

**KELLERSBERGER MEMORIAL LECTURE****EVOLUTION OF *MYCOBACTERIUM TUBERCULOSIS* AND IMPLICATIONS FOR VACCINE DEVELOPMENT**Sebastien Gagneux, PhD<sup>1</sup>**ABSTRACT**

*Tuberculosis (TB) is a growing public health threat, particularly in the face of the global epidemics of multidrug resistance. Given the limited efficacy of the current TB vaccine and the recent clinical failure of the most advanced new TB vaccine candidate, novel concepts for vaccine design should be explored. Most T cell antigens in the human-adapted Mycobacterium tuberculosis complex (MTBC) are evolutionarily conserved and under strong purifying selection, indicating that host immune responses targeting these antigens might not be protective. By contrast, a few highly variable T cell epitopes have recently been discovered, which could serve as alternative vaccine antigens. Moreover, there is increasing evidence that the human-adapted MTBC has been co-evolving with the human host for a long time. Hence, studying the interaction between bacterial and human genetic diversity might help identify additional targets that could be exploited for TB vaccine development.*

**INTRODUCTION**

Tuberculosis (TB) remains a global public health emergency. According to the World Health Organization, there were an estimated 1.5 million deaths and 9.6 million new cases of TB in 2014 (1). The global TB epidemics are worsening due to increases in antibiotic resistance in many parts of the world (2). TB is caused by a group of genetically closely related bacteria known as the *Mycobacterium tuberculosis* complex (MTBC). The MTBC can be divided into seven human-adapted phylogenetic lineages (known as Lineage 1 to 7) and two lineages adapted to various wild and domestic animal species (3). In humans, the outcome of TB infection and disease is highly variable, ranging from latent infection with no symptoms to classical pulmonary TB, extra pulmonary disease and TB meningitis (4).

In the past, this variation in outcome has primarily been attributed to host and environmental factors (5). However, there is growing evidence that bacterial variation also plays a role. Indeed, many experimental studies have shown that MTBC clinical strains differ in virulence and immunogenicity (reviewed in Ref. 6), and there is increasing evidence that some of these laboratory phenotypes are also reflected in epidemiological settings (reviewed in Ref. 3). For example, the Beijing family of strains, which belongs to the evolutionarily “modern” MTBC Lineage 2 (also known as the East Asian Lineage), has been associated with hyper-virulence and a delayed pro-inflammatory immune response in infection models (7, 8). In clinical settings, Lineage 2 has been associated with faster progression

to active disease (9), increased transmissibility (10) and recent emergence (11, 12). Lineage 2 has also repeatedly been associated with antibiotic resistance (reviewed in Ref. 13), which might be a consequence of its higher transmission potential (14) and/or its elevated mutation rate compared to other MTBC lineages (15). By contrast, Lineage 6 (also known as *Mycobacterium africanum* West Africa II) and other members of the evolutionarily “ancient” MTBC lineages have been associated with lower virulence (16,17) and early pro-inflammatory responses (8). Lineage 6 also exhibits a slower progression to active disease (9), and is less likely to be drug-resistant (18). Because a large proportion of the MTBC genome is dedicated to lipid metabolism (19), and lipids are important cell wall components involved in virulence and immunity to TB (20), it is likely that some of these lineage-specific phenotypic differences are mediated by variation in lipid content. Indeed, some of the strain-specific differences in virulence and pro-inflammatory immune responses in TB have been linked to the presence or absence of particular bacterial lipids (7,21). For example, quantitative and qualitative lineage-specific differences have been reported in mycolic acids (22), which are the most abundant cell wall lipids in the MTBC.

