphic DNA analysis, might be perfect optical illusions. Depending on the resolution power of the technique used, the common ancestor of the clone might be either a few weeks or months old, which is quite relevant to epidemiological follow-up, or hundreds of years old, which is relevant to the evolutionist but not to the epidemiologist. Before aiming at characterizing clones, it is therefore indispensable ***(i)*** to ascertain that the species and the population under study are truly clonal, which can be done by reliable population genetics analysis only; and ***(ii)*** to scale the resolution power of the markers to be used to the goal of the study.

4. Targeting Relevant Genes Rather than Whole Organisms

The object of molecular epidemiology is chiefly to survey what is medically relevant. What matters for health professionals is pathogenicity and resistance to treatments. Designing sophisticated phylogenies (even so-called not-so-phylogenies), characterizing acutely multilocus genotypes is highly relevant to the evolutionist. It may be less so to the health professional if the genes that drive medically relevant properties evolve largely independently from phylogenies and multilocus genotypes. In the case of bacteria, many important genes are harbored by plasmids, which are able to jump from a bacterial cell to another. Even nuclear genes could jump frequently from one genome to another (horizontal gene transfer), especially if they undergo great selective pressure. In medical research, it is therefore crucial to identify the genes to be followed and to design specific markers for them. Needless to say, however, ***(i)*** this is also very important to the evolutionist, and ***(ii)*** elaborating a sophisticated population genetics framework for the entire species remains quite informative for the follow-up of these culprit genes, precisely to see how independent they are from the general evolution and demography of the host species.

5. The Great Contribution of Evolutionary Biology to Our Knowledge of the Basic Biology of Pathogens

Evolutionary studies still have much to tell us about the world of pathogens. However, a consensus picture has emerged on the reproductive strategy of microbes: Many species play on a double keyboard and are capable of both sexual recombination and clonal propagation. This is wise from an evolutionary point

of view: Sexual recombination serves to quickly generate new genetic combinations able to respond to new selective challenges, and in turn, clonal propagation makes it possible to stabilize in long-term favorable genetic combinations.

The vast majority of pathogens have recombination as a side mechanism, but it is not at all mandatory for their reproduction. It is only a useful last resort on an evolutionary scale to allow successful genotypes to make an appearance. *Trypanosoma cruzi* is an illustrative example. Two of these genotypes that appear to be hybrids behave as conquistadores and, once generated by recombination of two ancestors, now propagate themselves clonally over vast geographical ranges, mainly in human transmission cycles. In spite of these occasional bouts of recombination, the species as a whole is profoundly structured into six persistent DTUs ***(27,28)***, found over the entire geographical reach of the species.

To some extent, this is also true in the case of *E. coli*. Although the subdivisions visible within this species might be less sharply defined than *T. cruzi* DTUs are, it is remarkable that those uncovered by the pioneer isoenzyme work by Ochman and Selander ***(29)*** (A, B1, B2, and D) have been recognized by recent studies relying on totally different molecular tools ***(30,31)***, making these subdivisions perfectly honest DTUs. The population genetics framework thus elaborated for *E. coli* provides a remarkable evolutionary tool for the study of medically relevant features [*Shigella* strains, pathogenicity islands, mutator genes ***(32)***, and antibiotic resistance genes].

Plasmodium falciparum, the agent of the most malignant form of malaria, is another example of the relevance of population biology studies to biomedical research. It has been long considered ***(33)*** the paradigm of a panmictic organism (a species is panmictic when genetic exchange occurs at random with no other obstacles than geographical distance or isolation by time). The cautious proposal that some populations of *P. falciparum* might undergo some kind of uniparental propagation ***(34)*** has received a flurry of blows. However, many, if not most, populations of this parasite show a strong linkage disequilibrium (nonrandom association of genotypes occurring at different loci), which indicates a severe inhibition to recombination in these populations ***(35)***. Whatever the final explanation, the rough data show that many *P. falciparum* populations are by no means panmictic.

It is not an exaggeration to say that evolutionary studies have revolutionized our views on pathogen population biology and dynamics, with considerable payoffs in terms of medical research. It is all the more distressing that evolution science is still not considered a built-in component of molecular epidemiology, as it should be.

