



Research paper

Large genomics datasets shed light on the evolution of the *Mycobacterium tuberculosis* complexÁlvaro Chiner-Oms^a, Iñaki Comas^{b,c,*}^a Unidad Mixta “Infección y Salud Pública” FISABIO-CSISP/Universidad de Valencia, Instituto de Biología Integrativa de Sistemas-I2SysBio, Valencia, Spain^b CIBER en Epidemiología y Salud Pública, Valencia, Spain^c Instituto de Biomedicina de Valencia, IBV-CSIC, Valencia, Spain

ARTICLE INFO

Keywords:

Mycobacterium tuberculosis complex
Evolution
Positive selection
Genomics

ABSTRACT

Two strains of *Mycobacterium tuberculosis* complex can be separated as much as 2500 single nucleotide differences (Coscolla and Gagneux, 2014). In that limited amount of diversity, we find an astonishing range of clinical, epidemiological and biological phenotypes. The most striking is the strong host preferences depending on the infecting strain while more subtle differences can be found looking at different human tuberculosis isolates. Those subtle differences are the most difficult to spot given that analysis methods for so little diversity are limited and phenotypes like virulence are difficult to define and measure. Recent genomics advances allow to study the pathogen diversity at a resolution not available before from comparative species level, to global diversity to transmission in local settings. Here, we will review some of these recent results to highlight how population genomics approaches can aid not only to understand how MTBC evolved but also to identify relevant biomedical targets.

1. The *Mycobacterium* genus

The *Mycobacterium* genus comprises a relatively wide number of species (~170) (Fedrizzi et al., 2017). Most of them are free living bacteria (that is, bacteria with a life cycle mostly dissociated from a host) found in a wide range of ecological environments. A high number of these species are able to cause disease in humans, supposing an important threat to human health (i.e. *M. tuberculosis* complex, *M. leprae*, *M. ulcerans* among others). The increasing public availability of mycobacterial genomes allow us to draw a precise map of the phylogenetic relationships of this group, avoiding the limitations of classifications based on partial gene sequencing and/or phenotypes (Fedrizzi et al., 2017). Average Nucleotide Identities (ANI) analysis have shown that some members previously considered as independent taxa could represent in reality different variants of the same species (Tortoli et al., 2017) (understanding species as individuals having a 95% genomic ANI or more (Goris et al., 2007)). Importantly, the study by Fedrizzi et al. (Fedrizzi et al., 2017) ameliorates the problem of overrepresentation of pathogenic mycobacteria species in genomic databases. This information will be crucial to define specific genetic markers for identification and diagnostics of mycobacteriosis (Braet et al., 2018). Furthermore, the genome information has already allowed to amend species classifications (Tortoli et al., 2017), redefine mycobacterial complexes

(Tortoli et al., 2017) and open the debate about the status of the *Mycobacterium* genus itself (Gupta et al., 2018; Turenne, 2019).

Core genome analysis placed the *Mycobacterium tuberculosis* complex (MTBC) phylogenetically close to other slow-grower mycobacteria (that is, bacteria that take more than 7 days to form colonies in culture (Runyon, 1959)) that are associated with human diseases such as *M. leprae* and *M. ulcerans*. *Mycobacterium tuberculosis* complex species (*M. tuberculosis*, *M. africanum* and the animal-adapted strains) together with *M. leprae* are obligate intracellular pathogens with the probable exception of *Mycobacterium canettii* (Koeck et al., 2011). The infection with some opportunistic mycobacterial pathogens such as *M. kansasii*, *M. marinum*, *M. simiae* or the *M. avium* complex result in different symptoms. Interestingly, *M. avium* infection symptoms resembles more to those provoked by *M. tuberculosis* complex (van Ingen, 2017) despite being phylogenetically distant in contrast with other species. Although being occasional pathogens, those mycobacterial infections are difficult to deal with (Nahid et al., 2016). The mycobacterial cell wall is highly complex (Alderwick et al., 2015) and allows mycobacteria to survive under multiple stress conditions. As a consequence, the mycobacterium group is naturally resistant to multiple antibiotics due to their cell wall structure (Jankute et al., 2015). In the case of the MTBC strains infecting humans this fact, among some other factors, leads to a medical treatment of 6 months mixing 4 different drugs to overcome the disease.

* Corresponding author at: Instituto de Biomedicina de Valencia, IBV-CSIC, Valencia, Spain.

E-mail address: icomas@ibv.csic.es (I. Comas).