One of the biggest problems for global TB control is the lack of a universally effective vaccine (23). The only currently available vaccine against TB, Bacille-Calmette-Guerin (BCG) is an attenuated derivative of *Mycobacterium bovis* (24). BCG protects infants well against TB meningitis and military TB, but it is essentially ineffective against regular pulmonary disease in adults, which is the transmissible form of the disease (25). Efforts to develop a better TB vaccine have been ongoing on for many decades, but our fundamental

<sup>1</sup>Department of Medical Parasitology and Infectious Disease, Swiss Tropical and Public Health Institute / University of Basel, Switzerland, e-mail: sebastien.gagneux@unibas.ch

ignorance with respect to what constitutes a protective immune response to TB has been hampering progress (4). This lack of basic understanding of some of the key biological aspects of the host-pathogen interaction in TB is also illustrated by the recent clinical failure of MVA85A, the most advanced novel TB vaccine candidate (26). In part, this failure has to do with the fact that until now, both basic research and product development in TB have largely ignored the genotypic and phenotypic diversity in natural populations of the MTBC (27), as well as the evolutionary forces shaping this diversity (28). Because the human-adapted MTBC is an obligate human pathogen with no animal or environmental reservoir (29), the tight link between host and pathogen in human TB will promote co-evolution between these two interacting partners (30). This co-evolution needs to be taken into account when studying the biology and epidemiology of TB, and developing new tools and strategies to control the disease (31).

***Evidence for host-pathogen co-evolution in human TB:***

There is mounting evidence that the human-adapted MTBC lineages have co-evolved with humans for thousands of years (32). Similar to humans, the genetic diversity of MTBC is largest in Africa and decreases with increasing distance from Africa (33). Moreover, there is strong phylogeographical evidence that the MTBC originated in sub-Saharan Africa (34). This is also supported by the observation that *Mycobacterium canettii* and the other so-called “smooth mycobacteria” that share a common ancestor with the MTBC, have only been isolated in the Horn of Africa (35). Similarly, several human-adapted lineages of the MTBC are geographically restricted to parts of Africa. Specifically, Lineages 5 and 6 (also known as *M. africanum* West Africa I and II) are mainly found in West Africa (36) or in patient of West African descent (37), whereas Lineage 7 almost exclusively occurs in Ethiopia and recent Ethiopian immigrants (38). Hence, these MTBC lineages may be particularly well adapted to infect and transmit among their respective host populations (39). In support of this notion, the sympatric host-pathogen association between the different MTBC lineages and their respective human host populations are maintained in low-incidence cosmopolitan settings (40-43), even during active TB transmission abroad (43, 44).

Furthermore, allopatric host-pathogen combinations in these settings are strongly associated with HIV co-infection, an association that is further enhanced with increasing immune-suppression of the host (43). Even though social factors are likely contributing to these phenomena, the stable associations between different host and pathogen populations, and the fact that these associations are disrupted by HIV-induced immune-deficiency are consistent with host-pathogen co-evolution in human TB (30). Little work has been done so far comparing the genetic population structures of the MTBC and humans. Comas et al. found qualitative and quantitative similarities between the main MTBC lineages and the three main human mito-

chondrial macro-haplogroups (34). Pepperell et al. found no association between the MTBC lineages and human Y chromosomes (45), but when focusing on the main Y chromosome haplogroups, possible associations became evident (29). Finally, additional parallels between the MTBC and human population structures emerge when considering the frequency spectra of single nucleotide polymorphisms. Both humans and the human-adapted MTBC feature a high proportion of so-called singleton mutations, i.e. single nucleotide mutations that occur only in one individual (45-47). In humans, this phenomenon has been linked to strong population increases in the past and a reduction in selective constraint, leading to an accumulation of slightly deleterious mutations, which has important implications for human medical genetics (48).

In the human-adapted MTBC, signals of population expansion have also been detected and connected to the Agricultural and Industrial Revolutions (34, 49-51). Given the strong link between the MTBC and its human host (28), it is perhaps not surprising that bacterial population expansions would follow naturally from similar expansions in humans (29). A reduction in selective constraint linked to serial transmission bottlenecks has also been invoked for the MTBC to explain the high proportion of non synonymous mutations (46), about 50% of which have been predicted to impact protein function (52). Deciphering if and how these predicted changes in protein function are linked to some of the lineage-specific differences in experimental and clinical phenotypes outlined in the first section of this review will require considerably more work. Ideally, this bacterial variation should be studied in the context of the corresponding host variation (31).

