

*Review*

# Host–pathogen coevolution in human tuberculosis

Sebastien Gagneux\*

*Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute and University of Basel, Socinstrasse 57, 4002 Basel, Switzerland*

Tuberculosis (TB) is a disease of antiquity. Yet TB today still causes more adult deaths than any other single infectious disease. Recent studies show that contrary to the common view postulating an animal origin for TB, *Mycobacterium tuberculosis* complex (MTBC), the causative agent of TB, emerged as a human pathogen in Africa and colonized the world accompanying the Out-of-Africa migrations of modern humans. More recently, evolutionarily ‘modern’ lineages of MTBC expanded as a consequence of the global human population increase, and spread throughout the world following waves of exploration, trade and conquest. While epidemiological data suggest that the different phylogenetic lineages of MTBC might have adapted to different human populations, overall, the phylogenetically ‘modern’ MTBC lineages are more successful in terms of their geographical spread compared with the ‘ancient’ lineages. Interestingly, the global success of ‘modern’ MTBC correlates with a hypo-inflammatory phenotype in macrophages, possibly reflecting higher virulence, and a shorter latency in humans. Finally, various human genetic variants have been associated with different MTBC lineages, suggesting an interaction between human genetic diversity and MTBC variation. In summary, the biology and the epidemiology of human TB have been shaped by the long-standing association between MTBC and its human host.

**Keywords:** mycobacteria; infection; virulence; genotyping; phylogeography; fitness

## 1. INTRODUCTION

Tuberculosis (TB) remains a major global health problem. Ten million new TB cases and 2 million deaths are estimated to occur each year, more than any time in history [1]. Furthermore, an estimated 2 billion people are thought to be latently infected, providing a large reservoir for active TB that will last for decades [2]. Even though during the last centuries TB used to cause up to half of all human deaths in Europe and North-America [3], today the disease affects primarily the developing world. Notably, some of the largest emerging economies including China and India remain among the 22 high-burden countries for TB identified by the World Health Organization (WHO). These two countries collectively accounted for more than one-third of all new TB cases in 2009 [1].

TB is the prototype of a disease of poverty [4]. In addition to the well-established environmental and socio-economical risk factors for TB [5], recent times have witnessed the appearance of new forces driving the global TB pandemic: HIV/AIDS and the emergence of antimicrobial resistance. Since its initial appearance in the 1980s, the HIV/AIDS pandemic had a dramatic effect on TB incidence, particularly in sub-Saharan Africa [6]. Moreover, the introduction of

chemotherapy against TB, which started in the 1940s, has led to increasing levels of drug resistance [7]. The global emergence of multi-drug resistance has initiated the post-antibiotic era, making TB essentially incurable in many parts of the world [8]. Left untreated TB kills 50 per cent of patients on average [9], with HIV co-infected patients facing an even higher risk of death [10]. The only available vaccine against TB is ‘bacille Calmette Guérin’ (BCG), an attenuated form of *Mycobacterium bovis*, which is a pathogen of cattle (further discussed below) [11]. However, BCG only protects young children against TB meningitis, the most severe form of the disease. Why BCG does not protect adults reliably against the classical pulmonary form of TB is unknown [12].

TB is an ancient disease [13]. Recent data show that, contrary to the traditional belief linking the origin of TB to the development of agriculture and animal domestication during the Neolithic transition approximately 10 000 years ago, TB has been affecting humans for much longer. Hence, the long-term interaction between the TB microbe and its human host raises the question as to whether coevolution has occurred and if it can be detected. Here, I start by summarizing some of the most important features of TB and the agents causing it. I then discuss several recent findings which support the notion of coevolution in human TB. Some of the points discussed are purely speculative. Yet, I am addressing them in the hope of fuelling future research.

\*sebastien.gagneux@unibas.ch

One contribution of 14 to a Discussion Meeting Issue ‘Immunity, infection, migration and human evolution’.

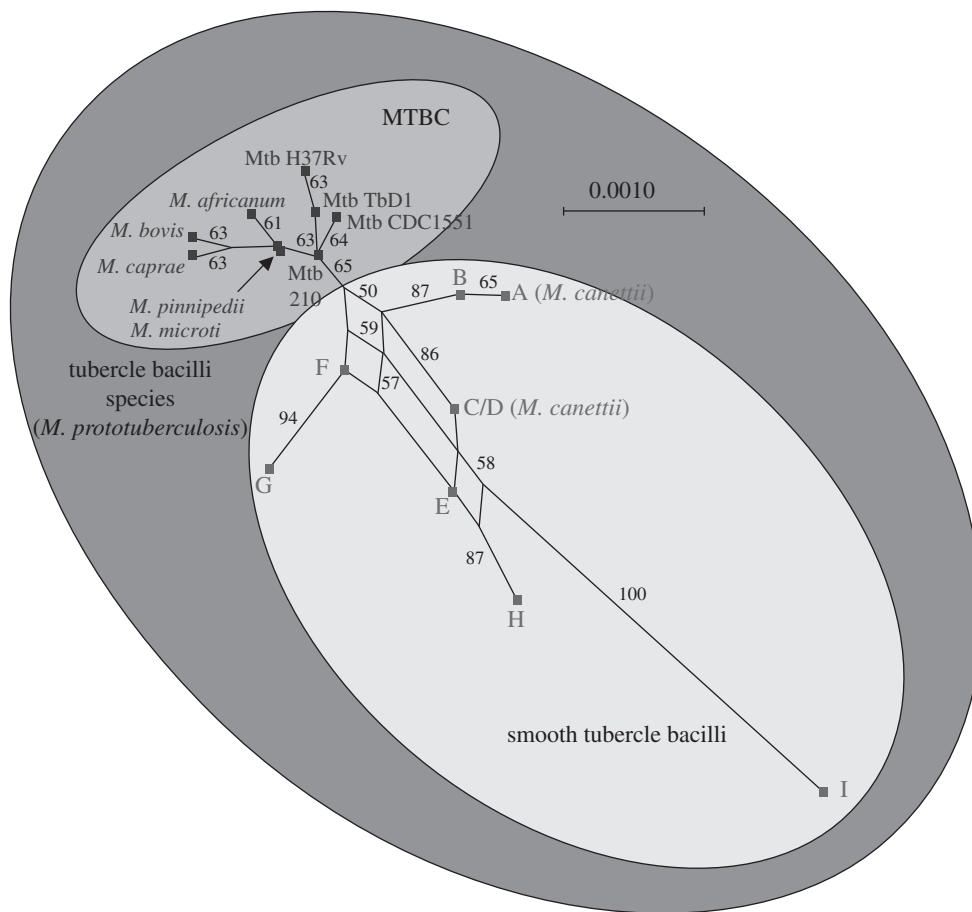


Figure 1. Phylogenetic relationship between ‘smooth’ tubercle bacilli and classical MTBC. Multi-locus sequence data of six housekeeping genes show that the smooth tubercle bacilli are more genetically diverse than MTBC. The network structure of the phylogenetic relationships between members of the smooth tubercle bacilli indicates ongoing horizontal gene exchange, which is absent in MTBC. Adapted from Gutierrez *et al.* [19].