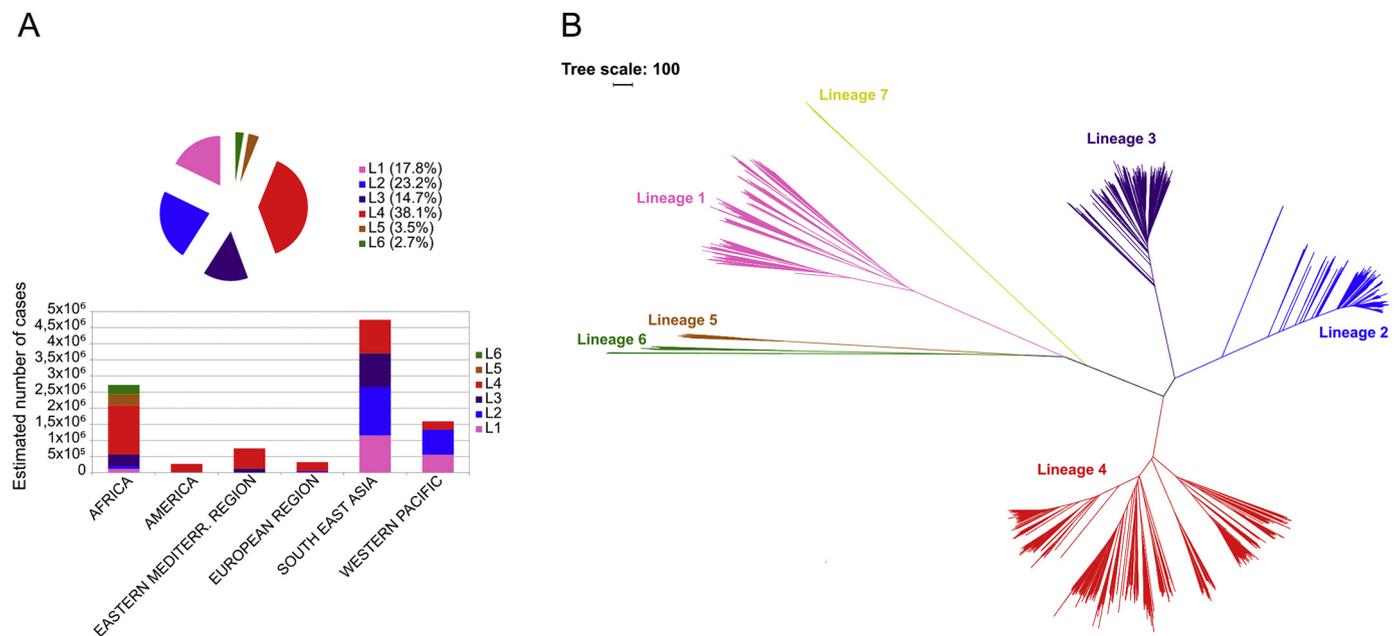


Fig. 1. A) Contribution of the major MTBC lineages to the global tuberculosis burden from Chiner-Oms et al., 2018a. B) MTBC global phylogeny.

2. Early steps in MTBC emergence

Recombination and horizontal gene transfer (HGT) are important processes for genomic plasticity and innovation in both, pathogenic and non-pathogenic bacteria (Sheppard et al., 2018). The role of Horizontal Gene Transfer (HGT) in mycobacterial evolution has been demonstrated in several studies (Gray and Derbyshire, 2018; Hatfull and Jacobs, 2000). Some species that form this group have mosaic genomes, product of multiple recombination events (Saini et al., 2012). In sharp contrast, MTBC members were considered to have a clonal population structure (Smith et al., 2003, 2006). Later studies showed that recombination with *Mycobacterium canettii*, a close group, have played a role during the early steps of the MTBC speciation process, but not after the emergence of the obligate pathogen (Boritsch et al., 2016a; Chiner-Oms et al., 2018b; Mortimer and Pepperell, 2014).

The average nucleotide identity between the *M. canettii* group and *M. tuberculosis* is about 98% (Chiner-Oms et al., 2018b; Supply et al., 2013), close to the 95% threshold used to define independent bacterial species (Goris et al., 2007). So that, all species of the MTBC including *M. canettii* should be considered a unique genomic species. However, in contrast with other members of the clade, *M. canettii* has a low epidemiological incidence, with less than 100 cases since its first isolation in 1969 (Gagneux, 2018). This fact, in addition with other evidences, suggests the likely existence of an environmental reservoir of *M. canettii* (Koeck et al., 2011).

Despite its elevated genetic similarity with other MTBC members, *M. canettii* shows a high recombinant profile in extant strains (Supply et al., 2013). As MTB strains and *M. canettii* belong to different ecological niches, it is reasonable to think that recombination is not taking place between both groups due to the lack physical interaction. However, experimental studies have shown that there is no recombination between *M. canettii* and MTBC strains even when they are grown in the same physical space (Boritsch et al., 2016b).

The role of on-going recombination in extant MTBC strains excluding *M. canettii* was a matter of debate in the past. Some studies have suggested the presence of recombination acting as a present evolutionary force in the MTBC (Liu et al., 2006; Namouchi et al., 2012). However, more recent studies taking into account information from hundreds to thousands of genomes as well as experimental studies have disproved this hypothesis (Boritsch et al., 2016b; Chiner-Oms et al.,

2018b; Comas et al., 2013; Mortimer and Pepperell, 2014). This is an important fact to understand the evolutionary processes that guide the adaptation of the tuberculosis pathogen to different environments and stress conditions (i.e. multiple hosts or antibiotics). It seems that genetic exchange between MTBC species, if exists, plays a minimum role in the evolution of this group. So, the emergence of new characteristics is, almost sure, lead by the appearance of new independent mutations and fixed by means of selection (Brynildsrud et al., 2018; Farhat et al., 2013; Merker et al., 2015) or genetic drift (Hershberg et al., 2008).

When the genome of the model strain H37Rv and the animal strain *M. bovis* were compared, unexpectedly no new gene was found in the animal strain confirming reductive evolution as an ongoing process in MTBC evolution (Garnier et al., 2003). Because of the MTBC clonal structure, some parts of the genome involved in past deletion events were never recovered by the descendants of the ancestral strain. In fact, the history of the BCG vaccine is a clear example of such a process (Behr et al., 1999; Brosch et al., 2007) where cumulative deleted regions have led to different immunogenic potential. Some of these deleted regions (known as RD from Regions of Difference) have been used for strain typing and lineage identification as they are accurate phylogenetic markers with virtually absence of homoplasmy (Comas et al., 2009).

3. Post-emergence diversification of MTBC in lineages and sub-lineages

Early efforts on studying MTBC diversity already notice the different distribution of genotypes across the world (Gagneux and Small, 2007). The advent of genome-based techniques allowed to corroborate early distinction among six main human lineages. Those lineages are not equally distributed across the world. For example the so-called lineage 4 has a global contribution and is likely responsible for roughly 4 million cases yearly ((Chiner-Oms et al., 2018a) Fig. 1A). Other lineages, while not equally dispersed, seem to be under expansion. For example, lineage 2 accounts for ~23% of the global MTBC burden (Fig. 1A), it is mainly distributed across east-Asia and has been associated to drug resistance (Borrell and Gagneux, 2009; Ford et al., 2013), hypervirulence (Reed et al., 2004) and high transmissibility (Holt et al., 2018; Merker et al., 2015). Lineage 2 strains were introduced in South Africa around 100 years ago (van Helden et al., 2002) and seem to be expanding. Likewise, there is increasing evidence on the presence of