It is well known that humans differ in susceptibility to TB (48). However, even though several human genetic loci have been implicated in TB susceptibility (53), replicating these findings across different human populations has proven difficult (54). Part of the reason for this poor reproducibility is that most studies in human genetics of TB have not taken into account pathogen diversity (54), which is particularly problematic in infectious diseases in which the host and the pathogen have a long co-evolutionary history (55). Yet, several studies have reported interactions between particular human polymorphisms and MTBC genotypes (56-61). This supports the notion that to be able to understand clinical phenotypes and predict the outcome of TB infection and disease, the genetic background of both the host and the pathogen need to be considered in conjunction (31). Similarly, host and pathogen diversity need to be considered for vaccine development.

***Consequences of host-pathogen co-evolution for TB vaccine development:***

Assuming the notion of host-pathogen co-evolution in human TB is valid, this has important implications for our understanding of the biology and epidemiology of the disease. Moreover, this will influence the way we think about developing new tools to combat TB, in particular vaccines. Tradi-

tionally, TB vaccine development has focused on dominant antigens that elicit strong T cell responses, in particular interferon- $\gamma$  (INF- $\gamma$ , 62). From other infectious diseases we know that pathogens vary their antigens to evade host immunity (63). This kind of evolutionary arms race is the hallmark of many chronic disease systems in which host and pathogen species interact over extended periods of time (64). Does this type of arms race also occur in human TB? Unexpectedly, we found that in the MTBC the known T cell epitopes were evolutionary highly conserved. First we showed this in 2010 using 491 experimentally confirmed T cell epitopes known at the time (65), which included many candidate vaccine antigens, and 21 MTBC genomes covering six of the seven main human-adapted MTBC lineages (66).

In 2015, we repeated this analysis using 1,226 experimentally confirmed T cell epitopes and 216 MTBC genomes from global sources covering all seven MTBC lineages (67). In both studies we found that the large majority (>70%) of T cell epitopes showed no amino acid substitution among the MTBC strains analyzed (66, 67). Moreover, the T cell epitopes were significantly more evolutionary conserved than essential genes. These results were consistent with a previous study, in which we analyzed 27 *pe\_pgrs* genes in 94 phylogenetically diverse MTBC clinical isolates (68). We found that although some *pe\_pgrs* genes were highly polymorphic, their sequence variation was largely confined to the C-terminal PGRS domain, whereas their T cell epitopes were concentrated in the conserved N-terminal PE domain (68).

Hence, despite being under constant host immune pressure, T cell epitopes in the MTBC are under strong purifying selection to remain unchanged, as opposed to experiencing diversifying selection for antigen variation. Based on these observations, we hypothesized that because the MTBC is an obligate human pathogen in which virulence contributes to pathogen transmission (28), the T cell responses elicited by most dominant T cell antigens will be beneficial to the pathogen rather than protecting the host (67). In support of this notion, lung cavitation is a hallmark of TB and a consequence of immune-pathological processes that strongly increase the transmission potential of TB patients (69).

Until now, efforts in TB vaccine development have been focusing on conserved antigens. The results summarized above suggest an alternative strategy, namely to focus on variable antigens, as these might reflect genomic loci that are evolving to evade the host immune responses deleterious to the pathogen. Stimulating host responses to such bacterial targets might therefore promote protective immunity. In an attempt to identify variable T cell epitopes, we recently screened 216 whole genome sequences from MTBC clinical strains covering all seven MTBC lineages (67).

We then selected seven genes that showed a high genetic diversity and a high proportion of non-synonymous variation, and computationally predicted CD4 and CD8 T cell epitopes in these seven candidate antigens using HLA class I and HLA class II alleles that are prevalent in diverse human populations (67). Finally, we synthesized 14 corresponding candidate peptides and used those to stimulate whole blood from 82 HIV-negative TB patients. We found that each of our candidate epitopes was immunogenic in at least one subject as assessed by INF- $\gamma$  release assay. Moreover, we evaluated the impact of naturally occurring amino acid substitutions in these candidate peptides with respect to their capacity to elicit an immune response in the same individual. We found that 72% of the subjects showed a differential response to the ancestral versus the alternative epitope variant. Taken together, these results show that even though the large majority of the known T cell epitopes in the MTBC are conserved, some variable epitopes exist and this variation can influence host immune recognition.

