PLoS PATHOGENS

The Past and Future of Tuberculosis Research

Iñaki Comas, Sebastien Gagneux*

Division of Mycobacterial Research, MRC National Institute for Medical Research, London, United Kingdom

Abstract: Renewed efforts in tuberculosis (TB) research have led to important new insights into the biology and epidemiology of this devastating disease. Yet, in the face of the modern epidemics of HIV/AIDS, diabetes, and multidrug resistance—all of which contribute to susceptibility to TB—global control of the disease will remain a formidable challenge for years to come. New high-throughput genomics technologies are already contributing to studies of TB's epidemiology, comparative genomics, evolution, and host–pathogen interaction. We argue here, however, that new multidisciplinary approaches—especially the integration of epidemiology with systems biology in what we call "systems epidemiology"—will be required to eliminate TB.

Introduction

Tuberculosis (TB) remains an important public health problem [1]. With close to 10 million new cases per year, and a pool of two billion latently infected individuals, control efforts are struggling in many parts of the world (Figure 1). Nevertheless, the renewed interest in research and improved funding for TB give reasons for optimism. Recently, the Stop TB Partnership, a network of concerned governments, organizations, and donors lead by the WHO (http://www.stoptb.org/stop_tb_initiative/), outlined a global plan to halve TB prevalence and mortality by 2015 and eliminate the disease as a public health problem by 2050 [2]. Attaining these goals will depend on both strong government commitment and increased interdisciplinary research and development. As existing diagnostics, drugs, and vaccines will be insufficient to achieve these objectives, a substantial effort in both basic science and epidemiology will be necessary to develop better tools and strategies to control TB [3]. Here we review the recent history of TB research and some of the latest insights into the evolutionary history of the disease. We then discuss ways in which we could benefit from a more comprehensive systems approach to control TB in the future.

Recent History of the Field

TB is caused by several species of gram-positive bacteria known as tubercle bacilli or *Mycobacterium tuberculosis* complex (MTBC). MTBC includes obligate human pathogens such as *Mycobacterium tuberculosis* and *Mycobacterium africanum*, as well as organisms adapted to various other species of mammal. In the developed world, TB incidence declined steadily during the second half of the 20th century and so funds available for research and control of TB decreased substantially during that time [4]. When TB started to reemerge in the early 1990s, fuelled by the growing pandemic of HIV/AIDS (Box 1), scientists and public health officials were caught off-guard; billions of dollars of emergency funds were necessary to control TB outbreaks [5]. Moreover, long-term neglect of basic TB research and product development meant that global TB control relied on a 100-year-old diagnostic method (i.e. sputum smear microscopy) of poor sensitivity, an 80-year-old and

largely ineffective vaccine (Bacille Calmette-Guérin [BCG]), and just a few drugs that were decades old (streptomycin, rifampicin, isoniazid, ethambutol, pyrozinamide) [3]. Tragically, these are the tools still in use today in most parts of the world where TB remains one of the most important public health problems (Figure 1).

In addition to the lack of appropriate tools to control TB globally, much about the disease was unknown in the early 1990s and many dogmas were guiding the field at the time. These included the view that differences in the clinical manifestation of TB were primarily driven by host variables and the environment as opposed to bacterial factors, a notion reinforced by early DNA sequencing studies that reported very limited genetic diversity in MTBC compared with other bacterial pathogens [6]. According to other dogmas, TB was mainly a consequence of reactivation of latent infections rather than ongoing disease transmission, and that mixed infections and exogenous reinfections with different strains were very unlikely.

The development of molecular techniques to differentiate between strains of MTBC made it possible to readdress some of these points. One of these methods, a DNA fingerprinting protocol based on the *Mycobacterium* insertion sequence IS6110, quickly evolved into the first international gold standard for genotyping of MTBC [7]. It also became a key component of pragmatic public health efforts, such as detecting disease outbreaks and ongoing TB transmission [8], and allowed differentiation between patients who relapsed due to treatment failure and those reinfected with a different strain [9]. This latter finding demonstrated for the first time that previous exposure to MTBC does not protect against subsequent exogenous reinfection and TB disease, which is a phenomenon with implications for vaccine design. Many other new insights were gained through these molecular epidemiological studies [10], which, for the most part, were performed in wealthy countries; corresponding data from most high-burden areas remained limited because of poor infrastructure and lack of funding.