lineage 2 strains across the African continent (Rutaiwa et al., 2018). It is important to note that the evidence for this different features is still scarce. In contrast, other lineages have more limited presence across the globe. Lineage 5 and Lineage 6 strains, historically pooled as *Mycobacterium africanum*, are responsible for up to 50% of cases in West Africa but are difficult to find in significant frequencies outside the region (de Jong et al., 2010). The two lineages are not equally distributed across west African countries suggesting that they cannot be considered as a unique entity as a result of different historical events and biological differences. In any case, we cannot make generalizations on the success of certain genotypes over others and what seems to be under expansion in one place is not in another. Take for example lineage 2 strains. Recently it has been described that the sub-lineages L2.2.5 and L2.2.6 are dominating the population structure of *Mycobacterium tuberculosis* in East Africa (Merker et al., 2015; Rutaiwa et al., 2018). On the contrary, the sub-lineage L2.2.10 has clearly succeeded in Central Asia in the face of other sub-lineages. This suggests that epidemiological success is locally dependent, most likely as a result of the interaction with different human genetic backgrounds and different “ecological” systems such as the nature of public health control measures or the socioeconomic status of the population (Comas and Gagneux, 2009).

In cosmopolitan clinical settings, we can still observe a strong association between pathogen lineage and country of origin of the patient. This was first noticed in San Francisco, where the association between both variables is also reflected in preferred transmission when there is a match between the country of origin and the geographic origin of the lineage (Gagneux et al., 2006). This association is lost when HIV is present a result that has been replicated in Switzerland (Fenner et al., 2013). Data from San Francisco thus suggests that beyond social contact there are biological factors associated to genotype-genotype interaction between the pathogen and the patient genetic background. While preferential transmission has not been replicated in other parts of the world, mostly because of the lack of population-based genotyping data over decades, there is still a clear association between lineages and patient country of origin (Reed et al., 2009).

A mirror case to the lineages global/local distribution is seen when we zoom into the most common lineage worldwide, lineage 4 (Stucki et al., 2016). The lineage 4 sublineages are also either globally or locally distributed and this has led to the hypothesis that some of them are generalist, able to interact and thrive in a wide number of human genetic backgrounds, or specialists infecting mainly particular human populations. The different geographical distribution maybe explained by historical contingency with some sublineages emerging in regions that later led the colonization and globalization era (Brynildsrud et al., 2018; Stucki et al., 2016). By contrast the data from San Francisco suggests that there is also some biological features that in fact promotes the association of lineages (or sublineages) to specific human populations. In an attempt to identify some of the biological factors Stucki et al. analyzed the patterns of diversity in the epitope regions of thousands of genome sequences from the different lineage 4 sublineages. For most of them the epitope regions were more conserved than the non-epitope regions. The result confirms that epitope hyperconservation is the rule as observed already in early studies with global populations (Comas et al., 2010; Coscolla et al., 2015). However, this general rule was reverted for the L4.3 (also known as LAM) sublineage, arguably the most successful lineage 4 sublineage. The L4.3 sublineage is widespread across the globe as a consequence of early European colonization movements being highly frequent in many countries of Asia, Europe, Africa and America (Brynildsrud et al., 2018). It is still to further investigate whether the higher variation in epitope regions in L4.3 are associated to interaction with different human populations.

Analysis of genome sequence data has also revealed considerable additional diversity that was difficult to spot with previous techniques. A seventh human lineage was described in Ethiopia (Comas et al., 2015;

Firdessa et al., 2013). The lineage falls between the split of lineage 1 and the so called modern lineage (Fig. 1B). The lineage is extremely rare outside certain areas in Ethiopia showing that new branches of the overall phylogeny can be revealed by deeply looking within countries. By far, the branch that have seen a larger diversity undercovered in the last years is the one associated to animals and lineage 6 strains. Lineage 6 is considered one human lineage, historically called *Mycobacterium africanum* together with lineage 5. Early phylogenetic analysis based on the distribution of large sequence polymorphisms showed that *M. africanum* had the unique characteristics of being associated to animal strains (Brosch et al., 2002). Initially it was thought that lineage 6 strains and animal strains were sisters clades (Comas et al., 2013; Hershberg et al., 2008). However, more recent data of tuberculosis strains isolated from wild animals mainly in Africa has shown that some animal strains cluster earlier but closer to the lineage 6 than to animal clade. Among those early branching strains there is the first description of a bacillus isolated from a wild chimpanzee (Brites et al., 2018; Coscolla et al., 2013). Altogether, it is becoming clear that our knowledge of *Mycobacterium tuberculosis* complex diversity particularly associated to wild animals and the *Mycobacterium africanum* lineages is still limited compared to what we know from “modern” strains.

4. Evolutionary genomics identifies genome regions under positive selection in *Mycobacterium tuberculosis* complex

Since the sequencing of the first *Mycobacterium tuberculosis* genome, that of the reference strain H37Rv, new virulence factors among the predicted proteome have been identified (Cole et al., 1998). Different techniques have been used in what can be called post-genomic approaches. For example the random mutagenesis deletion of mutants allows to identify essential and non-essential genes across different conditions as well as the components of different metabolic responses (DeJesus et al., 2017). However, most of this work was based on a single reference strain. The advent of cheaper genomic sequencing led to a complementary approach in which comparative genomics with other mycobacteria and population genomics of thousands of clinical strains has impacted our understanding of virulence factors in MTBC.