## 2. TUBERCULOSIS AND THE *MYCOBACTERIUM TUBERCULOSIS* COMPLEX

TB is caused by a group of phylogenetically closely related bacteria, collectively known as the *Mycobacterium tuberculosis* complex (MTBC) [14]. TB in humans is primarily caused by *M. tuberculosis* and *Mycobacterium africanum*, a phylogenetic variant of MTBC limited to West Africa [15]. In addition, several animal-adapted members of MTBC exist, which affect a range of wild and domestic animal species [16]. These include *M. bovis* (a pathogen of cattle), *Mycobacterium caprae* (sheep and goats), *Mycobacterium microti* (voles) and *Mycobacterium pinnipedii* (seals and sea lions). *Mycobacterium bovis* used to be a significant cause of human TB, primarily in children who consumed raw milk [17]. However, *M. bovis* infections in humans decreased markedly following the introduction of pasteurization and meat-control practices. Moreover, *M. bovis* does not easily transmit between humans and, similarly, while *M. tuberculosis* has been isolated from various animal species, including cattle, there is currently no evidence of animal-to-animal transmission of *M. tuberculosis* or *M. africanum*. Hence, the different members of MTBC appear to be best adapted to their particular host species [18].

One particular member of MTBC deserves special mention: *Mycobacterium canettii*, which is part of the so-called ‘smooth TB bacilli’ [19]. This somewhat

enigmatic microbe was first described in the 1960s, and so far only about 60 isolates have been reported [20]. Intriguingly, the large majority of these were recovered from TB patients in Djibouti or from individuals who spent some time at the Horn of Africa prior to developing TB. *Mycobacterium canettii* differs in many ways from classical MTBC. This organism produces smooth and shiny colonies on solid growth media, which are distinct from the rough colony morphology characteristic of classical MTBC. Moreover, *M. canettii* and the other ‘smooth TB bacilli’ harbour much more genetic diversity compared with classical MTBC (figure 1). Most importantly, *M. canettii* shows clear evidence of ongoing horizontal gene exchange (figure 1), which does not occur in classical MTBC [21,22]. This latter observation, combined with the lack of evidence for human-to-human transmission of *M. canettii*, and the fact that *M. canettii* clinical isolates are rare, suggests that this organism is an opportunist, and that an environmental reservoir exists somewhere in the Horn of Africa [23].

By contrast, *M. tuberculosis* and *M. africanum* are obligate human pathogens with limited survival outside of the human body and no known animal reservoir [15]. Moreover, these microbes have to cause active disease in order to transmit to secondary hosts [24]. In other words, in MTBC there is a direct link between virulence and transmission,

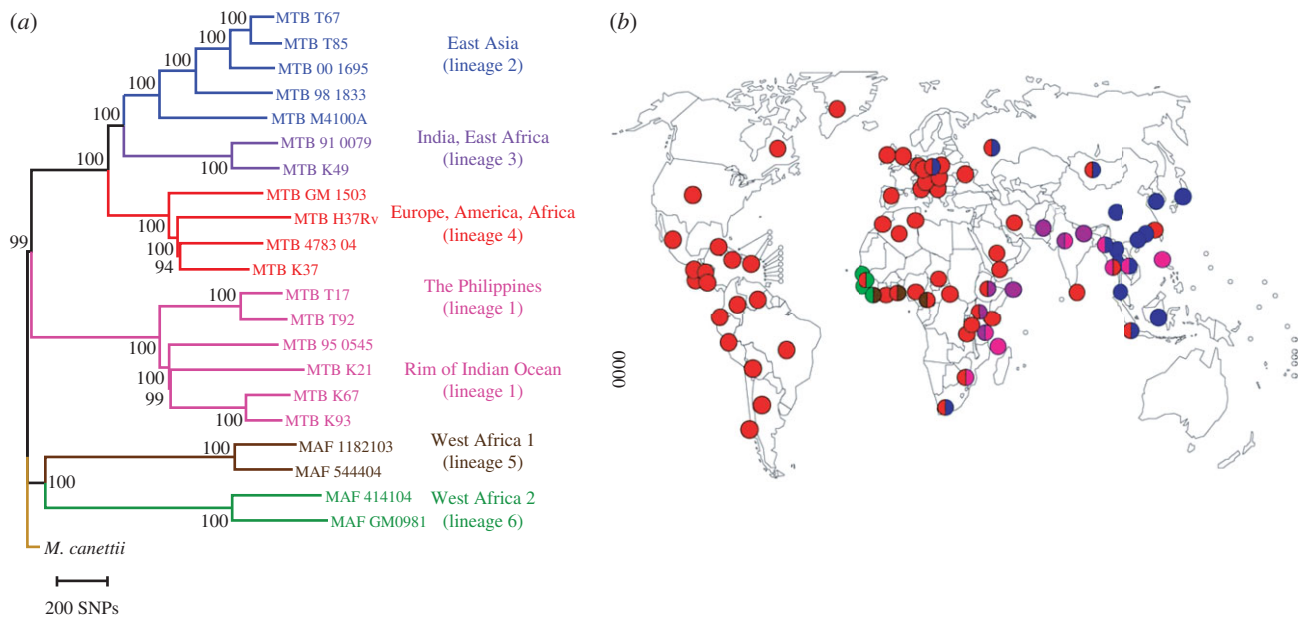


Figure 2. Global phylogeography of MTBC. (a) Global phylogeny of human-adapted MTBC is based on 22 whole-genome sequences. Adapted from Comas *et al.* [35]. (b) Global distribution of the six main human-adapted MTBC lineages. Coloured dots represent the dominant MTBC lineage(s) in each country. Adapted from Gagneux *et al.* [36].

which is in contrast to many other pathogens [25,26]. Because latently infected individuals are considered non-infectious [2], global efforts in controlling TB almost exclusively focus on detecting and treating patients with active TB [1]. However, because culture facilities are often not available, most places still diagnose TB by sputum-smear microscopy. Unfortunately, this technique suffers from limited sensitivity and up to 50 per cent of active TB cases are routinely missed. Individuals harbouring the largest numbers of acid-fast bacilli in their sputum are considered the most infectious [27].

MTBC can infect virtually any organ. However, from a public health point of view, classical pulmonary TB is the most important form of the disease because of its infectious nature. One of the main immunopathological features of active pulmonary TB is the formation of lung cavities. These structures tend to harbour large numbers of bacteria, which is why cavitary TB tends to be the most infectious form of the disease. By contrast, extrapulmonary TB is generally non-infectious and therefore has a lower public health priority [1].

### 3. THE ORIGIN OF HUMAN TUBERCULOSIS

Until a few years ago, the prevalent view was that TB originated in animals and was transferred to humans during the Neolithic transition [28]. However, comparative genomics and population genetic studies have challenged this notion [29,30]. In particular, *M. bovis* and the other animal-adapted members of MTBC have a genome which is about 60 000 base pairs smaller than the genome of the human-adapted *M. tuberculosis* [14,31]. Because MTBC exhibits essentially no ongoing horizontal gene exchange [21], the genomic regions present in *M. tuberculosis* but absent from *M. bovis* represent deletions in the latter rather than insertions in the former. More recently,

phylogenetic analyses of 89 gene concatenates in 108 strains representative of the global MTBC diversity revealed that the animal-adapted members of the MTBC clustered together, representing an ingroup with respect to the global phylogeny of MTBC, while all other taxa represented human-adapted forms [32]. In summary, although *M. tuberculosis* and *M. bovis* do share a common ancestor, the most parsimonious scenario suggests that humans gave TB to animals rather than the other way around [28]. This latter observation also suggests that TB in humans pre-dates the Neolithic transition.