Whether this natural variation in MTBC results from immune escape and whether the immune responses to such variable epitopes could be exploited for TB vaccine development should be explored further. In addition, a “genome-to-genome” approach could be a new way forward for identifying novel vaccine antigens with a high probability of eliciting protective immune responses in TB (31). In the context of the co-evolution between humans and the MTBC, interrogating both the host and the pathogen genome in conjunction might reveal particular loci that are in conflict with each other (55). Focusing TB vaccine development efforts on such conflicting loci could be particularly rewarding.

In conclusion, a novel, more effective vaccine is urgently needed to control TB globally, particularly in view of the growing epidemics of multidrug resistance. Future efforts to develop new vaccine candidates should consider the natural variation of MTBC populations. Given the mounting evidence that this variation partially reflects host-pathogen co-evolution, studying the interaction between human and MTBC diversity might point to novel targets and strategies for the development and deployment of new TB vaccines.

## ACKNOWLEDGEMENTS

I thank Joel Ernst, Richard Copin and all the members of my group for the stimulating discussions. Work in my group is supported by the Swiss National Science Foundation (grant PP00P3\_150750), the European Research Council (309540-EVODRTB), and SystemsX.ch.

## REFERENCES

1. World Health Organization. Global tuberculosis control - surveillance, planning, financing. Geneva, Switzerland: WHO, 2015.
2. Trauner A, Borrell S, Reither K, Gagneux S. Evolution of drug resistance in tuberculosis: recent progress and implications for diagnosis and therapy. *Drugs*. 2014; 74(10):1063-72.
3. Coscolla M, Gagneux S. Consequences of genomic diversity in *Mycobacterium tuberculosis*. *Semin Immunol*. 2014; 26(6):431-44.
4. Ernst JD. The immunological life cycle of tuberculosis. *Nat Rev Immunol*. 2012; 12(8):581-91.
5. Comas I, Gagneux S. The past and future of tuberculosis research. *PLoS Pathog*. 2009; 5(10):e1000600.
6. Coscolla M, Gagneux S. Does *M. tuberculosis* genomic diversity explain disease diversity? *Drug Discov Today Dis Mech*. 2010; 7(1):e43-e59.
7. Reed MB, Domenech P, Manca C, Su H, Barczak AK, Kreiswirth BN, et al. A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature*. 2004; 431(7004):84-7.
8. Portevin D, Gagneux S, Comas I, Young D. Human macrophage responses to clinical isolates from the *Mycobacterium tuberculosis* complex discriminate between ancient and modern lineages. *PLoS Pathog*. 2011; 7(3):e1001307.
9. de Jong BC, Hill PC, Aiken A, Awine T, Antonio M, Adetifa IM, et al. Progression to active tuberculosis, but not transmission, varies by *Mycobacterium tuberculosis* lineage in The Gambia. *J Infect Dis*. 2008; 198(7):1037-43.
10. Yang C, Luo T, Sun G, Qiao K, Deriemer K, Mei J, et al. The *Mycobacterium tuberculosis* Beijing strains favor transmission but not drug resistance in China. *Clin Infect Dis*. 2012; 55(9):1179-87.