Routine genotyping of MTBC for public health purposes also revived discussions about the role of pathogen variation in
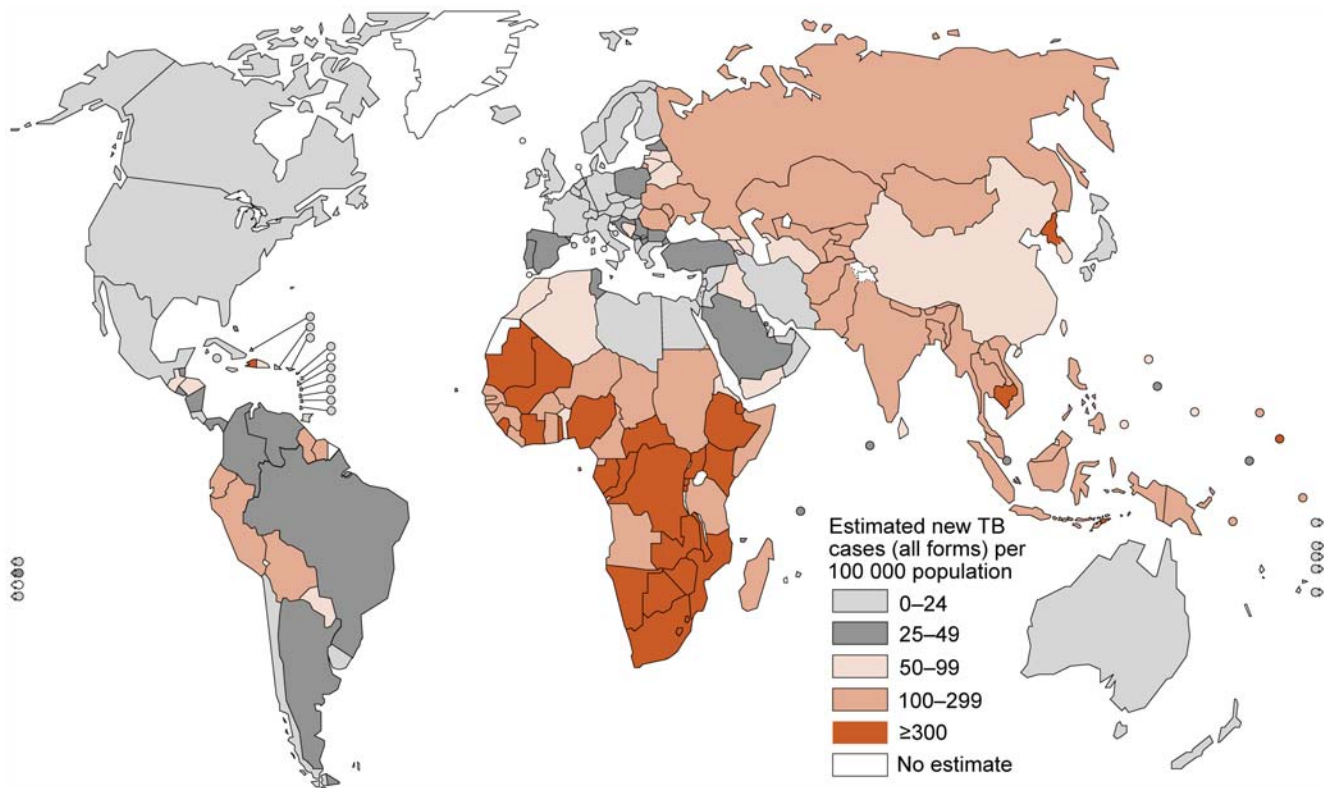
Figure 1. The global incidence of TB. The number of new TB cases per 100,000 population for the year 2007 according to WHO estimates (adapted from [1]).
doi:10.1371/journal.ppat.1000600.g001

outcome of infection and disease. Some strains of MTBC appeared over-represented in particular patient populations, which suggested that strain diversity may have epidemiological implications. The completion of the first whole genome sequence of *M. tuberculosis* in 1998 [11] and the development of DNA microarrays offered a new opportunity to address this question by interrogating the entire genome of multiple clinical strains of MTBC. These comparative genomics studies revealed that genomic deletions, also known as large sequence polymorphisms (LSPs), are an important source of genome plasticity in MTBC [12]. Furthermore, statistical analyses of patient data suggested possible associations between strain genomic content and disease severity in humans [13]. Clinical phenotypes in TB are difficult to

standardize, however, and whether MTBC genotype plays a meaningful role in TB severity remains controversial [14].

Comparative genomics of MTBC also yielded interesting insights into the evolution and geographic distribution of the organism. Because MTBC has essentially no detectible horizontal gene transfer [15,16], LSPs can be used as phylogenetic markers to trace the evolutionary relationships of different strain families. Following such an approach, studies have shown that humans did not, as previously believed, acquire MTBC from animals during the initiation of animal domestication, rather the human- and animal-adapted members of MTBC share a common ancestor, which might have infected humans even before the Neolithic transition [17,18]. LSPs also allowed researchers to define several discrete strain lineages within the human-adapted members of MTBC, which are associated with different human populations and geographical regions (Figures 2 and 3) [15,19,20]. Because of the lack of horizontal gene exchange in MTBC, phylogenetic trees derived using various molecular markers define the same phylogenetic groupings [21], and several studies based on single nucleotide polymorphisms (SNPs) and other molecular makers have gathered additional support for the highly phylogeo-graphical population structure of MTBC [22–25].

Ancient History of the Pathogen

Although LSPs have proven very useful for defining different lineages within MTBC, these markers do not reflect actual genetic distances, and the mode of molecular evolution in MTBC cannot be easily inferred from them [21]. By contrast, DNA sequence-based methods can provide important clues about the evolutionary forces shaping bacterial populations. Multilocus sequence typing (MLST), in which fragments of seven structural genes are