That human TB appears to be significantly older than approximately 10 000 years is supported by several additional lines of evidence. A recent study from Turkey reported a 500 000 year old fossil of *Homo erectus*, which shows lesions characteristic of TB [33]. Maybe not surprisingly, this finding is controversial, as it suggests that TB predated the origin of modern humans. Similarly, a study by Gutierrez *et al.* [19] proposed that *M. canettii* and the other smooth TB bacilli represent a group of organisms likely to contain the ancestor of MTBC. The authors used DNA sequence data of six housekeeping genes and calculated an approximate age of 3 million years for the origin of what they have been referring to as '*Mycobacterium prototuberculosis*' (figure 1). However, these dating results, too, have been contested [34].

More recently, Comas *et al.* [35] generated the first whole-genome-based global phylogeny of human-adapted MTBC (figure 2a). This analysis revealed that the two MTBC lineages generally referred to as *M. africanum* (lineage 5 and 6 in figure 2a) are the most basal. These two lineages occur almost exclusively in West Africa for reasons unknown (figure 2b) [15]. Furthermore, Africa is the only region of the world that harbours all six main human-adapted MTBC lineages (figure 2b) [37]. In other words, Africa harbours the largest diversity of human-adapted

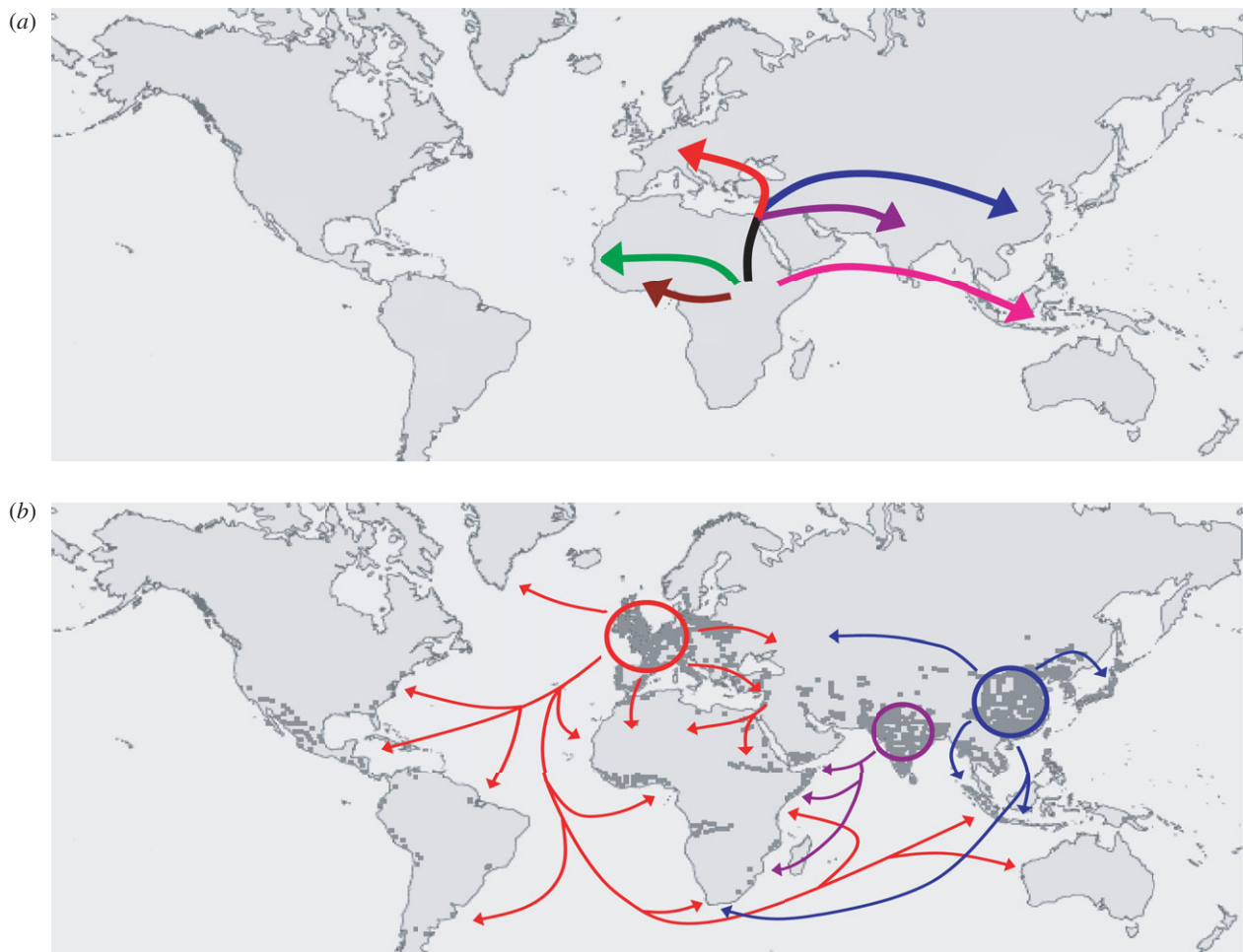


Figure 3. ‘Out-of-and-back-to-Africa’ scenario for the evolutionary history of human TB. Adapted from Hershberg *et al.* [32]. (a) MTBC originated in Africa and some lineages accompanied the Out-of-Africa migrations of modern humans. (b) The three evolutionary ‘modern’ MTBC lineages seeded Europe, India and China, respectively, and expanded as a consequence of the sharp increases in human populations in these regions starting a few centuries ago (each dark grey dot corresponds to 1 million people). These lineages then spread throughout the world via exploration, trade and conquest. Lineage colours correspond to figure 2.

MTBC in the world. Both of these observations are reminiscent of the situation in modern humans; i.e. *Homo sapiens* from Africa are known to be phylogenetically ‘ancestral’ and harbour most of the known human genetic diversity [38]. Finally, human-adapted MTBC exhibits a phylogeographic population structure with different lineages associated with different human populations (figure 2) [36].

Based on these observations, Hershberg *et al.* [32] proposed an evolutionary scenario for the origin and global spread of human TB. This scenario was supported by multi-locus sequence data from 108 global MTBC strains and postulates that human MTBC originated in Africa and accompanied the Out-of-Africa migrations of modern humans approximately 70 000 years ago (figure 3). While two phylogenetically ‘ancient’ MTBC lineages remained in Africa, other lineages left Africa and spread into Eurasia where the three phylogenetically ‘modern’ lineages seeded Europe, India and China, respectively. These areas experienced strong human population growth during the last few centuries. As a consequence, the MTBC populations in these areas expanded, and concomitantly spread globally through waves of human

exploration, trade and conquest. In support of this scenario, a study by Wirth *et al.* [39] based on mini-satellite data in another global collection of MTBC isolates detected molecular signatures of recent (i.e. less than 200 years) population expansion in the different MTBC populations. These signatures were more pronounced in MTBC populations from Europe and Asia compared with those from Africa.

Assuming the ‘Out-of-and-back-to-Africa’ scenario of TB evolution by Hershberg *et al.* is correct, then several important questions remain. For a long time, the general view was that TB did not exist in the Americas before European contact [40]. This notion was to a large extent based on the high susceptibility to TB observed in Native Americans. This gave rise to the so-called ‘Virgin soil’ hypothesis, according to which Native Americans were immunologically naive because they (presumably) had not been exposed to TB before European contact [41]. This view has now changed quite radically, with an emerging consensus among historians and paleopathologists suggesting that TB was in fact present in pre-Columbian America [42]. Note that Native Americans remain highly susceptible to TB, but other, more general risk factors

for TB are increasingly being invoked to explain this phenomenon, including poor living and nutritional conditions and alcohol abuse [41]. Even though ancient DNA studies were able to detect mycobacterial DNA in pre-Columbian human remains [43], the more detailed characteristics of this DNA remains unknown. In other words, we do not know what the exact etiologic agent of TB was at the time [42]. Today, by far, most TB in the Americas is caused by lineage 4, also known as the Euro-American lineage (figure 2*b*) [37]. Future studies will tell whether pre-Columbian TB was caused by MTBC related to Asian forms, as would be expected given the original human colonization of the Americas via the Bering Strait. Alternatively, pre-Columbian TB might have been caused by mycobacterial lineages which are now extinct, perhaps because they were outcompeted by lineage 4/Euro-American lineage following the massive influx of Europeans with active or latent TB into the Americas between the early eighteenth and early twentieth century [44].