11. Cowley D, Govender D, February B, Wolfe M, Steyn L, Evans J, et al. Recent and rapid emergence of W-Beijing strains of *Mycobacterium tuberculosis* in Cape Town, South Africa. *Clin Infect Dis*. 2008; 47(10):1252-9.
12. van der Spuy GD, Kremer K, Ndabambi SL, Beyers N, Dunbar R, Marais BJ, et al. Changing *Mycobacterium tuberculosis* population highlights clade-specific pathogenic characteristics. *Tuberculosis (Edinb)*. 2008; 89(2):120-5.
13. Borrell S, Gagneux S. Infectiousness, reproductive fitness and evolution of drug-resistant *Mycobacterium tuberculosis*. *Int J Tuberc Lung Dis*. 2009; 13(12):1456-66.
14. Muller B, Borrell S, Rose G, Gagneux S. The heterogeneous evolution of multidrug-resistant *Mycobacterium tuberculosis*. *Trends Genet*. 2013; 29(3):160-9.
15. Ford CB, Shah RR, Maeda MK, Gagneux S, Murray MB, Cohen T, et al. *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nat Genet*. 2013; 45(7):784-90.
16. Bold TD, Davis DC, Penberthy KK, Cox LM, Ernst JD, de Jong BC. Impaired Fitness of *Mycobacterium africanum* despite secretion of ESAT-6. *J Infect Dis*. 2012; 205(6):984-90.
17. Mitchison DA, Bhatia AL, Radhakrishna S, Selkon JB, Subbaiah TV, Wallace JG. The virulence in the guinea-pig of tubercle bacilli isolated before treatment from South Indian patients with pulmonary tuberculosis. I. Homogeneity of the investigation and a critique of the virulence test. *Bull World Health Organ*. 1961; 25:285-312.
18. Yeboah-Manu D, Asante-Poku A, Bodmer T, Stucki D, Koram K, Bonsu F, et al. Genotypic diversity and drug susceptibility patterns among *M. tuberculosis* complex isolates from South-Western Ghana. *PLoS ONE*. 2011; 6(7):e21906.
19. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*. 1998; 393(6685):537-44.
20. Jackson M. The mycobacterial cell envelope-lipids. *Cold Spring Harb Perspect Med*. 2014; 4(10).
21. Manca C, Tsenova L, Barry CE, 3rd, Bergtold A, Freeman S, Haslett PA, et al. *Mycobacterium tuberculosis* CDC1551 induces a more vigorous host response in vivo and in vitro, but is not more virulent than other clinical isolates. *J Immunol*. 1999; 162(11):6740-6.
22. Portevin D, Sukumar S, Coscolla M, Shui G, Li B, Guan XL, et al. Lipidomics and genomics of *Mycobacterium tuberculosis* reveal lineage-specific trends in mycolic acid biosynthesis. *Microbiologyopen*. 2014; 3(6):823-35.
23. Evans TG, Schrager L, Thole J. Status of vaccine research and development of vaccines for tuberculosis. *Vaccine*. 2016 epublished.
24. Brosch R, Gordon SV, Garnier T, Eiglmeier K, Frigui W, Valenti P, et al. Genome plasticity of BCG and impact on vaccine efficacy. *Proc Natl Acad Sci U S A*. 2007; 104(13):5596-601.
25. Andersen P, Doherty TM. Opinion: The success and failure of BCG - implications for a novel tuberculosis vaccine. *Nat Rev Microbiol*. 2005; 3(8):656-62.
26. Tameris MD, Hatherill M, Landry BS, Scriba TJ, Snowden MA, Lockhart S, et al. Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomized, placebo-controlled phase 2b trial. *Lancet*. 2013; 381(9871):1021-8.
27. Gagneux S, Small PM. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect Dis*. 2007; 7(5):328-37.