Early studies in *Mycobacterium marinum* identified a *ppe38* knockout strain unable to secrete ESX5 dependent substrates involved in virulence (Dong et al., 2012). This gene is part of a larger family, difficult to study with current short-read sequencing techniques, with some of its members linked to virulence and host-pathogen crosstalk. More recent studies using whole genome comparisons corroborated that strains across the MTBC and particularly the Beijing clade had naturally insertions/deletions occurring in this gene (Ates et al., 2018; McEvoy et al., 2009). The mutant strains with *ppe38* inactivated had no capacity to secrete a large number of PE_PPE members leading to a higher virulence (Ates et al., 2018). As Beijing strains are thought to be hypervirulent (Merker et al., 2015) this led to the hypothesis that *ppe38* is under selection in natural populations. With this idea in mind, Ates et al. identify strains across the world with likely *ppe38* inactivating mutations. First, the authors show that the situation in *M. marinum* *ppe38* deleted strains is paralleled by *M. tuberculosis* strains when different parts of the *ppe38*-related region are knocked out. Second, the authors show that clinical strains of MTBC with putative inactivating mutations does not secrete the ESX-5 associated proteins. Third, they show that virulent strains of the Beijing family had inactivation of *pp38* and thus impaired secretion of the PE_PPE protein cluster. Those strains were also high virulent in a mice model of infection showing again that natural inactivation of *ppe38* leads to high virulent strains in vitro and in vivo. Furthermore, the inactivating mutations is shared by a cluster of Beijing strains called “modern” which is particularly successful across the globe (Liu et al., 2018).

Recently, by analyzing a large number of strains collected from patient over 2.5 years in Vietnam a gene belonging to *esx* family, *esxW*, has been shown to be under positive selection (Holt et al., 2018). In this

case, the authors identify a mutation common to all strains of a particular sublineage of Beijing strains. The mutation is also homoplastic, and was identified in other unrelated phylogenetic clades including lineage 1 and lineage 3 strains. Homoplasmy in *M. tuberculosis* is very rare, only around 1–2% of all mutations appear more than once in the phylogeny of the MTBC. Homoplasmy represents the result of a common selective pressure leading to the same nucleotide change in unrelated strains. The most common selective force associated to homoplasmy is drug resistance (Mortimer et al., 2018). Detecting positive selection in *M. tuberculosis* depends on the existence of this convergent episodes of homoplasmy, at the same time the number of homoplasies of a character largely depend on the strength of the selective force. Thus those residues under stronger selective forces are easier to spot than those masked by other phenomena or when selection is weak. Thus, is not surprising that most residues described under positive selection are link to drug resistance, arguably the strongest selective force that *M. tuberculosis* encounters (Farhat et al., 2013). In comparison much more difficult is to find positively associated variants to virulence as the trait is not that easy to define and thus to phenotype.

Analysis of genomics scars of past and present evolutionary pressures can also lead to the identification of virulence factors that are important for extant circulating strains. The MTBC ancestor shared ecological niche with *Mycobacterium canettii* (Blouin et al., 2014; Comas et al., 2015; Supply et al., 2013). An analysis of recombination patterns showed that recombination was a major force before the MTBC emergence. This has allowed to study which regions were important during the emergence, i.e. important to establish an obligate association to the host, and to compare to the selective forces acting on those regions today. An analyses of the selective forces before and after the emergence identified *phoR* under purifying selection before the emergence of the obligate pathogen and then under diversifying selection (Chiner-Oms et al., 2018b). *phoR* encodes a protein that forms part of the PhoPR two-component system. PhoPR is a major virulence regulator in *M. tuberculosis* controlling the expression of many genes involved in host pathogen interaction (Gonzalo-Asensio et al., 2008). It is known that, early in the diversification of the MTBC, a PhoP mutation was linked to adaptation of *M. bovis* to animal strains (Gonzalo-Asensio et al., 2014). That same SNP introduced in an *M. tuberculosis* strain genetic background decreased virulence of the strain in human macrophages and animal models of infection (Gonzalo-Asensio et al., 2014). The decrease in virulence was associated to the lower secretion of key host-pathogen molecules such as the SL and PAT lipids (from the polyacyltrehalose and sulfolipids families) and the lipolytic enzyme LipF. Population genomics of human tuberculosis strains have identified that *phoR* is under positive selection in human populations (Chiner-Oms et al., 2018b). The ratio of non-synonymous to synonymous substitutions is well above 1 (dN/dS = 2,37) and much higher than the mean dN/dS in essential and non-essential genes. Thus PhoR mutations seems to play a role modulating the virulence of MTBC during infection. To show that this is the case, an analysis of all the strains circulating in a region in Malawi during ten years shows that new mutations are arising in the clinical setting and that those mutations are associated to large transmission cluster. PhoR is the sensor component of the two-component system, it is still unclear what PhoR sense and what triggers the need to adapt to changes although it is clear that should associated to the dynamic microenvironment inside the macrophage.

The examples outlined above, as a part of a growing list of research works, shows how the detailed evolutionary study of patterns of diversity in thousands of strains is allowing to identify relevant virulence factors in the tuberculosis host-pathogen interaction. It is still not clear why some of the genetic changes found are confined to certain strains while others seem to be affecting strains across the complex. For example, the impaired function of PPE38 or the *esxW* mutations, although can be convergently found in strains from other lineages, are common to the lineage 2.2.1 strains. The *phoR* case is different, as some mutations map to complete lineages (i.e. lineages 5 and 6) but also more

recently diversified clades or even single strains. Altogether, the examples illustrates the increasing power of population and evolutionary genomics to identify bacterial genetic factors associated to virulence and transmission.

Acknowledgements

IC group is currently funded by projects of the European Research Council (ERC) research grant 638553-TB-ACCELERATE, Ministerio de Economía y Competitividad (Spanish Government) research grant SAF2016-77346-R and Consellería de Educación, Investigación, Cultura i Esport (Valencian government) research grant AICO/2018/113. ACO is recipient of a FPU fellowship from Ministerio de Ciencia, Innovación y Universidades FPU13/00913 (Spanish Government).

References

Alderwick, L.J., Harrison, J., Lloyd, G.S., Birch, H.L., 2015. The mycobacterial cell wall—peptidoglycan and arabinogalactan. *Cold Spring Harb. Perspect. Med.* 5, a021113. <https://doi.org/10.1101/cshperspect.a021113>.