Another open question is why despite centuries of cross-Atlantic slave trade, *M. africanum* lineages are almost never isolated in the Americas [15]. It is reasonable to assume that a proportion of the 10–15 million African slaves taken to the Americas would have been latently infected with TB. Similar to the argument above, perhaps, *M. africanum* in the Americas was outcompeted by lineage 4 following the European exodus into America. Interestingly, there is evidence that at least in some areas of West Africa, *M. africanum* has been decreasing and *M. tuberculosis* increasing over the last few decades [45]. We will return to the phenotypic differences between *M. africanum* and ‘modern’ MTBC which could account for this phenomenon further below.

In summary, the available evidence suggests that human MTBC originated in Africa and has been infecting humans for millennia. Considering this long-standing host–pathogen association, one wonders whether some degree of coevolution might have occurred between MTBC and its human host.

#### 4. SIGNS OF HOST–PATHOGEN COEVOLUTION IN HUMAN TUBERCULOSIS

Host–pathogen coevolution can broadly be defined as ‘reciprocal, adaptive genetic changes in interacting host and pathogen species’ [46]. While such changes will be difficult to detect in humans, approaching the subject from the pathogen side appears more promising. Over recent years, data have accumulated suggesting that the different MTBC lineages might be adapted to different human populations. Not only does MTBC exhibit a global biogeographic population structure (figure 2*b*), but the associations between the particular MTBC lineages and human populations are maintained in cosmopolitan settings where human populations and their associated MTBC strains experience at least some degree of intermingling [21,47,48]. Moreover, a study in San Francisco found that MTBC lineages primarily transmitted among their sympatric host populations (figure 3) [36]. Intriguingly, transmission of allopatric strains was strongly associated with known risk factors for TB such as HIV co-infection.

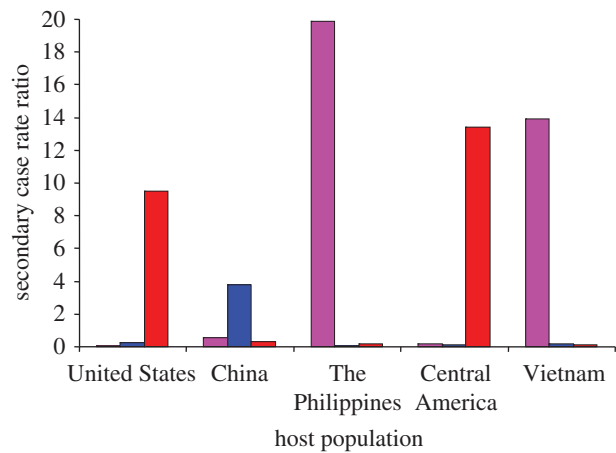


Figure 4. Preferential sympatric transmission of MTBC lineages in San Francisco. The number of secondary TB cases generated between 2001 and 2009 is plotted as a function of the number of index cases (secondary case rate ratio) and the human population in which these secondary cases occurred. In each human population, the number of secondary cases caused by sympatric MTBC lineages was significantly higher compared with the secondary cases caused by allopatric lineages. Sympatric and allopatric lineages were defined based on their usual phylogeographic association shown in figure 2. (Red, lineage 4 or Euro-American lineage; Blue, lineage 2 or East-Asian lineage; Pink, lineage 1 or Indo-Oceanic lineage.) Adapted from Gagneux *et al.* [36].

Whether these findings reflect true biological phenomena rather than mere sociological factors need to be explored further. Nevertheless, they are consistent with host-specific adaptation of MTBC lineages, which if confirmed, would also be consistent with past host–pathogen coevolution.

#### 5. IMMUNE SUBVERSION INSTEAD OF EVOLUTIONARY ARMS-RACE

If host-specific pathogen adaptation has occurred in human MTBC, we would be most likely to detect the associated molecular changes in parts of the MTBC genome that interact with the human immune system, i.e. in genes encoding antigens. More generally, if host–pathogen coevolution is occurring in MTBC, we should be able to detect its signature in antigenic genes. One of the key features of host–pathogen coevolution in other disease systems is the ongoing evolutionary arms-race between the pathogen and the host immune system [49]. Many pathogenic viruses, bacteria and protozoa evade host immunity by varying their antigenic genes. Is MTBC using a similar strategy to evade human immune responses? Comas *et al.* [35] aimed at answering this question by comparing the genetic diversity in different gene classes of MTBC. The authors sequenced the genomes of 21 human MTBC strains representative of the global diversity of MTBC. As a proof of concept, Comas *et al.* first showed that, as expected, essential genes were more evolutionarily conserved than non-essential genes (figure 4). Surprisingly however, the authors found that the known human T-cell epitopes of MTBC were the most conserved regions of the MTBC genome (figure 4). In

fact, more than 95 per cent of the 491 individual epitopes analysed had no amino acid change at all. Based on their findings, the authors speculated that the immune responses elicited by these T-cell epitopes might in fact be beneficial to the bacteria rather than to the host. In other words, rather than escaping host immunity, MTBC wants to be recognized because the ensuing host immune responses contribute to tissue destruction and the formation of cavities in the host lung, which ultimately enhances transmission [50]. The notion that T-cell immunity contributes to the formation of lung cavities in TB patients is supported by the observation that HIV-positive TB patients with low CD4 T-cell counts are less likely to present with cavitory TB than HIV-positive patients with higher CD4 cell counts [6]. In summary, the reason these T-cell epitopes are evolutionarily hyper-conserved might be that they perform an essential function in the sense of contributing to the successful spread of MTBC. This of course does not exclude the possibility that MTBC might harbour antigenic variation in other regions of its genome [51].

## 6. HUMAN DEMOGRAPHY AND THE EVOLUTION OF VIRULENCE

Despite the lack of genetic diversity in the known T-cell epitopes, much genetic diversity exists within and between MTBC lineages, which so far has not been explored experimentally [32]. This diversity could be at the basis of the phenotypic variation, which is increasingly being reported in human-adapted MTBC [52,53]. Our unpublished data from whole genome comparisons in MTBC show that two human-adapted MTBC strains can differ by more than 1700 single-nucleotide polymorphisms (SNPs), which is close to the genetic distance between *M. tuberculosis* and the cattle-adapted *M. bovis*. Furthermore, about two-thirds of these SNPs are non-synonymous (figure 4), and the majority of these non-synonymous SNPs are predicted to have a functional consequence [32]. Thus, any of these differences on their own, or in combination, could influence the various MTBC phenotypes [52].