Box 1. The Influence of Modern Epidemics on TB Incidence

HIV/AIDS and diabetes are important comorbidities that dramatically increase the susceptibility to TB. The synergy between TB and HIV/AIDS is a particular problem in sub-Saharan Africa, while the impact of diabetes on TB is increasing in many rapidly growing world economies; it may already be a more important risk factor for TB than HIV/AIDS in places like India and Mexico. The emergence of multidrug-resistant strains represents an additional threat to global TB control. The strong association between HIV/AIDS and drug-resistant TB has been well established, but whether similar interactions exist between drug-resistant TB and diabetes needs to be explored further.
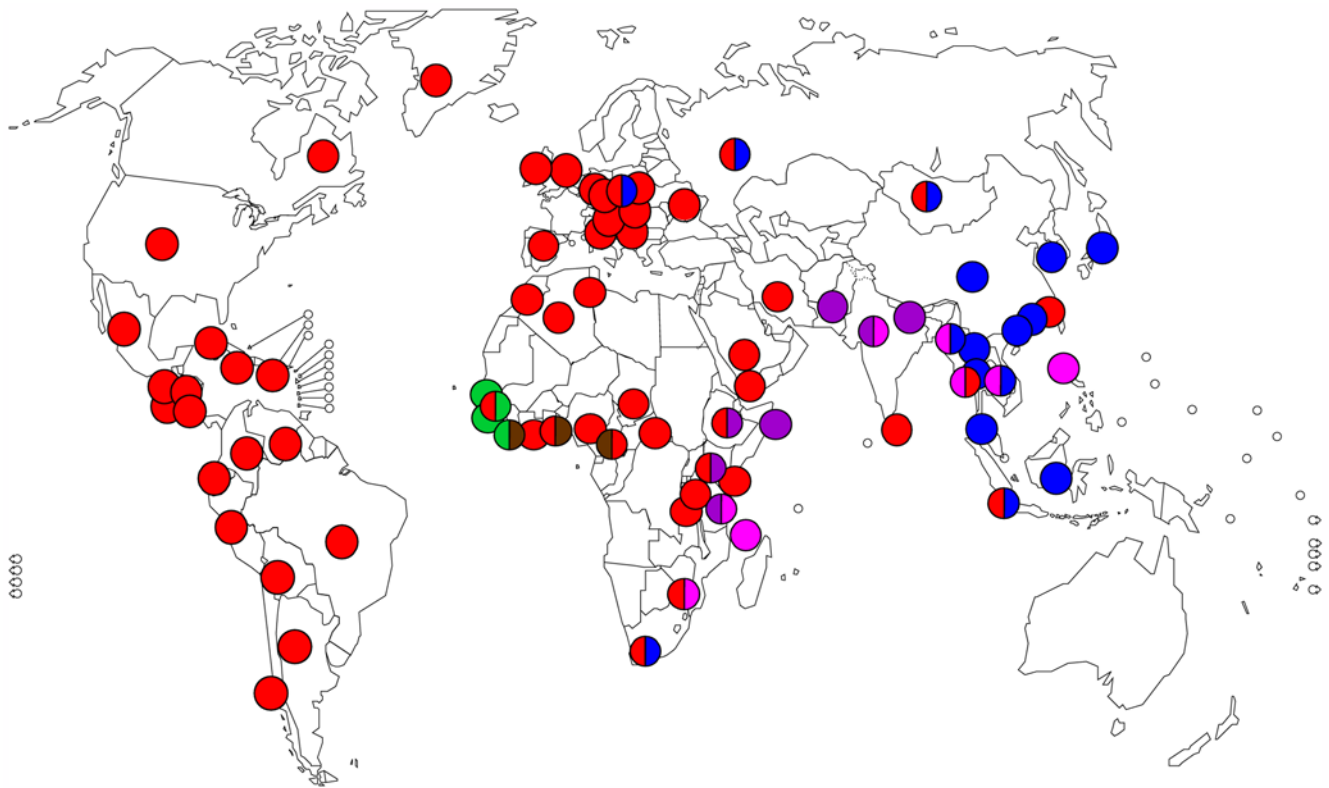
Figure 2. Global distribution of the six main lineages of human MTBC. Each dot represents the most frequent lineage(s) circulating in a country. Colours correspond to the lineages defined in Figure 3 (adapted from [20]).
doi:10.1371/journal.ppat.1000600.g002

sequenced for each strain [26], has been used very successfully to define the genetic population structure of many bacterial species [27]. Because of the low degree of sequence polymorphisms in MTBC, however, standard MLST is uninformative [28]. A recent study of MTBC extended the traditional MLST scheme by sequencing 89 complete genes in 108 strains, covering 1.5% of the genome of each strain [29]. Phylogenetic analysis of this extended multilocus sequence dataset resulted in a tree that was highly congruent with that generated previously using LSPs (Figure 3). The new sequence-based data also revealed that the MTBC strains that are adapted to various animal species represent just a subset of the global genetic diversity of MTBC that affects different human populations [29]. Furthermore, by comparing the geographical distribution of various human MTBC strains with their position on the phylogenetic tree, it became evident that MTBC most likely originated in Africa and that human MTBC originally spread out of Africa together with ancient human migrations along land routes. This view is further supported by the fact that the so-called "smooth tubercle bacilli," which are the closest relatives of the human MTBC, are highly restricted to East Africa [30]. The multilocus sequence data reported by Hershberg et al. [29] further suggested a scenario in which the three "modern" lineages of MTBC (purple, blue, and red in Figure 3) seeded Eurasia, which experienced dramatic human population expansion in more recent times. These three lineages then spread globally out of Europe, India, and China, respectively, accompanying waves of colonization, trade and conquest. In contrast to the ancient human migrations, however, this more recent dispersal of human MTBC occurred primarily along water routes [29].