Ates, L.S., Dippenaar, A., Ummels, R., Piersma, S.R., van der Woude, A.D., van der Kuij, K., Le Chevalier, F., Mata-Espinosa, D., Barrios-Payán, J., Marquina-Castillo, B., Guapillo, C., Jiménez, C.R., Pain, A., Houben, E.N.G., Warren, R.M., Brosch, R., Hernández-Pando, R., Bitter, W., 2018. Mutations in ppe38 block PE_{PGRS} secretion and increase virulence of *Mycobacterium tuberculosis*. *Nat. Microbiol.* 3, 181–188. <https://doi.org/10.1038/s41564-017-0090-6>.

Behr, M.A., Wilson, M.A., Gill, W.P., Salamon, H., Schoolnik, G.K., Rane, S., Small, P.M., 1999. Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 284, 1520–1523.

Blouin, Y., Cazajous, G., Dehan, C., Soler, C., Vong, R., Hassan, M.O., Hauck, Y., Boulais, C., Andriamanantena, D., Martinaud, C., Martin, É., Pourcel, C., Vergnaud, G., 2014. Progenitor *Mycobacterium canettii* clone responsible for lymph node tuberculosis epidemic, Djibouti. *Emerg. Infect. Dis.* 20, 21–28. <https://doi.org/10.3201/eid2001.130652>.

Boritsch, E.C., Frigui, W., Cascioferro, A., Malaga, W., Etienne, G., Laval, F., Pawlik, A., Le Chevalier, F., Orgeur, M., Ma, L., Bouchier, C., Stinear, T.P., Supply, P., Majlessi, L., Daffé, M., Guilhot, C., Brosch, R., 2016a. *pkv5*-recombination-mediated surface re-modelling in *Mycobacterium tuberculosis* emergence. *Nat. Microbiol.* 1, 15019. <https://doi.org/10.1038/nmicrobiol.2015.19>.

Boritsch, E.C., Khanna, V., Pawlik, A., Honoré, N., Navas, V.H., Ma, L., Bouchier, C., Seemann, T., Supply, P., Stinear, T.P., Brosch, R., 2016b. Key experimental evidence of chromosomal DNA transfer among selected tuberculosis-causing mycobacteria. *Proc. Natl. Acad. Sci.* 113, 9876–9881. <https://doi.org/10.1073/pnas.1604921113>.

Borrell, S., Gagneux, S., 2009. Infectiousness, reproductive fitness and evolution of drug-resistant *Mycobacterium tuberculosis*. *Int. J. Tuberc. Lung Dis.* 13, 1456–1466.

Braet, S., Vandelanoot, K., Meehan, C.J., Brum Fontes, A.N., Hasker, E., Rosa, P.S., Lucena-Silva, N., Rigouts, L., Suffys, P.N., De Jong, B.C., 2018. The repetitive element RLEP is a highly specific target for detection of *Mycobacterium leprae*. *J. Clin. Microbiol.* 56 <https://doi.org/10.1128/JCM.01924-17>. (e01924-17).

Brites, D., Loiseau, C., Menardo, F., Borrell, S., Boniotti, M.B., Warren, R., Dippenaar, A., Parsons, S.D.C., Beisel, C., Behr, M.A., Fyfe, J.A., Coscolla, M., Gagneux, S., 2018. A new phylogenetic framework for the animal-adapted *Mycobacterium tuberculosis* complex. *Front. Microbiol.* 9 (2820). <https://doi.org/10.3389/fmicb.2018.02820>.

Brosch, R., Gordon, S.V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., Pym, A.S., Samper, S., van Soolingen, D., Cole, S.T., 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc. Natl. Acad. Sci. U. S. A.* 99, 3684–3689. <https://doi.org/10.1073/pnas.052548299>.

Brosch, R., Gordon, S.V., Garnier, T., Eiglmeier, K., Frigui, W., Valenti, P., Dos Santos, S., Duthoy, S., Lacroix, C., Garcia-Pelayo, C., Inwald, J.K., Golby, P., Garcia, J.N., Hewinson, R.G., Behr, M.A., Quail, M.A., Churcher, C., Barrell, B.G., Parkhill, J., Cole, S.T., 2007. Genome plasticity of BCG and impact on vaccine efficacy. *Proc. Natl. Acad. Sci.* 104, 5596–5601. <https://doi.org/10.1073/pnas.0700869104>.

Brynildsrud, O.B., Pepperell, C.S., Suffys, P., Grandjean, L., Monteserin, J., Debeck, N., Bohlin, J., Alfsnes, K., Pettersson, J.O.-H., Kirkeleite, I., Fandinho, F., da Silva, M.A., Perdigão, J., Portugal, I., Viveiros, M., Clark, T., Caws, M., Dunstan, S., Thai, P.V.K., Lopez, B., Ritacco, V., Kitchen, A., Brown, T.S., van Soolingen, D., O'Neill, M.B., Holt, K.E., Feil, E.J., Mathema, B., Balloux, F., Eldholm, V., 2018. Global expansion of *Mycobacterium tuberculosis* lineage 4 shaped by colonial migration and local adaptation. *Sci. Adv.* 4, eaat5869. <https://doi.org/10.1126/sciadv.aat5869>.

Chiner-Oms, Á., González-Candelas, F., Comas, I., 2018a. Gene expression models based on a reference laboratory strain are poor predictors of *Mycobacterium tuberculosis* complex transcriptional diversity. *Sci. Rep.* 8 (3813). <https://doi.org/10.1038/s41598-018-22237-5>.

Chiner-Oms, Á., Sánchez-Busó, L., Corander, J., Gagneux, S., Harris, S., Young, D., González-Candelas, F., Comas, I., 2018b. Genomic Determinants of Sympatric Speciation of the *Mycobacterium tuberculosis* Complex across Evolutionary Timescales. *bioRxiv*, pp. 314559. <https://doi.org/10.1101/314559>.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., Tekaia, F., Badcock, K., Basham, D., Brown, D.,

Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.-A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S., Barrell, B.G., 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393, 537–544. <https://doi.org/10.1038/31159>.