As mentioned above, the six main human-adapted lineages of MTBC can be divided into evolutionarily ‘ancient’ and ‘modern’. As can be seen from figure 2, the three ‘modern’ lineages are overall more successful than the other lineages in terms of their geographical spread [37]. What could account for this differential success? According to the evolutionary scenario by Hershberg *et al.* [32] (figure 3), the ‘modern’ lineages expanded as a consequence of the sharp increases in human populations during the last few centuries in Europe, India and China. Ecological theory predicts that when virulence is positively correlated with transmission, as is the case in TB, access to a larger number of susceptible hosts favours higher virulence and a shorter latency period [26]. If we believe the proposed coevolutionary history of MTBC with humans, most of this evolutionary history would have occurred during hunter–gatherer times, when human population densities were low. In fact, it has been argued that the characteristic latency period of human TB, followed

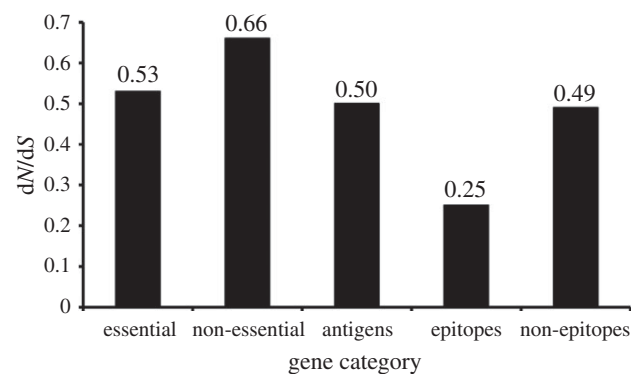


Figure 5. Human T-cell epitopes of MTBC are evolutionarily hyper-conserved. Ratio of non-synonymous mutations to synonymous mutations in different gene classes based on whole genome sequencing of 22 MTBC strains representative of the global MTBC diversity. Adapted from Comas *et al.* [35].

by reactivation and active disease several decades later, might have evolved to allow MTBC to access new birth cohorts of susceptible hosts and, at the same time, avoid a burn-out situation whereby all susceptible hosts would be decimated by an overly virulent pathogen [54]. By contrast, high host densities, similar to the ones encountered in the crowded cities of eighteenth and nineteenth century Europe [3], might have selected for less ‘prudent’ forms of MTBC [55], as access to susceptible hosts was no more a limiting factor. Based on this scenario, and because the ‘modern’ MTBC lineages were exposed to this rapid human expansion, we would predict that ‘modern’ MTBC will be more virulent and associated with a shorter latency period compared with ‘ancient’ MTBC. Intriguingly, several recent data from the laboratory and the field support this prediction.

## 7. FROM MYCOBACTERIUM TUBERCULOSIS COMPLEX GENOTYPE TO PHENOTYPE

The first human cells encountered by MTBC during the course of an initial infection are believed to be alveolar macrophages [24]. These cells engulf MTBC but cannot destroy them. In fact, MTBC evolved mechanisms to inhibit endosome maturation and phagosome–lysosome fusion, which allows the bacteria to survive and even multiply within these macrophages [56]. Human monocyte-derived macrophages (MDMs) have thus been used as a model to study the cell biology and early mechanisms of infection in TB. Portevin *et al.* [57] recently used such a model to study the innate immune responses to 28 MTBC strains spanning the whole global diversity of human MTBC. The authors used MDMs from eight different healthy donors, infected them with one of each of these 28 strains, and measured a series of cytokines and chemokines after 24 h of infection. The results showed that the MDMs infected with different MTBC strains differed markedly in the levels of cytokines and chemokines produced. Interestingly, MDMs infected with ‘modern’ strains produced significantly less pro-inflammatory cytokines compared with ‘ancient strains’ (figure 5). The same results were

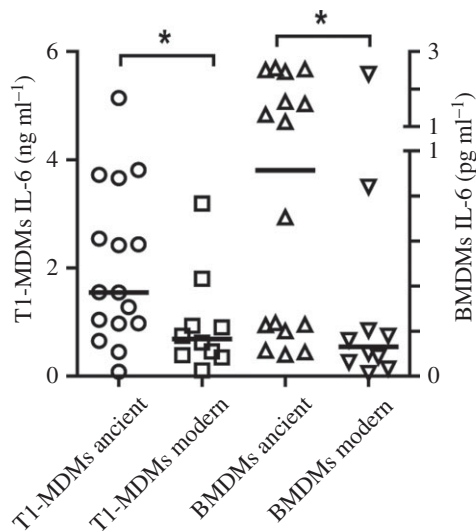


Figure 6. Evolutionarily ‘modern’ MTBC lineages are hypo-inflammatory in human and murine macrophages. Human monocyte-derived (MDM) and murine bone marrow-derived (BMDM) macrophages were infected with one of 28 MTBC strains representative of the global MTBC diversity. Cytokines were measured in supernatants after 24 h of infection. \* $p < 0.05$ , Mann–Whitney U test. Adapted from Portevin *et al.* [57].

obtained when infecting human monocyte-derived dendritic cells or bone marrow-derived murine macrophages. Furthermore, the results were reproducible across the eight human donors. Previous work has shown that low pro-inflammatory innate immune responses to MTBC infection were associated with a higher virulence in animal models [58,59]. Moreover, work from the 1960s showed that MTBC strains from South India were less virulent in guinea pigs than strains from the UK [60]. We now know that these South-Indian strains belonged to the ‘ancient’ lineage 1 while most strains in the UK belong to the ‘modern’ lineage 4 (figure 2).

In a molecular epidemiological study of TB transmission in the Gambia, de Jong *et al.* [61] recruited 317 TB patients and followed their 1808 HIV-negative household contacts during 2 years. The authors then categorized these contacts depending on the MTBC lineage the index case was infected with, and determined how many contacts would become newly infected (i.e. indicating transmission from the index case) and how many would develop active TB. The results showed that, while there was no difference in transmission, contacts exposed to ‘modern’ MTBC were more likely to develop active TB compared with individuals exposed to ‘ancient’ MTBC (figure 6).

Taken together, these observations are consistent with the notion that ‘modern’ strains have evolved a strategy of higher virulence and shorter latency, perhaps to better adapt to the increasing availability of susceptible hosts brought by the large human population increases of the last few centuries.

## 8. THE ROLE OF HUMAN GENETIC DIVERSITY

So far our discussion has primarily focused on the pathogen, maybe because the case for unilateral ‘host-specific pathogen adaptation’ might be more plausible

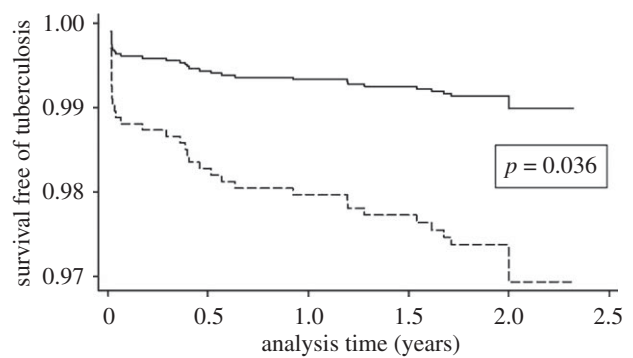


Figure 7. Evolutionarily ‘modern’ MTBC lineages are associated with a shorter latency in human TB. TB patients of 1808 HIV-negative household contacts were followed over 2 years. The number of active TB cases occurring in these contacts during the study period was determined and stratified by the MTBC lineage of the index cases (solid line, ‘ancient’ MTBC (i.e. *M. africanum*), dashed line, ‘modern’ MTBC. Adapted from de Jong *et al.* [61].

and easier to demonstrate than true reciprocal ‘coevolution’ between MTBC and humans, given the shorter generation-time of the former. Nevertheless, variation in TB susceptibility across humans is well established and best illustrated by early twin studies [62]. Since then, many TB susceptibility and resistance loci have been reported in humans [63]. Except for a few extreme and very rare cases of Mendelian disorders, TB is a classical example of a complex disease in which each individual TB-associated genetic locus contributes a small proportion of the observed variation in disease susceptibility. Ten years ago, a theoretical study explored whether the natural selection for resistance to TB could have been sufficiently intense to account for the sharp decrease in TB-associated mortality during the later half of the nineteenth century in Europe [3]. The authors concluded that even when applying a time window of 300 years, time would not have been enough to increase the frequency of TB resistance loci to any significant level in these European populations [64]. However, we now know that TB has been affecting humans for millennia, albeit perhaps imposing lower mortality on low-density hunter–gatherer populations (figure 7).