6. The Future

It will always be useful to trace pathogen genes and genotypes responsible for epidemics, especially those genotypes designated as "superspreaders," causing the majority of infections in a given species. However, this classical and restricted conception of molecular epidemiology is bound to be outshined. Technological progress makes it possible to envisage a thorough characterization of pathogens, integrating genomic, proteomic, metabolomic, clinical, and epidemiological data. This is rendered accessible by automatic sequencing, microarrays, and geographical information systems. It is the concept of pathogen profiling (***15***). Integrating complex sets of data will be made possible by the emerging Web portals and portals of portals (MLSTNet for MLST data, PulseNet for pulsed-field gel electrophoresis data, among others). It would be a pity not to interpret these abundant and complementary data in terms of evolutionary biology, which could lead to the population genomics and population proteomics needed.

Finally, as already advocated many times (***25***,***36***), research investigating the pathogen, its host, and in the case of vector-borne diseases, its vector, should not be artificially compartmentalized when obviously these organisms do not evolve separately and on the contrary follow a pattern of coevolution. The host is a characteristic of the pathogen, and vice versa, and the same is true for pathogens and vectors. Pathogen profiling could not be complete without parallel evolutionary studies on the host and the vector. The MEEGID (Molecular Epidemiology and Evolutionary Genetics) congresses and the journal *Infection, Genetics and Evolution* are the privileged tribunes for this integrated approach.

References

1. Tibayrenc, M. (2005). Bridging up the gap between molecular epidemiologists and evolutionists. *Trends Microbiol.* **13**, 575–580.
2. Tibayrenc, M. (2007). Molecular epidemiology and evolutionary genetics of pathogens, in *Encyclopedia of Infectious Diseases: Modern Methodologies* (Tibayrenc, M., ed.), Wiley, pp. 337–355.
3. Centers for Disease Control and Prevention. (1994). Addressing emerging infectious diseases threat. A prevention strategy for the United States, **27**.
4. Tibayrenc, M. (1995). Population genetics of parasitic protozoa and other microorganisms. *Adv. Parasitol.* **36**, 47–115.
5. Humphreys, H. (2004). Does molecular typing make any contribution to the care of patients with infection? *Clin. Microbiol. Infect.* **10**, 269–271.
6. Arañaz, A., Liébana, E., Mateos, A., Dominguez, L., Vidal, D., Domingo, M., et al. (1996). Spacer oligonucleotide typing of *Mycobacterium bovis* strains from cattle and other animals: a tool for studying epidemiology of tuberculosis. *J. Clin. Microbiol.* **34**, 2734–2740.
7. Awadalla, P. (2003). The evolutionary genomics of pathogen recombination. *Nat. Rev. Genet.* 4, 50–60.
8. Brosch, R., Gordon, S. V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., et al. (2002). A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc. Natl. Acad. Sci. U S A* **99**, 3684–3689.