The availability of comprehensive DNA sequence data has also allowed researchers to address questions about the molecular

evolution of MTBC. In-depth population genetic analyses by Hershberg et al. highlight the fact that purifying selection against slightly deleterious mutations in this organism is strongly reduced compared to other bacteria [29]. As a consequence, nonsynonymous SNPs tend to accumulate in MTBC, leading to a high ratio of nonsynonymous to synonymous mutations (also known as dN/dS). The authors hypothesized that the high dN/dS in MTBC compared to most other bacteria might indicate increased random genetic drift associated with serial population bottlenecks during past human migrations and patient-to-patient transmission. If confirmed, this would indicate that "chance," not just natural selection, has been driving the evolution of MTBC. Although these kinds of fundamental evolutionary questions are often underappreciated by clinicians and biomedical researchers, studying the evolution of a pathogen ultimately allows for better epidemiological predictions by contributing to our understanding of basic biology, particularly with respect to antibiotic resistance.

A Vision for the Future

Thanks to recent increases in research funding for TB [4], substantial progress has been made in our understanding of the basic biology and epidemiology of the disease. Unfortunately, this increased knowledge has not yet had any noticeable impact on the current global trends of TB (Figure 1). While TB incidence appears to have stabilized in many countries, the total number of cases is still increasing as a function of global human population growth [1]. Of particular concern are the ongoing epidemics of multidrug-resistant TB [31], as well as the synergies between TB and the ongoing epidemics of HIV/AIDS and other comorbidities such as diabetes (Box 1).
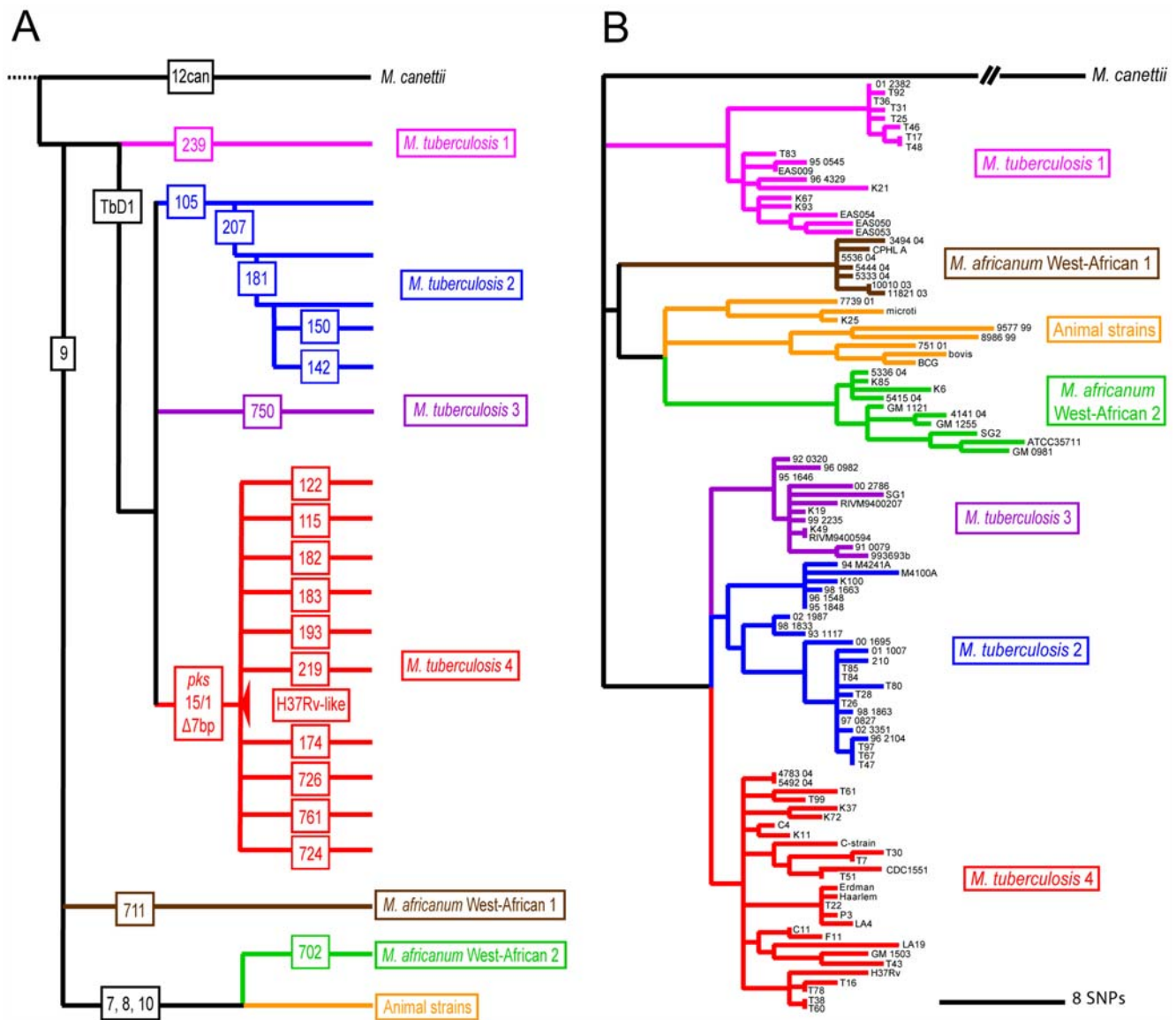
Figure 3. The global phylogeny of *Mycobacterium tuberculosis* complex (MTBC). The phylogenic relationships between various human- and animal-adapted strains and species are largely consistent when defined by using either (A) large sequence polymorphisms (LSPs) or (B) single nucleotide polymorphisms (SNPs) identified by sequencing 89 genes in 108 MTBC strains. Numbers inside the squares in (A) refer to specific lineage-defining LSPs. Colors indicate congruent lineages (adapted from [20] and [29]).
doi:10.1371/journal.ppat.1000600.g003