One of the most important questions in TB research is why only approximately 10 per cent of latently infected individuals develop active disease any time during their lifetime [12]. If we knew the answer to this question, we would be in a much better position to design an effective TB vaccine [65]. From an evolutionary point of view, the low attack rate in active TB suggests that MTBC is not a particularly good pathogen in terms of causing disease, yet, considering the 2 billion latently infected individuals, MTBC is extraordinarily successful at generating secondary infections once active TB develops. On a more provocative and highly speculative note, infection with MTBC might in fact be beneficial to at least some human beings (i.e. individuals who never develop active TB), perhaps because a low level of constant immune-stimulation might protect against other diseases. It is well known that the BCG vaccine strain is a potent immune adjuvant which is used to treat bladder cancer. Nevertheless, considering

the high global morbidity and mortality owing to TB, it would seem inappropriate to label MTBC as a mere commensal or even a symbiont.

Getting back to the coevolutionary question, some evidence is gradually accumulating which suggests that human genetic variation is indeed interacting with MTBC diversity. At least four recent studies have reported associations between particular human genetic variants in immunity-related genes and different MTBC lineages [66–69]. For example, one study in Vietnam found that a particular SNP in Toll-like receptor 2, a molecule known to be key in innate immune recognition of MTBC, was associated with infection by MTBC lineage 2 in TB patients. Future studies will tell whether these associations reflect reciprocal genetic changes characteristic of host–pathogen evolution [46].

## 9. CONCLUDING REMARKS

In summary, current evidence suggests that human TB originated in Africa and that MTBC has been associated with its human host for a very long time. It is therefore plausible that this extended host–pathogen association has facilitated some degree of coevolution, and data are accumulating that support this notion. Furthermore, changes in human demography could also have played a role in the evolution of MTBC's life history. The thought that 'modern' MTBC has gained an evolutionary advantage compared with 'ancient' strains is supported by the larger geographical spread of the former. In addition, 'ancient' MTBC lineages appear to be more constrained to their particular geography and associated human populations, while 'modern' strains seem to be more promiscuous. Hence from an ecological point of view, 'ancient' strains might be referred to as 'specialists', while 'modern' MTBC might represent 'generalists' [70]. Looking into the future and considering the continued global human population increase, coupled with ongoing globalization and urbanization, one wonders where the evolution of human TB is headed. Will MTBC become even more virulent and will particular strain variants become the dominant forms all over the world [71]? Although WHO has devised an ambitious plan to eliminate TB as a global public health problem by 2050 [72], some of the questions discussed here, while primarily of academic interest, should be taken into account by TB researchers and public health officials alike [73]. Answering them will help develop better tools and strategies to control one of the world's oldest and yet still most important human diseases.

I thank Douglas Young and all members of my group for stimulating discussions and comments on the manuscript. Work in my laboratory is supported by the Medical Research Council, UK (MRC\_U117588500), the Leverhulme-Royal Society Africa Award (AA080019), the Swiss National Science Foundation (PP00A-119205) and the US National Institutes of Health (AI090928 and AI034238).

## REFERENCES

- 1 WHO 2010 *Global tuberculosis control—surveillance, planning, financing*. Geneva, Switzerland: WHO. (Contract No: WHO/HTM/TB/2008.393.)
- 2 Barry 3rd, C. E., Boshoff, H. I., Dartois, V., Dick, T., Ehrt, S., Flynn, J., Schnappinger, D., Wilkinson, R. J. & Young, D. 2009 The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat. Rev. Microbiol.* **7**, 845–855. (doi:10.1038/nrmicro2236)
- 3 Wilson, L. G. 1990 The historical decline of tuberculosis in Europe and America: its causes and significance. *J. Hist. Med. Allied Sci.* **45**, 366–396. (doi:10.1093/jhmas/45.3.366)
- 4 Dye, C. & Williams, B. G. 2010 The population dynamics and control of tuberculosis. *Science* **328**, 856–861. (doi:10.1126/science.1185449)
- 5 Lonnroth, K., Jaramillo, E., Williams, B. G., Dye, C. & Raviglione, M. 2009 Drivers of tuberculosis epidemics: the role of risk factors and social determinants. *Soc. Sci. Med.* **68**, 2240–2246. (doi:10.1016/j.socscimed.2009.03.041)
- 6 Kwan, C. K. & Ernst, J. D. 2011 HIV and tuberculosis: a deadly human syndemic. *Clin. Microbiol. Rev.* **24**, 351–376. (doi:10.1128/CMR.00042-10)
- 7 Borrell, S. & Gagneux, S. 2009 Infectiousness, reproductive fitness and evolution of drug-resistant *Mycobacterium tuberculosis*. *Int. J. Tuberc. Lung Dis.* **13**, 1456–1466.
- 8 Raviglione, M. C. & Smith, I. M. 2007 XDR tuberculosis—implications for global public health. *N. Engl. J. Med.* **356**, 656–659. (doi:10.1056/NEJMp068273)
- 9 Murray, C. J., Styblo, K. & Rouillon, A. 1990 Tuberculosis in developing countries: burden, intervention and cost. *Bull. Int. Union Tuberc. Lung Dis.* **65**, 6–24.
- 10 Gandhi, N. R., Moll, A., Sturm, A. W., Pawinski, R., Govender, T., Laloo, U., Zeller, K., Andrews, J. & Friedland, G. 2006 Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet* **368**, 1575–1580. (doi:10.1016/S0140-6736(06)69573-1)
- 11 Andersen, P. & Doherty, T. M. 2005 Opinion: the success and failure of BCG—implications for a novel tuberculosis vaccine. *Nat. Rev. Microbiol.* **3**, 656–662. (doi:10.1038/nrmicro1211)
- 12 Russell, D. G., Barry 3rd, C. E. & Flynn, J. L. 2010 Tuberculosis: what we don't know can, and does, hurt us. *Science* **328**, 852–856. (doi:10.1126/science.1184784)
- 13 Donoghue, H. D., Spigelman, M., Greenblatt, C. L., Lev-Maor, G., Bar-Gal, G. K., Matheson, C., Vernon, K., Gnerlich, A. & Rzink, A. 2004 Tuberculosis: from prehistory to Robert Koch, as revealed by ancient DNA. *Lancet Infect. Dis.* **4**, 584–592. (doi:10.1016/S1473-3099(04)01133-8)
- 14 Cole, S. T. *et al.* 1998 Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544.
- 15 de Jong, B. C., Antonio, M. & Gagneux, S. 2010 *Mycobacterium africanum*—review of an important cause of human tuberculosis in West Africa. *PLoS Negl. Trop. Dis.* **4**, e744. (doi:10.1371/journal.pntd.0000744)
- 16 Smith, N. H., Kremer, K., Inwald, J., Dale, J., Driscoll, J. R., Gordon, S. V., van Soolingen, D., Glyn Hewinson, R. & Maynard Smith, J. 2005 Ecotypes of the *Mycobacterium tuberculosis* complex. *J. Theor. Biol.* **239**, 220–225. (doi:10.1016/j.jtbi.2005.08.036)
- 17 Grange, J. M. 2001 *Mycobacterium bovis* infection in human beings. *Tuberculosis (Edinb)*. **81**, 71–77. (doi:10.1054/tube.2000.0263)
- 18 Smith, N. H., Gordon, S. V., de la Rua-Domenech, R., Clifton-Hadley, R. S. & Hewinson, R. G. 2006 Bottle-necks and broomsticks: the molecular evolution of *Mycobacterium bovis*. *Nat. Rev. Microbiol.* **4**, 670–681. (doi:10.1038/nrmicro1472)
- 19 Gutierrez, C., Brisse, S., Brosch, R., Fabre, M., Omais, B., Marmiesse, M., Supply, P. & Vincent, V. 2005



- Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. *PLoS Pathogens* **1**, 1–7. (doi:10.1371/journal.ppat.0010005)
- 20 Fabre, M., Hauck, Y., Soler, C., Koeck, J. L., van Ingen, J., van Soolingen, D., Vergnaud, G. & Pourcel, C. 2010 Molecular characteristics of ‘*Mycobacterium canettii*’ the smooth *Mycobacterium tuberculosis* bacilli. *Infect. Genet. Evol.* **10**, 1165–1173. (doi:10.1016/j.meegid.2010.07.016)
  - 21 Hirsh, A. E., Tsolaki, A. G., DeRiemer, K., Feldman, M. W. & Small, P. M. 2004 Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. *Proc. Natl Acad. Sci. USA* **101**, 4871–4876. (doi:10.1073/pnas.0305627101)
  - 22 Supply, P., Warren, R. M., Banuls, A. L., Lesjean, S., Van Der Spuy, G. D., Lewis, L. A., Tibayrenc, M., Van Helden, P. D. & Locht, C. 2003 Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area. *Mol. Microbiol.* **47**, 529–538. (doi:10.1046/j.1365-2958.2003.03315.x)
  - 23 Koeck, J.-L. *et al.* 2010 Clinical characteristics of the smooth tubercle bacilli ‘*Mycobacterium canettii*’ infection suggest the existence of an environmental reservoir. *Clin. Microbiol. Infect.* **17**, 1013–1019. (doi:10.1111/j.1469-0691.2010.03347.x)
  - 24 Russell, D. G. 2011 *Mycobacterium tuberculosis* and the intimate discourse of a chronic infection. *Immunol. Rev.* **240**, 252–268. (doi:10.1111/j.1600-065X.2010.00984.x)
  - 25 Ebert, D. & Bull, J. J. 2003 Challenging the trade-off model for the evolution of virulence: is virulence management feasible? *Trends Microbiol.* **11**, 15–20. (doi:10.1016/S0966-842X(02)00003-3)
  - 26 Levin, B. R. 1996 The evolution and maintenance of virulence in microparasites. *Emerg. Infect. Dis.* **2**, 93–102. (doi:10.3201/eid0202.960203)
  - 27 Behr, M. A., Warren, S. A., Salamon, H., Hopewell, P. C., Ponce, dL., Daley, C. L. & Small, P. M. 1999 Transmission of *Mycobacterium tuberculosis* from patients smear-negative for acid-fast bacilli. *Lancet* **353**, 444–449. [Erratum *Lancet* 1999 **353**, 1714.] (doi:10.1016/S0140-6736(98)03406-0)
  - 28 Smith, N. H., Hewinson, R. G., Kremer, K., Brosch, R. & Gordon, S. V. 2009 Myths and misconceptions: the origin and evolution of *Mycobacterium tuberculosis*. *Nat. Rev. Microbiol.* **7**, 537–544. (doi:10.1038/nrmicro2165)
  - 29 Brosch, R. *et al.* 2002 A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc. Natl Acad. Sci. USA* **99**, 3684–3689. (doi:10.1073/pnas.052548299)
  - 30 Mostowy, S., Cousins, D., Brinkman, J., Aranaz, A. & Behr, M. A. 2002 Genomic deletions suggest a phylogeny for the *Mycobacterium tuberculosis* complex. *J. Infect. Dis.* **186**, 74–80. (doi:10.1086/341068)
  - 31 Garnier, T. *et al.* 2003 The complete genome sequence of *Mycobacterium bovis*. *Proc. Natl Acad. Sci. USA* **100**, 7877–7882. (doi:10.1073/pnas.1130426100)
  - 32 Hershberg, R. *et al.* 2008 High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol.* **6**, e311. (doi:10.1371/journal.pbio.0060311)
  - 33 Roberts, C. A., Pfister, L. A. & Mays, S. 2009 Letter to the editor: was tuberculosis present in *Homo erectus* in Turkey? *Am. J. Phys. Anthropol.* **139**, 442–444. (doi:10.1002/ajpa.21056)
  - 34 Smith, N. H. 2006 A re-evaluation of *M. prototuberculosis*. *PLoS Pathog.* **2**, e98. (doi:10.1371/journal.ppat.0020098)
  - 35 Comas, I., Chakravarti, J., Small, P. M., Galagan, J., Niemann, S., Kremer, K., Ernst, J. D. & Gagneux, S. 2010 Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat. Genet.* **42**, 498–503. (doi:10.1038/ng.590)
  - 36 Gagneux, S. *et al.* 2006 Variable host–pathogen compatibility in *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **103**, 2869–2873. (doi:10.1073/pnas.0511240103)
  - 37 Gagneux, S. & Small, P. M. 2007 Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect. Dis.* **7**, 328–337. (doi:10.1016/S1473-3099(07)70108-1)
  - 38 Tishkoff, S. A. *et al.* 2009 The genetic structure and history of Africans and African Americans. *Science* **324**, 1035–1044. (doi:10.1126/science.1172257)
  - 39 Wirth, T. *et al.* 2008 Origin, spread and demography of the *Mycobacterium tuberculosis* complex. *PLoS Pathog.* **4**, e1000160. (doi:10.1371/journal.ppat.1000160)
  - 40 Paulsen, H. J. 1987 Tuberculosis in the native American: indigenous or introduced? *Rev. Infect. Dis.* **9**, 1180–1186. (doi:10.1093/clinids/9.6.1180)
  - 41 McMillen, C. W. 2008 ‘The red man and the white plague’: rethinking race, tuberculosis, and American Indians, ca. 1890–1950. *Bull. Hist. Med.* **82**, 608–645. (doi:10.1353/bhm.0.0094)
  - 42 Stone, A. C., Wilbur, A. K., Buikstra, J. E. & Roberts, C. A. 2009 Tuberculosis and leprosy in perspective. *Am. J. Phys. Anthropol.* **140**, 66–94. (doi:10.1002/ajpa.21185)
  - 43 Salo, W. L., Aufderheide, A. C., Buikstra, J. & Holcomb, T. A. 1994 Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. *Proc. Natl Acad. Sci. USA* **91**, 2091–2094. (doi:10.1073/pnas.91.6.2091)
  - 44 McNeill, W. 1984 Human migration in historical perspective. *Popul. Dev. Rev.* **10**, 1–18. (doi:10.2307/1973159)
  - 45 Niobe-Eyangoh, S. N., Kuaban, C., Sorlin, P., Cunin, P., Thonnon, J., Sola, C., Rastogi, N., Vincent, V. & Gutierrez, M. C. 2003 Genetic biodiversity of *Mycobacterium tuberculosis* complex strains from patients with pulmonary tuberculosis in Cameroon. *J. Clin. Microbiol.* **41**, 2547–2553. (doi:10.1128/JCM.41.6.2547-2553.2003)
  - 46 Woolhouse, M. E., Webster, J. P., Domingo, E., Charlesworth, B. & Levin, B. R. 2002 Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat. Genet.* **32**, 569–577. (doi:10.1038/ng1202-569)
  - 47 Baker, L., Brown, T., Maiden, M. C. & Drobniowski, F. 2004 Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. *Emerg. Infect. Dis.* **10**, 1568–1577.
  - 48 Reed, M. B. *et al.* 2009 Major *Mycobacterium tuberculosis* lineages associate with patient country of origin. *J. Clin. Microbiol.* **47**, 1119–1128. (doi:10.1128/JCM.02142-08)
  - 49 Dawkins, R. & Krebs, J. R. 1979 Arms races between and within species. *Proc. R. Soc. Lond. B* **205**, 489–511. (doi:10.1098/rspb.1979.0081)
  - 50 Rodrigo, T., Cayla, J. A., Garcia de Olalla, P., Galdos-Tanguis, H., Jansa, J. M., Miranda, P. & Brugal, T. 1997 Characteristics of tuberculosis patients who generate secondary cases. *Int. J. Tuberc. Lung Dis.* **1**, 352–357.
  - 51 Banu, S., Honore, N., Saint-Joanis, B., Philpott, D., Prevost, M. C. & Cole, S. T. 2002 Are the PE-PGRS proteins of *Mycobacterium tuberculosis* variable surface antigens? *Mol. Microbiol.* **44**, 9–19. (doi:10.1046/j.1365-2958.2002.02813.x)
  - 52 Coscolla, M. & Gagneux, S. 2010 Does *M. tuberculosis* genomic diversity explain disease diversity? *Drug Discov. Today Dis. Mech.* **7**, e43–e59. (doi:10.1016/j.ddmec.2010.09.004)
  - 53 Nicol, M. P. & Wilkinson, R. J. 2008 The clinical consequences of strain diversity in *Mycobacterium tuberculosis*. *Trans. R. Soc. Trop. Med. Hyg.* **102**, 955–965. (doi:10.1016/j.trstmh.2008.03.025)