Comas, I., Gagneux, S., 2009. The past and future of tuberculosis research. *PLoS Pathog.* 5, e1000600. <https://doi.org/10.1371/journal.ppat.1000600>.

Comas, I., Homolka, S., Niemann, S., Gagneux, S., 2009. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS One* 4, e7815. <https://doi.org/10.1371/journal.pone.0007815>.

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. <https://doi.org/10.1038/ng.590>.

Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K.E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S.R., Niemann, S., Diel, R., Aseffa, A., Gao, Q., Young, D., Gagneux, S., 2013. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat. Genet.* 45, 1176–1182.

Comas, I., Hailu, E., Kiro, T., Bekele, S., Mekonnen, W., Gumi, B., Tschopp, R., Ameni, G., Hewinson, R.G., Robertson, B.D., Goig, G.A., Stucki, D., Gagneux, S., Aseffa, A., Young, D., Berg, S., 2015. Population genomics of *Mycobacterium tuberculosis* in Ethiopia contradicts the virgin soil hypothesis for human tuberculosis in Sub-Saharan Africa. *Curr. Biol.* 25, 3260–3266. <https://doi.org/10.1016/j.cub.2015.10.061>.

Coscolla, M., Gagneux, S., 2014. Consequences of genomic diversity in *Mycobacterium tuberculosis*. *Semin. Immunol.* 26, 431–444.

Coscolla, M., Lewin, A., Metzger, S., Maetz-Rensing, K., Calvignac-Spencer, S., Nitsche, A., Dabrowski, P.W., Radonic, A., Niemann, S., Parkhill, J., Couacy-Hymann, E., Feldman, J., Comas, I., Boesch, C., Gagneux, S., Leendertz, F.H., 2013. Novel *Mycobacterium tuberculosis* complex isolate from a wild chimpanzee. *Emerg. Infect. Dis.* 19, 969–976. <https://doi.org/10.3201/eid1906.121012>.

Coscolla, M., Copin, R., Sutherland, J., Gehre, F., de Jong, B., Owolabi, O., Mbayo, G., Giardina, F., Ernst, J.D., Gagneux, S., 2015. M. tuberculosis T cell epitope analysis reveals paucity of antigenic variation and identifies rare variable TB antigens. *Cell Host Microbe* 18, 538–548. <https://doi.org/10.1016/j.chom.2015.10.008>.

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. <https://doi.org/10.1371/journal.pntd.0000744>.

DeJesus, M.A., Gerrick, E.R., Xu, W., Park, S.W., Long, J.E., Boutte, C.C., Rubin, E.J., Schnappinger, D., Ehr, S., Fortune, S.M., Sasseti, C.M., Ioerger, T.R., 2017. Comprehensive essentiality analysis of the *Mycobacterium tuberculosis* genome via saturating transposon mutagenesis. *MBio* <https://doi.org/10.1128/mBio.02133-16>. (e02133–16).

Dong, D., Wang, D., Li, M., Wang, H., Yu, J., Wang, C., Liu, J., Gao, Q., 2012. PPE38 modulates the innate immune response and is required for *Mycobacterium marinum* virulence. *Infect. Immun.* 80, 43–54. <https://doi.org/10.1128/IAI.05249-11>.

Farhat, M.R., Shapiro, B.J., Kieser, K.J., Sultana, R., Jacobson, K.R., Victor, T.C., Warren, R.M., Streicher, E.M., Calver, A., Sloutsky, A., Kaur, D., Posey, J.E., Plikaytis, B., Oggioni, M.R., Gardy, J.L., Johnston, J.C., Rodrigues, M., Tang, P.K.C., Kato-Maeda, M., Borowsky, M.L., Muddukrishna, B., Kreiswirth, B.N., Kurepina, N., Galagan, J., Gagneux, S., Birren, B., Rubin, E.J., Lander, E.S., Sabeti, P.C., Murray, M., 2013. Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis*. *Nat. Genet.* 45, 1183–1189. <https://doi.org/10.1038/ng.2747>.

Fedrizzi, T., Meehan, C.J., Grotto, A., Giacobazzi, E., Fregni Serpini, G., Tagliacucchi, S., Fabio, A., Bettua, C., Bertorelli, R., De Sanctis, V., Rumpianesi, F., Pecorari, M., Jousson, O., Tortoli, E., Segata, N., 2017. Genomic characterization of non-tuberculous mycobacteria. *Sci. Rep.* 7, 45258. <https://doi.org/10.1038/srep45258>.

Fenner, L., Egger, M., Bodmer, T., Furrer, H., Ballif, M., Bategay, M., Helbling, P., Fehr, J., Gspomer, T., Rieder, H.L., Zwahlen, M., Hoffmann, M., Bernasconi, E., Cavassini, M., Calmy, A., Dolina, M., Frei, R., Janssens, J.-P., Borrell, S., Stucki, D., Schrenzel, J., Böttger, E.C., Gagneux, S., Groups, for the S.H.C. and M.E. of T.S., 2013. HIV infection disrupts the sympatric host–pathogen relationship in human tuberculosis. *PLoS Genet.* e1003318, 9. <https://doi.org/10.1371/journal.pgen.1003318>.

Firdessa, R., Berg, S., Hailu, E., Schelling, E., Gumi, B., Erenso, G., Gadisa, E., Kiro, T., Habtamu, M., Hussein, J., Zinsstag, J., Robertson, B.D., Ameni, G., Lohan, A.J., Loftus, B., Comas, I., Gagneux, S., Tschopp, R., Yamuah, L., Hewinson, G., Gordon, S.V., Young, D.B., Aseffa, A., 2013. Mycobacterial lineages causing pulmonary and extrapulmonary tuberculosis, Ethiopia. *Emerg. Infect. Dis.* 19, 460–463. <https://doi.org/10.3201/eid1903.120256>.