28. Brites D, Gagneux S. Old and new selective pressures on *Mycobacterium tuberculosis*. *Infect Genet Evol.* 2012; 12(4):678-85.
29. Brites D, Gagneux S. Co-evolution of *Mycobacterium tuberculosis* and *Homo sapiens*. *Immunol Rev.* 2015; 264(1):6-24.
30. Woolhouse ME, Webster JP, Domingo E, Charlesworth B, Levin BR. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat Genet.* 2002; 32(4):569-77.
31. Comas I, Gagneux S. A role for systems epidemiology in tuberculosis research. *Trends Microbiol.* 2011; 19(10):492-500.
32. Gagneux S. Host-pathogen coevolution in human tuberculosis. *Philos Trans R Soc Lond B Biol Sci.* 2012; 367(1590):850-9.
33. Comas I, Hailu E, Kiros T, Bekele S, Mekonnen W, Gumi B, et al. Population genomics of *Mycobacterium tuberculosis* in Ethiopia contradicts the Virgin Soil Hypothesis for human tuberculosis in Sub-Saharan Africa. *Curr Biol.* 2015; 25(24):3260-6.
34. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Genet.* 2013; 45(10):1176-82.
35. Gutierrez C, Brisse S, Brosch R, Fabre M, Omais B, Marmiesse M, et al. Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. *PLoS Pathogens.* 2005; 1:1-7.
36. de Jong BC, Antonio M, Gagneux S. *Mycobacterium africanum*--review of an important cause of human tuberculosis in West Africa. *PLoS Negl Trop Dis.* 2010; 4(9):e744.
37. Sharma A, Bloss E, Heilig CM, Click ES. Tuberculosis caused by *Mycobacterium africanum*, United States, 2004-2013. *Emerg Infect Dis.* 2016; 22(3):396-403.
38. Firdessa R, Berg S, Hailu E, Schelling E, Gumi B, Erenso G, et al. Mycobacterial lineages causing pulmonary and extrapulmonary tuberculosis, Ethiopia. *Emerg Infect Dis.* 2013; 19:460-3.
39. Asante-Poku A, Yeboah-Manu D, Otchere ID, Aboagye SY, Stucki D, Hattendorf J, et al. *Mycobacterium africanum* is associated with patient ethnicity in Ghana. *PLoS Negl Trop Dis.* 2015; 9(1):e3370.
40. Baker L, Brown T, Maiden MC, Drobniewski F. Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. *Emerg Infect Dis.* 2004; 10(9):1568-77.
41. Reed MB, Pichler VK, McIntosh F, Mattia A, Fallow A, Masala S, et al. Major *Mycobacterium tuberculosis* lineages associate with patient country of origin. *J Clin Microbiol.* 2009; 47(4):1119-28.
42. Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small PM. Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. *Proc Natl Acad Sci U S A.* 2004; 101(14):4871-6.
43. Fenner L, Egger M, Bodmer T, Furrer H, Ballif M, Battegay M, et al. HIV infection disrupts the sympatric host-pathogen relationship in human tuberculosis. *PLoS Genetics.* 2013; 9(3):e1003318.
44. Gagneux S, Deriemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, et al. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A.* 2006; 103:2869-73.
45. Pepperell CS, Casto AM, Kitchen A, Granka JM, Cornejo OE, Holmes EC, et al. The role of selection in shaping diversity of natural *M. tuberculosis* populations. *PLoS Pathog.* 2013; 9(8):e1003543.
46. Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, Homolka S, et al. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biology.* 2008; 6(12):e311.
47. Keinan A, Clark AG. Recent explosive human population growth has resulted in an excess of rare genetic variants. *Science.* 2012;336(6082): 740-3.
48. Abel L, El-Baghdadi J, Bousfiha AA, Casanova JL, Schurr E. Human genetics of tuberculosis: a long and winding road. *Philos Trans R Soc Lond B Biol Sci.* 2014; 369(1645):20130428.
49. Bos KI, Harkins KM, Herbig A, Coscolla M, Weber N, Comas I, et al. Pre-Columbian mycobacterial genomes reveal seals as a source of New World human tuberculosis. *Nature.* 2014; 514(7523):494-7.
50. Luo T, Comas I, Luo D, Lu B, Wu J, Wei L, et al. Southern East Asian origin and coexpansion of *Mycobacterium tuberculosis* Beijing family with Han Chinese. *Proc Natl Acad Sci U S A.* 2015; 112(26):8136-41.
51. Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, et al. Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nat Genet.* 2015; 47(3):242-9.
52. Rose G, Cortes T, Comas I, Coscolla M, Gagneux S, Young DB. Mapping of genotype-phenotype diversity among clinical isolates of *Mycobacterium tuberculosis* by sequence-based transcriptional profiling. *Genome Biology and Evolution.* 2013; 5(10):1849-62.
53. Moller M, de Wit E, Hoal EG. Past, present and future directions in human genetic susceptibility to tuberculosis. *FEMS Immunol Med Microbiol.* 2010; 58(1):3-26.
54. Karlsson EK, Kwiatkowski DP, Sabeti PC. Natural selection and infectious disease in human populations. *Nat Rev Genet.* 2014; 15(6):379-93.
55. Kodaman N, Sobota RS, Mera R, Schneider BG, Williams SM. Disrupted human-pathogen co-evolution: a model for disease. *Front Genet.* 2014; 5:290.
56. Caws M, Thwaites G, Dunstan S, Hawn TR, Thi Ngoc Lan N, Thuong NT, et al. The influence of host and bacterial genotype on the development of disseminated disease with *Mycobacterium tuberculosis*. *PLoS Pathog.* 2008; 4(3):e1000034.