- 54 Blaser, M. J. & Kirschner, D. 2007 The equilibria that allow bacterial persistence in human hosts. *Nature* **449**, 843–849. (doi:10.1038/nature06198)
- 55 Kerr, B., Neuhauser, C., Bohannan, B. J. & Dean, A. M. 2006 Local migration promotes competitive restraint in a host–pathogen ‘tragedy of the commons’. *Nature* **442**, 75–78. (doi:10.1038/nature04864)
- 56 Russell, D. G. 2001 *Mycobacterium tuberculosis*: here today, and here tomorrow. *Nat. Rev. Mol. Cell. Biol.* **2**, 569–577. (doi:10.1038/35085034)
- 57 Portevin, D., Gagneux, S., Comas, I. & Young, D. 2011 Human macrophage responses to clinical isolates from the *Mycobacterium tuberculosis* complex discriminate between ancient and modern lineages. *PLoS Pathog.* **7**, e1001307. (doi:10.1371/journal.ppat.1001307)
- 58 Reed, M. B., Domenech, P., Manca, C., Su, H., Barczak, A. K., Kreiswirth, B. N., Kaplan, G. & Barry, C. E. 2004 A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature* **431**, 84–87. (doi:10.1038/nature02837)
- 59 Tsenova, L., Ellison, E., Harbacheuski, R., Moreira, A. L., Kurepina, N., Reed, M. B., Mathema, B., Barry III, C. E. & Kaplan, G. 2005 Virulence of selected *Mycobacterium tuberculosis* clinical isolates in the rabbit model of meningitis is dependent on phenolic glycolipid produced by the bacilli. *J. Infect. Dis.* **192**, 98–106. (doi:10.1086/430614)
- 60 Mitchison, D. A., Wallace, J. G., Bhatia, A. L., Selkon, J. B., Subbaiah, T. V. & Lancaster, M. C. 1960 A comparison of the virulence in guinea-pigs of South Indian and British tubercle bacilli. *Tubercle* **41**, 1–22. (doi:10.1016/S0041-3879(60)80019-0)
- 61 de Jong, B. C. *et al.* 2008 Progression to active tuberculosis, but not transmission, varies by *Mycobacterium tuberculosis* lineage in The Gambia. *J. Infect. Dis.* **198**, 1037–1043. (doi:10.1086/591504)
- 62 Comstock, G. W. 1978 Tuberculosis in twins: a re-analysis of the Prophit survey. *Am. Rev. Respir. Dis.* **117**, 621–624.
- 63 Casanova, J. L. & Abel, L. 2002 Genetic dissection of immunity to mycobacteria: the human model. *Annu. Rev. Immunol.* **20**, 581–620. (doi:10.1146/annurev.immunol.20.081501.125851)
- 64 Lipsitch, M. & Sousa, A. O. 2002 Historical intensity of natural selection for resistance to tuberculosis. *Genetics* **161**, 1599–1607.
- 65 Young, D. B., Perkins, M. D., Duncan, K. & Barry, C. E. 2008 Confronting the scientific obstacles to global control of tuberculosis. *J. Clin. Invest.* **118**, 1255–1265. (doi:10.1172/JCI34614)
- 66 Caws, M. *et al.* 2008 The influence of host and bacterial genotype on the development of disseminated disease with *Mycobacterium tuberculosis*. *PLoS Pathog.* **4**, e1000034. (doi:10.1371/journal.ppat.1000034)
- 67 Intemann, C. D. *et al.* 2009 Autophagy gene variant IRGM -261T contributes to protection from tuberculosis caused by *Mycobacterium tuberculosis* but not by *M. africanum* strains. *PLoS Pathog.* **5**, e1000577. (doi:10.1371/journal.ppat.1000577)
- 68 Herb, F. *et al.* 2008 ALOX5 variants associated with susceptibility to human pulmonary tuberculosis. *Hum. Mol. Genet.* **17**, 1052–1060. (doi:10.1093/hmg/ddm378)
- 69 van Crevel, R. *et al.* 2009 Infection with *Mycobacterium tuberculosis* Beijing Genotype strains is associated with polymorphisms in SLC11A1/NRAMP1 in Indonesian patients with tuberculosis. *J. Infect. Dis.* **200**, 1671–1674. (doi:10.1086/648477)
- 70 Straub, C. S., Ives, A. R. & Gratton, C. 2011 Evidence for a trade-off between host-range breadth and host-use efficiency in aphid parasitoids. *Am. Nat.* **177**, 389–395. (doi:10.1086/658177)
- 71 Cowley, D., Govender, D., February, B., Wolfe, M., Steyn, L., Evans, J., Wilkinson, R. J. & Nicol, M. P. 2008 Recent and rapid emergence of W-Beijing strains of *Mycobacterium tuberculosis* in Cape Town, South Africa. *Clin. Infect. Dis.* **47**, 1252–1259. (doi:10.1086/592575)
- 72 WHO STPOP TB 2006 *The global plan to stop TB 2006–2015*. Geneva, Switzerland: World Health Organization.
- 73 Comas, I. & Gagneux, S. 2009 The past and future of tuberculosis research. *PLoS Pathog.* **5**, e1000600. (doi:10.1371/journal.ppat.1000600)