9. Frank, S. A. (2002). *Immunology and Evolution of Infectious Disease*. Princeton University Press, Princeton, NJ.
10. Gaunt,M.W.,Yeo,M.,Frame,I.A.,Tothard,J.R., Carrasco, H. J., Taylor, M. C., et al. (2003). Mechanism of genetic exchange in American trypanosomes. *Nature* **421**, 936–939.
11. Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., et al. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. U S A* **95**, 3140–3145.
12. Hall, B., and Barlow, M. (2006). Phylogenetic analysis as a tool in molecular epidemiology of infectious diseases. *Ann. Epidemiol.* **16**, 157–169.
13. Blanc, D. (2004). The use of molecular typing for epidemiological surveillance and investigation of endemic nosocomial infections. *Infect. Genet. Evol.* **4**, 193–197.
14. Foxman, B. (2007). Contributions of molecular epidemiology to the understanding of infectious disease transmission, pathogenesis, and evolution. *Ann. Epidemiol.* **17**, 148–156.
15. Sintchenko, V., Iredell, J. R., and Gilbert, G. L. (2007). Pathogen profiling for disease management and surveillance. *Nat. Rev. Microbiol.* **5**, 464–470.
16. Taylor, J. W., Geiser, D. M., Burt, A., and Koufopanou, V. (1999). The evolutionary biology and population genetics underlying fungal strain typing. *Clin. Microbiol. Rev.* **12**, 126–146.
17. Tibayrenc, M. (2006). The species concept in parasites and other pathogens: a pragmatic approach? *Trends Parasitol.* **22**, 66–70.
18. De Meeûs, T., Durand, P., and Renaud, F. (2003). Species concepts: what for? *Trends Parasitol.* **19**, 425–427.
19. Cracraft, J. (1983). Species concept and speciation analysis, in *Current Ornithology* (Johnson, R. F., ed.), Plenum Press, New York, pp. 159–187.
20. Ørskov, F., and Ørskov, I. (1983). Summary of a workshop on the clone concept in the epidemiology, taxonomy, and evolution of the Enterobacteriaceae and other Bacteria. *J. Infect. Dis.* **148**, 346–357.
21. Maynard Smith, J., Smith, N. H., O'Rourke, M., and Spratt, B. G. (1993). How clonal are bacteria? *Proc. Natl. Acad. Sci. U S A* **90**, 4384–4388.
22. Tibayrenc, M., and Ayala, F. J. (2002). The clonal theory of parasitic protozoa: 12 years on. *Trends Parasitol.* **18**, 405–410.
23. Jenni, L., Marti, S., Schweizer, J., Betschart, B., Le Page, R. W. F., Wells, J. M., et al. (1986). Hybrid formation between African trypanosomes during cyclical transmission. *Nature* **322**, 173–175.
24. Hennig, W. (1966). *Phylogenetic Systematics*. University of Illinois Press, Urbana, IL.
25. Tibayrenc, M. (1998). Genetic epidemiology of parasitic protozoa and other infectious agents: the need for an integrated approach. *Int. J. Parasitol.* **28**, 85–104.
26. Tibayrenc, M., and Ayala, F. J. (1991). Towards a population genetics of microorganisms: the clonal theory of parasitic protozoa. *Parasitol. Today* **7**, 228–232.
27. Barnabé, C., Brisse, S., and Tibayrenc, M. (2000). Population structure and genetic typing of *Trypanosoma cruzi*, the agent of Chagas' disease: a multilocus enzyme electrophoresis approach. *Parasitology* **150**, 513–526.
28. Brisse, S., Barnabé, C., and Tibayrenc, M. (2000). Identification of six *Trypanosoma cruzi* phylogenetic lineages by random amplified polymorphic DNA and multilocus enzyme electrophoresis. *Int. J. Parasitol.* **30**, 35–44.
29. Ochman, H., and Selander, R. K. (1984). Evidence for clonal population structure in *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A* **81**, 198–201.
30. Feil, E. J., Holmes, E. C., Bessen, D. E., Chan, M. S., Day, N. P. J., Enright, M. C., et al. (2001). Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. *Proc. Natl. Acad. Sci. U S A* **98**, 182–187.
31. Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L. H., et al. (2006). Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol. Microbiol.* **60**, 1136–1151.
32. Taddei, F., Matic, I., Godelle, B., and Radman, M. (1997). To be a mutator, or how pathogenic and commensal bacteria can evolve rapidly. *Trends Microbiol.* **5**, 427–428.
33. Walliker, D. (1985). Characterization of *Plasmodium falciparum* of different countries. *Ann. Soc. Belge Méd. Trop.* **65**, 69–77.
34. Tibayrenc, M., Kjellberg, F., and Ayala, F. J. (1990). A clonal theory of parasitic protozoa: the population structure of *Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas and Trypanosoma* , and its medical and taxonomical consequences. *Proc. Natl. Acad. Sci. U S A* 87, 2414–2418.
35. Anderson, T. J., Haubold, B., Williams, J. T., Estrada-Franco Section Sign, J. G., Richardson, L., Mollinedo, R., et al. (2000). Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Mol. Biol. Evol.* **17**, 1467–1482.
36. Tibayrenc, M. (1999). Towards an integrated genetic epidemiology of parasitic protozoa and other pathogens. *Annu. Rev. Genet.* **33**, 449–477.