57. Salie M, van der Merwe L, Moller M, Daya M, van der Spuy GD, van Helden PD, et al. Associations between human leukocyte antigen class I variants and the *Mycobacterium tuberculosis* subtypes causing disease. *J Infect Dis.* 2014; 209(2):216-23.
58. Thye T, Niemann S, Walter K, Homolka S, Intemann CD, Chinbuah MA, et al. Variant G57E of Mannose Binding Lectin associated with protection against tuberculosis caused by *Mycobacterium africanum* but not by *M. tuberculosis*. *PLoS ONE.* 2011; 6(6):e20908.
59. Intemann CD, Thye T, Niemann S, Browne EN, Amanua Chinbuah M, Enimil A, et al. Autophagy gene variant IRGM -261T contributes to protection from tuberculosis caused by *Mycobacterium tuberculosis* but not by *M. africanum* strains. *PLoS Pathog.* 2009; 5(9):e1000577.
60. Herb F, Thye T, Niemann S, Browne EN, Chinbuah MA, Gyapong J, et al. ALOX5 variants associated with susceptibility to human pulmonary tuberculosis. *Human molecular genetics.* 2008; 17(7):1052-60.
61. van Crevel R, Parwati I, Sahiratmadja E, Marzuki S, Ottenhoff TH, Netea MG, et al. Infection with *Mycobacterium tuberculosis* Beijing genotype strains is associated with polymorphisms in SLC11A1/NRAMP1 in Indonesian patients with tuberculosis. *J Infect Dis.* 2009; 200(11):1671-4.
62. Karp CL, Wilson CB, Stuart LM. Tuberculosis vaccines: barriers and prospects on the quest for a transformative tool. *Immunol Rev.* 2015; 264(1):363-81.
63. Deitsch KW, Lukehart SA, Stringer JR. Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nat Rev Microbiol.* 2009; 7(7):493-503.
64. Dawkins R, Krebs JR. Arms races between and within species. *Proc R Soc Lond B Biol Sci.* 1979; 205(1161):489-511.
65. Ernst JD, Lewinson DM, Behar S, Blythe M, Schlesinger LS, Kornfeld H, et al. Meeting Report: NIH Workshop on the Tuberculosis Immune Epitope Database. *Tuberculosis (Edinb).* 2008; 88(4):366-70.
66. Comas I, Chakravarti J, Small PM, Galagan J, Niemann S, Kremer K, et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat Genet.* 2010; 42(6):498-503.
67. Coscolla M, Copin R, Sutherland J, Gehre F, de Jong B, Owolabi O, et al. *M. tuberculosis* T cell epitope analysis reveals paucity of antigenic variation and identifies rare variable TB antigens. *Cell Host & Microbe.* 2015; 18(5):538-48.
68. Copin R, Coscolla M, Seiffert SN, Bothamley G, Sutherland J, Mbayo G, et al. Sequence diversity in the pe\_pgrs genes of *Mycobacterium tuberculosis* is independent of human T cell recognition. *MBio.* 2014; 5(1):e00960-13.
69. Reichler MR, Reves R, Bur S, Thompson V, Mangura BT, Ford J, et al. Evaluation of investigations conducted to detect and prevent transmission of tuberculosis. *JAMA.* 2002; 287(8):991-5.