As our understanding of TB improves, we would like to be able to make better predictions about the future trajectory of the disease and to develop new tools to control the disease better and ultimately reverse global trends. For this to be feasible, TB epidemiology needs to evolve into a more predictive, interdisciplinary endeavour; a discipline we might refer to as "systems epidemiology" (Figure 4). Systems biology is already a rapidly emerging field, in which cycles of mathematical modelling and experiments using various large-scale "-omics" datasets are integrated in an iterative manner [32]. Novel biological processes are being discovered through these systems approaches, which might not have been possible using more traditional methods [33–35].

Last year, Young et al. argued that systems biology approaches will be necessary to elucidate some of the key aspects of host–pathogen interactions in TB [36] and to develop new drugs, vaccines, and biomarkers to evaluate new interventions [3]. For

example, according to another dogma in the TB field, latent TB infections are caused by physiologically dormant bacilli and can thus be differentiated from active disease where MTBC is actively growing and dividing [37]. In reality, however, the phenomenon of TB latency most likely reflects a whole spectrum of responses to TB infection, involving phenotypically distinct bacterial subpopulations and spanning various degrees of bacterial burden and associated host immune responses [38]. We agree with Young et al. [36] that TB latency and similar biological complexities will only be adequately addressed using systems approaches, and we argue further that to comprehend the current TB epidemic as a whole, and to better predict its future trajectory, a complementary systems epidemiology approach will be necessary (Figure 4).

Mathematical models are already being used extensively to study the epidemiology of TB and to guide control policies [39]. Recent applications have shown that socioeconomic factors are key drivers
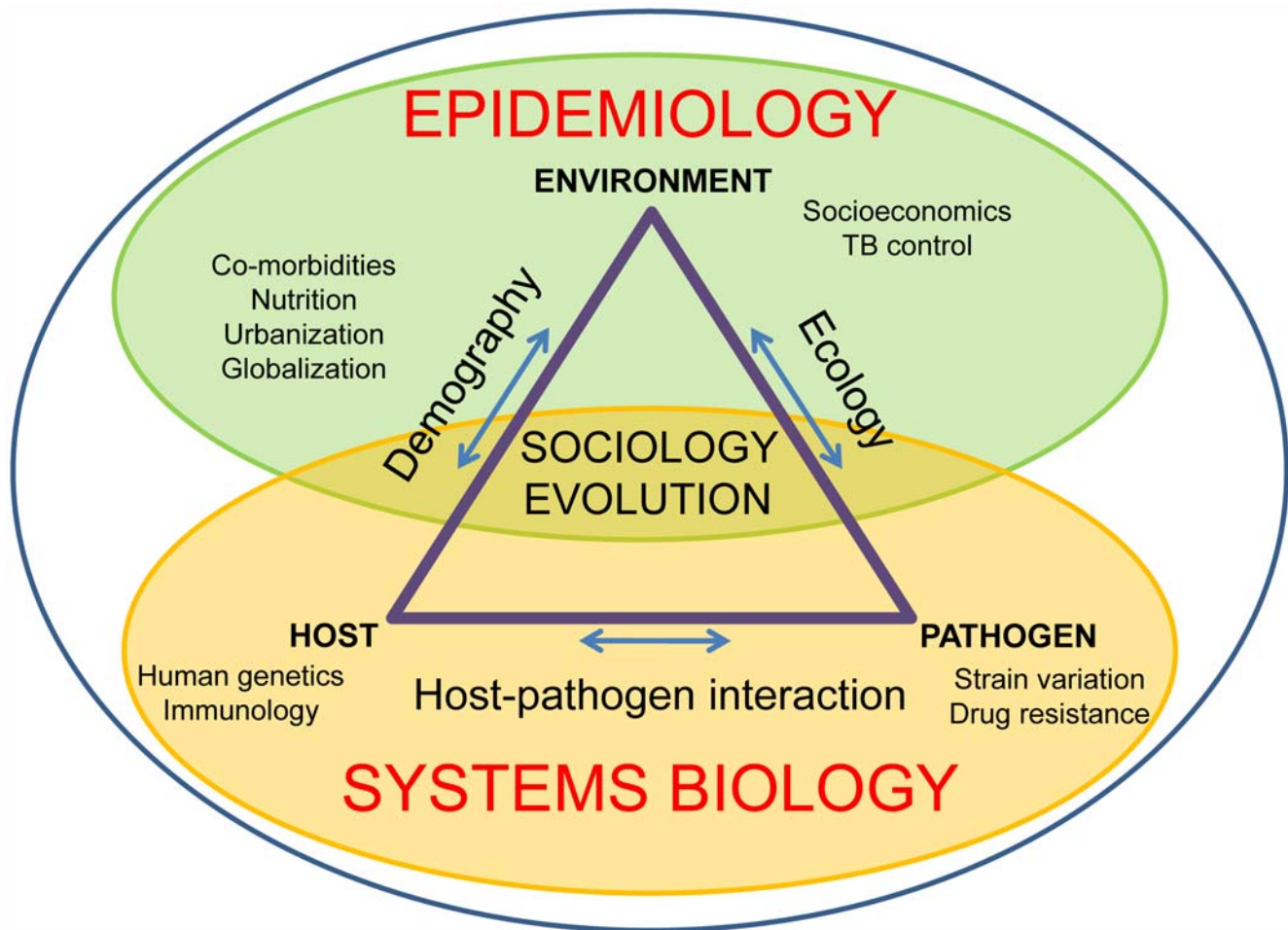
Figure 4. A systems epidemiology approach to TB research. The spread of TB is influenced by social and biological factors. On the one hand, the new discipline of systems biology integrates approaches that address the host, the pathogen, and interactions between the two. On the other hand, epidemiology addresses the burden of the disease and the social, economic, and ecological causes of its frequency and distribution. There is little crosstalk between these two disciplines at the moment. "Systems epidemiology" is an attempt to take into account the interactions between these various fields of research.
doi:10.1371/journal.ppat.1000600.g004