Ford, C.B., Shah, R.R., Maeda, M.K., Gagneux, S., Murray, M.B., Cohen, T., Johnston, J.C., Gardy, J., Lipsitch, M., Fortune, S.M., 2013. *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis. *Nat. Genet.* 45, 784–790. <https://doi.org/10.1038/ng.2656>.

Gagneux, S., 2018. Ecology and evolution of *Mycobacterium tuberculosis*. *Nat. Rev. Microbiol.* 16, 202–213. <https://doi.org/10.1038/nrmicro.2018.8>.

Gagneux, S., Small, P.M., 2007. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect. Dis.* 7, 328–337. [https://doi.org/10.1016/S1473-3099\(07\)70108-1](https://doi.org/10.1016/S1473-3099(07)70108-1).

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B.C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., Hilty, M., Hopewell, P.C., Small, P.M., 2006. Variable host–pathogen compatibility in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* 103, 2869–2873.

Garnier, T., Eiglemer, K., Camus, J.-C., Medina, N., Mansoor, H., Pryor, M., Duthoy, S., Grondin, S., Croaix, C., Monsempe, C., Simon, S., Harris, B., Atkin, R., Doggett, J., Mayes, R., Keating, L., Wheeler, P.R., Parkhill, J., Barrell, B.G., Cole, S.T., Gordon, S.V., Hewinson, R.G., 2003. The complete genome sequence of *Mycobacterium bovis*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 7877–7882. <https://doi.org/10.1073/pnas.1130426100>.

Gonzalo-Asensio, J., Mostowy, S., Harders-Westerveen, J., Huygen, K., Hernández-Pando, R., Thole, J., Behr, M., Gicquel, B., Martin, C., 2008. PhoP: a missing piece in the intricate puzzle of *Mycobacterium tuberculosis* virulence. *PLoS One* 3 (10), e3496.

Gonzalo-Asensio, J., Malaga, W., Pawlik, A., Astarie-Dequeker, C., Passemar, C., Moreau, F., Laval, F., Daffé, M., Martin, C., Brosch, R., Guilhot, C., 2014. Evolutionary history of tuberculosis shaped by conserved mutations in the PhoPR virulence regulator. *Proc. Natl. Acad. Sci.* 111, 11491–11496.

Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., Tiedje, J.M., 2007. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. <https://doi.org/10.1099/ijs.0.64483-0>.

Gray, T.A., Derbyshire, K.M., 2018. Blending genomes: distributive conjugal transfer in mycobacteria, a sexier form of HGT. *Mol. Microbiol.* <https://doi.org/10.1111/mmi.13971>.

Gupta, R.S., Lo, B., Son, J., 2018. Phylogenomics and comparative genomic studies robustly support division of the genus *Mycobacterium* into an emended genus *Mycobacterium* and four novel genera. *Front. Microbiol.* 9 (67). <https://doi.org/10.3389/fmicb.2018.00067>.

Hatfull, G.F., Jacobs, W.R., 2000. *Molecular Genetics of Mycobacteria*. ASM Press.

Hersheyberg, R., Lipatov, M., Small, P.M., Sheffer, H., Niemann, S., Homolka, S., Roach, J.C., Kremer, K., Petrov, D.A., Feldman, M.W., Gagneux, S., 2008. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol.* 6, e311. <https://doi.org/10.1371/journal.pbio.0060311>.

Holt, K.E., McAdam, P., Thai, P.V.K., Thuong, N.T.T., Ha, D.T.M., Lan, N.N., Lan, N.H., Nhu, N.T.Q., Hai, H.T., Ha, V.T.N., Thwaites, G., Edwards, D.J., Nath, A.P., Pham, K., Ascher, D.B., Farrar, J., Khor, C.C., Teo, Y.Y., Inouye, M., Caws, M., Dunstan, S.J., 2018. Frequent transmission of the *Mycobacterium tuberculosis* Beijing lineage and positive selection for the EsxW Beijing variant in Vietnam. *Nat. Genet.* 50, 849–856. <https://doi.org/10.1038/s41588-018-0117-9>.

Jankute, M., Cox, J.A.G., Harrison, J., Besra, G.S., 2015. Assembly of the Mycobacterial Cell Wall. <https://doi.org/10.1146/annurev-micro-091014-104121>.

Koeck, J.-L., Fabre, M., Simon, F., Daffé, M., Garnotel, É., Matan, A.B., Jérôme, P., Bernatas, J.-J., Buisson, Y., Pourcel, C., 2011. Clinical characteristics of the smooth tubercle bacilli *Mycobacterium canettii* infection suggest the existence of an environmental reservoir. *Clin. Microbiol. Infect.* 17, 1013–1019. <https://doi.org/10.1111/J.1469-0691.2010.03347.X>.

Liu, X., Gutacker, M.M., Musser, J.M., Fu, Y.-X., 2006. Evidence for recombination in *Mycobacterium tuberculosis*. *J. Bacteriol.* 188, 8169–8177. <https://doi.org/10.1128/JB.01062-06>.

Liu, Q., Ma, A., Wei, L., Pang, Y., Wu, B., Luo, T., Zhou, Y., Zheng, H.-X., Jiang, Q., Gan, M., Zuo, T., Liu, M., Yang, C., Jin, L., Comas, I., Gagneux, S., Zhao, Y., Pepperell, C.S., Gao, Q., 2018. China's tuberculosis epidemic stems from historical expansion of four strains of *Mycobacterium tuberculosis*. *Nat. Ecol. Evol.* 1. <https://doi.org/10.1038/s41559-018-0680-6>.

McEvoy, C.R.E., van Helden, P.D., Warren, R.M., Gey van Pittius, N.C., 2009. Evidence for a rapid rate of molecular evolution at the hypervariable and immunogenic *Mycobacterium tuberculosis* PPE38 gene region. *BMC Evol. Biol.* 9, 237. <https://doi.org/10.1186/1471-2148-9-237>.