of today's TB epidemic [40]. In addition, much theoretical emphasis has been placed on trying to define the impact that drug resistance will have on the global TB epidemic [41]. Some of this theoretical work has become more complex by incorporating new biological insights obtained empirically and through targeted experimental studies. Early theoretical studies on the spread of drug-resistant MTBC were based on the assumption that all drug-resistant bacteria had an inherent fitness disadvantage compared to drug-susceptible strains [42]; however, as is becoming clear from experimental and molecular epidemiological investigation, substantial heterogeneity exists with respect to the reproductive success of drug-resistant strains [43–46]. Newer mathematical models account for some of this heterogeneity [47–49].

One could imagine an expansion of such mathematical approaches—much as systems biology operates—in which epidemiological modelling is combined with more comprehensive biological data related to the host, the pathogen, and their interactions (Figure 4). Of course, environmental and sociological data would also need to be considered [40]. As mathematical models become more finely tuned, they could in turn inform future experimental work to test some of the specific predictions. The genomics revolution now offers the opportunity to study host–

pathogen interactions at an unprecedented depth. To be able to make sense out of the current and upcoming deluge of -omics data, however, scientists will have to rely on a mathematically and statistically robust analytical framework. Ideally, some of these theoretical approaches will be able to accommodate increasingly diverse sets of data in order to capture the various biological, environmental, and social aspects of TB.

Among the newly emerging technologies, we believe that next-generation DNA sequencing will play an important role in improving our understanding of TB [50]. Whole-genome sequencing could potentially become the new gold standard for strain typing in routine molecular epidemiology [51]. For host genetics and TB susceptibility, too, de novo DNA sequencing based approaches could have advantages over traditional SNP typing [52]. For example, many of the human populations carrying the largest proportion of the global TB burden have not been sufficiently characterised genetically (Figure 1) [53,54], and screening for currently limited human SNP collections might have little relevance for these populations [55]. Furthermore, comprehensive DNA sequencing of TB patients and controls in various human populations could help unveil rare but biologically relevant mutations [56]. Another approach increasingly being

used to study both the host and the pathogen is sequence-based transcriptomics, in which gene expression is measured by whole genome sequencing of RNA transcripts; a method referred to as RNA-seq [57]. One of the advantages of this approach over existing microarray-based methods is that changes in the expression of noncoding RNAs and other novel transcripts can be easily detected. RNA-seq is particularly useful for genome-wide studies of small regulatory RNAs, as such studies are more difficult to perform using standard DNA microarrays. Recent studies, for example, have reported a role for small regulatory RNAs in *M. tuberculosis* [58], and there is little doubt more regulatory RNAs will soon be identified by RNA-seq [57].

Challenges for the Future

Advances in TB research are hampered by the fact that MTBC is a Biosafety Level 3 pathogen with a long generation time, making it slow and complex to culture. Moreover, TB is a chronic disease that can develop over many years, and is characterised by extended periods of latency during which MTBC cannot be isolated from infected individuals. All of these factors complicate and prolong the development of new interventions and their assessment in clinical trials. As we have already mentioned, the field has been marked by a number of dogmas that, in some cases, might have contributed to the slow progress in TB research. New insights are now questioning some of these views, but at the same time, new opinions could well evolve into new dogmas. For example, we and others have spent much of our scientific careers seeking convincing evidence for the role of MTBC strain diversity in human disease. Although some pieces of evidence have recently started to emerge [59–61], the subject needs more work. One of the problems has been that the macrophage and mouse infection models used in these studies relied on poorly characterised strains, and finding relevant links to human disease has been all but impossible [14,21].

In TB control, too, potential new dogmas might emerge to limit future progress. A strong T cell–derived interferon gamma (INFγ) response appears to be crucial for the immunological control of TB, and many MTBC antigens have been identified based on their capacity to elicit INFγ responses in TB patients or their infected contacts [62]. Some of these antigens are being developed into new TB diagnostics and vaccines, but the potential impact of MTBC diversity on immune responses is not generally being considered [21]. A recent study in The Gambia showed that INFγ responses to one of the key MTBC antigens differed in an MTBC lineage–specific manner [63]. Developing a universally effective vaccine might be the only way to eliminate TB in the future [3]. This is particularly true given the large reservoir of latently infected individuals in the world, which would be impossible to eliminate through prophylactic drug treatment. Considering that natural TB infection does not protect against exogenous reinfection and disease, however, mimicking natural infection using attenuated strains or a cocktail of traditional INFγ-inducing antigens might not necessarily be the most promising vaccine strategy. Indeed, the largely unsuccessful implementation of BCG vaccination might serve as a warning [64].

References

1. World Health Organization (2009) Global tuberculosis control - surveillance, planning, financing. Geneva, Switzerland: WHO.
2. Stop TB Partnership (2006) The global plan to stop TB 2006–2015. Geneva: WHO.
3. Young DB, Perkins MD, Duncan K, Barry CE (2008) Confronting the scientific obstacles to global control of tuberculosis. J Clin Invest 118: 1255–1265.
4. Kaufmann SH, Parida SK (2007) Changing funding patterns in tuberculosis. Nat Med 13: 299–303.
5. Frieden TR, Fujiwara PI, Washko RM, Hamburg MA (1995) Tuberculosis in New York City–turning the tide. N Engl J Med 333: 229–233.
6. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, et al. (1997) Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. Proc Natl Acad Sci U S A 94: 9869–9874.
7. van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, et al. (1993) Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J Clin Microbiol 31: 406–409.
8. Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, et al. (1994) The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods. N Engl J Med 330: 1703–1709.
9. Small PM, Shafer RW, Hopewell PC, Singh SP, Murphy MJ, et al. (1993) Exogenous reinfection with multidrug-resistant *Mycobacterium tuberculosis* in patients with advanced HIV infection. N Engl J Med 328: 1137–1144.
10. Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN (2006) Molecular epidemiology of tuberculosis: current insights. Clin Microbiol Rev 19: 658–685.
11. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393: 537–544.
12. Tsolaki AG, Hirsh AE, DeRiemer K, Enciso JA, Wong MZ, et al. (2004) Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains. Proc Natl Acad Sci U S A 101: 4865–4870.
13. Kato-Maeda M, Rhee JT, Gingeras TR, Salamon H, Drenkow J, et al. (2001) Comparing genomes within the species *Mycobacterium tuberculosis*. Genome Res 11: 547–554.
14. Nicol MP, Wilkinson RJ (2008) The clinical consequences of strain diversity in *Mycobacterium tuberculosis*. Trans R Soc Trop Med Hyg 102: 955–65.
15. Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small PM (2004) Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. Proc Natl Acad Sci U S A 101: 4871–4876.

16. Supply P, Warren RM, Banuls AL, Lesjean S, Van Der Spuy GD, et al. (2003) Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area. Mol Microbiol 47: 529–538.
17. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, et al. (2002) A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. Proc Natl Acad Sci U S A 99: 3684–3689.
18. Mostowy S, Cousins D, Brinkman J, Aranaz A, Behr MA (2002) Genomic deletions suggest a phylogeny for the *Mycobacterium tuberculosis* complex. J Infect Dis 186: 74–80.
19. Reed MB, Pichler VK, McIntosh F, Mattia A, Fallow A, et al. (2009) Major *Mycobacterium tuberculosis* lineages associate with patient country of origin. J Clin Microbiol 47: 1119–28.
20. Gagneux S, Deriemer K, Van T, Kato-Maeda M, de Jong BC, et al. (2006) Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A 103: 2869–2873.
21. Gagneux S, Small PM (2007) Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. Lancet Infect Dis 7: 328–337.
22. Baker L, Brown T, Maiden MC, Drobniewski F (2004) Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. Emerg Infect Dis 10: 1568–1577.
23. Gutacker MM, Mathema B, Soini H, Shashkina E, Kreiswirth BN, et al. (2006) Single-nucleotide polymorphism-based population genetic analysis of *Mycobacterium tuberculosis* strains from 4 geographic sites. J Infect Dis 193: 121–128.
24. Filliol I, Motiwala AS, Cavatore M, Qi W, Hernando Hazbon M, et al. (2006) Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. J Bacteriol 188: 759–772.
25. Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, et al. (2006) *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC Microbiol 6: 23.
26. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, 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.
27. Maiden MC (2006) Multilocus sequence typing of bacteria. Annu Rev Microbiol 60: 561–588.

28. Achtman M (2008) Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu Rev Microbiol 62: 53–70.
29. Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, et al. (2008) High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. PLoS Biol 6: e311.
30. Gutierrez C, Brisse S, Brosch R, Fabre M, Omais B, et al. (2005) Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. PLoS Pathogens 1: 1–7.
31. World Health Organization (2008) Anti-tuberculosis drug resistance in the world report no. 4. Geneva, Switzerland: WHO.
32. Zak DE, Aderem A (2009) Systems biology of innate immunity. Immunol Rev 227: 264–282.
33. Gilchrist M, Thorsson V, Li B, Rust AG, Korb M, et al. (2006) Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. Nature 441: 173–178.
34. Querec TD, Akondy RS, Lee EK, Cao W, Nakaya HI, et al. (2009) Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. Nat Immunol 10: 116–125.
35. Stuart LM, Boulais J, Charriere GM, Hennessy EJ, Brunet S, et al. (2007) A systems biology analysis of the Drosophila phagosome. Nature 445: 95–101.
36. Young D, Stark J, Kirschner D (2008) Systems biology of persistent infection: tuberculosis as a case study. Nat Rev Microbiol 6: 520–8.
37. Gill WP, Harik NS, Whiddon MR, Liao RP, Mittler JE, et al. (2009) A replication clock for *Mycobacterium tuberculosis*. Nat Med 15: 211–4.
38. Young DB, Gideon HP, Wilkinson RJ (2009) Eliminating latent tuberculosis. Trends Microbiol 17: 183–188.
39. Cohen T, Dye C, Colijn C, Murray M (2009) Mathematical models of the epidemiology and control of drug-resistant TB. Expert Rev Resp Med in press.
40. Lonnroth K, Jaramillo E, Williams BG, Dye C, Raviglione M (2009) Drivers of tuberculosis epidemics: The role of risk factors and social determinants. Soc Sci Med 68: 2240–6.
41. Borrell S, Gagneux S (2009) Infectiousness, reproductive fitness, and evolution of drug-resistant *Mycobactyerium tuberculosis*. Int J Tuberc Lung Dis in press.
42. Dye C, Williams BG, Espinal MA, Raviglione MC (2002) Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis. Science 295: 2042–2046.
43. Bottger EC, Springer B, Pletschette M, Sander P (1998) Fitness of antibiotic-resistant microorganisms and compensatory mutations. Nat Med 4: 1343–1344.
44. Gagneux S, Burgos MV, DeRiemer K, Encisco A, Munoz S, et al. (2006) Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis*. PLoS Pathog 2: e61.
45. Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, et al. (2006) The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. Science 312: 1944–1946.
46. van Soolingen D, de Haas PE, van Doorn HR, Kuijper E, Rinder H, et al. (2000) Mutations at amino acid position 315 of the katG gene are associated with high-level resistance to isoniazid, other drug resistance, and successful transmission of *Mycobacterium tuberculosis* in the Netherlands. J Infect Dis 182: 1788–1790.
47. Cohen T, Murray M (2004) Modeling epidemics of multidrug-resistant *M. tuberculosis* of heterogeneous fitness. Nat Med 10: 1117–1121.
48. Blower SM, Chou T (2004) Modeling the emergence of the 'hot zones': tuberculosis and the amplification dynamics of drug resistance. Nat Med 10: 1111–1116.
49. Dye C (2009) Doomsday postponed? Preventing and reversing epidemics of drug-resistant tuberculosis. Nat Rev Microbiol 7: 81–87.
50. Mardis ER (2008) Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 9: 387–402.
51. MacLean D, Jones JD, Studholme DJ (2009) Application of 'next-generation' sequencing technologies to microbial genetics. Nat Rev Microbiol 7: 287–296.
52. Hardy J, Singleton A (2009) Genomewide association studies and human disease. N Engl J Med 360: 1759–1768.
53. Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, et al. (2009) The genetic structure and history of Africans and African Americans. Science 324: 1035–44.
54. Basu A, Mukherjee N, Roy S, Sengupta S, Banerjee S, et al. (2003) Ethnic India: a genomic view, with special reference to peopling and structure. Genome Res 13: 2277–2290.
55. Campbell MC, Tishkoff SA (2008) African Genetic Diversity: Implications for human demographic history, modern human origins, and complex disease mapping. Annu Rev Genomics Hum Genet 9: 403–33.
56. Goldstein DB (2009) Common genetic variation and human traits. N Engl J Med 360: 1696–1698.
57. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57–63.
58. Arnvig KB, Young DB (2009) Identification of small RNAs in *Mycobacterium tuberculosis*. Mol Microbiol 73: 397–408.
59. de Jong BC, Hill PC, Aiken A, Awine T, Antonio M, et al. (2008) Progression to active tuberculosis, but not transmission, varies by *Mycobacterium tuberculosis* lineage in the Gambia. J Infect Dis 198: 1037–43.
60. Caws M, Thwaites G, Dunstan S, Hawn TR, Thi Ngoc Lan N, et al. (2008) The influence of host and bacterial genotype on the development of disseminated disease with *Mycobacterium tuberculosis*. PLoS Pathog 4: e1000034.
61. Thwaites G, Caws M, Chau TT, D'Sa A, Lan NT, et al. (2008) The relationship between *Mycobacterium tuberculosis* genotype and the clinical phenotype of pulmonary and meningeal tuberculosis. J Clin Microbiol 46: 1363–8.
62. Ernst JD, Lewinsohn DM, Behar S, Blythe M, Schlesinger LS, et al. (2007) Meeting report: NIH workshop on the Tuberculosis Immune Epitope Database. Tuberculosis (Edinb) 88: 366–70.
63. de Jong BC, Hill PC, Brookes RH, Gagneux S, Jeffries DJ, et al. (2006) *Mycobacterium africanum* elicits an attenuated T Cell response to Early Secreted Antigenic Target, 6 kDa, in patients with tuberculosis and their household contacts. J Infect Dis 193: 1279–1286.
64. Andersen P, Doherty TM (2005) Opinion: The success and failure of BCG - implications for a novel tuberculosis vaccine. Nat Rev Microbiol 3: 656–62.