Merker, M., Blin, C., Mona, S., Duforet-Frebourg, N., Lecher, S., Willery, E., Blum, M.G.B., Rüsche-Gerdes, S., Mokrousov, I., Aleksic, E., Allix-Béguec, C., Antierens, A., Augustynowicz-Kopec, E., Ballif, M., Barletta, F., Beck, H.P., Barry, C.E., Bonnet, M., Borroni, E., Campos-Herrero, I., Cirillo, D., Cox, H., Crowe, S., Crudu, V., Diel, R., Drobniewski, F., Fauville-Dufaux, M., Gagneux, S., Ghebremichael, S., Hanekom, M., Hoffner, S., Jiao, W., Kalon, S., Kohl, T.A., Kontsevaya, I., Lillebaek, T., Maeda, S., Nikolayevskyy, V., Rasmussen, M., Rastogi, N., Samper, S., Sanchez-Padilla, E., Savic, B., Shamputa, I.C., Shen, A., Sng, L.-H., Stakenas, P., Toit, K., Varaine, F., Vukovic, D., Wahl, C., Warren, R., Supply, P., Niemann, S., Wirth, T., 2015. Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. *Nat. Genet.* 47, 242–249. <https://doi.org/10.1038/ng.3195>.

Mortimer, T.D., Pepperell, C.S., 2014. Genomic signatures of distributive conjugal transfer among mycobacteria. *Genome Biol. Evol.* 6, 2489–2500. <https://doi.org/10.1093/gbe/evu175>.

Mortimer, T.D., Weber, A.M., Pepperell, C.S., 2018. Signatures of selection at drug resistance loci in *Mycobacterium tuberculosis*. *mSystems* 3, e00108–e00117. <https://doi.org/10.1128/MSYSTEMS.00108-17>.

Nahid, P., Dorman, S.E., Alipanah, N., Barry, P.M., Brozek, J.L., Cattamanchi, A., Chaisson, L.H., Chaisson, R.E., Daley, C.L., Grzemska, M., Higashi, J.M., Ho, C.S., Hopewell, P.C., Keshavjee, S.A., Lienhardt, C., Menzies, R., Merrifield, C., Narita, M., O'Brien, R., Peloquin, C.A., Raftery, A., Saukkonen, J., Schaaf, H.S., Sotgiu, G., Starke, J.R., Migliori, G.B., Vernon, A., 2016. Executive summary: official American Thoracic Society/Centers for Disease Control and Prevention/Infectious Diseases Society of America clinical practice guidelines: treatment of drug-susceptible tuberculosis. *Clin. Infect. Dis.* 63, 853–867. <https://doi.org/10.1093/cid/ciw566>.

Namouchi, A., Didelot, X., Schöck, U., Gicquel, B., Rocha, E.P.C., 2012. After the bottleneck: genome-wide diversification of the *Mycobacterium tuberculosis* complex by mutation, recombination, and natural selection. *Genome Res.* 22, 721–734. <https://doi.org/10.1101/gr.129544.111>.

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. <https://doi.org/10.1038/nature02837>.

Reed, M.B., Pichler, V.K., McIntosh, F., Mattia, A., Fallow, A., Masala, S., Domenech, P.,

- Zwerling, A., Thibert, L., Menzies, D., Schwartzman, K., Behr, M.A., 2009. Major *Mycobacterium tuberculosis* lineages associate with patient country of origin. *J. Clin. Microbiol.* 47, 1119–1128. <https://doi.org/10.1128/JCM.02142-08>.
- Runyon, E.H., 1959. Anonymous mycobacteria in pulmonary disease. *Med. Clin. North Am.* 43, 273–290.
- Rutaihua, L.K., Menardo, F., Stucki, D., Gygli, S.M., Ley, S.D., Malla, B., Feldmann, J., Farnov, S.B., Beisel, C., Middelkoop, K., Carter, J.E., Diero, L., Ballif, M., Jugheli, L., Reither, K., Fenner, L., Brites, D., Gagneux, S., 2018. Multiple Introductions of the *Mycobacterium tuberculosis* Lineage 2 Beijing into Africa over Centuries. *bioRxiv*, pp. 413039. <https://doi.org/10.1101/413039>.
- Saini, V., Raghuvanshi, S., Khurana, J.P., Ahmed, N., Hasnain, S.E., Tyagi, A.K., Tyagi, A.K., 2012. Massive gene acquisitions in *Mycobacterium indicus pranii* provide a perspective on mycobacterial evolution. *Nucleic Acids Res.* 40, 10832–10850. <https://doi.org/10.1093/nar/gks793>.
- Sheppard, S.K., Guttman, D.S., Fitzgerald, J.R., 2018. Population genomics of bacterial host adaptation. *Nat. Rev. Genet.* 19, 549–565. <https://doi.org/10.1038/s41576-018-0032-z>.
- Smith, N.H., Dale, J., Inwald, J., Palmer, S., Gordon, S.V., Hewinson, R.G., Smith, J.M., 2003. The population structure of *Mycobacterium bovis* in Great Britain: clonal expansion. *Proc. Natl. Acad. Sci. U. S. A.* 100, 15271–15275. <https://doi.org/10.1073/pnas.2036554100>.
- Smith, N.H., Gordon, S.V., de la Rúa-Domenech, R., Clifton-Hadley, R.S., Hewinson, R.G., 2006. Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*. *Nat. Rev. Microbiol.* 4, 670–681. <https://doi.org/10.1038/nrmicro1472>.
- Stucki, D., Brites, D., Jeljeli, L., Coscolla, M., Liu, Q., Trauner, A., Fenner, L., Rutaihua, L., Borrell, S., Luo, T., Gao, Q., Kato-Maeda, M., Ballif, M., Egger, M., Macedo, R., Mardassi, H., Moreno, M., Tundo Vilanova, G., Fyfe, J., Globan, M., Thomas, J., Jamieson, F., Guthrie, J.L., Asante-Poku, A., Yeboah-Manu, D., Wampande, E., Ssengooba, W., Joloba, M., Henry Boom, W., Basu, I., Bower, J., Saraiva, M., Vacconcellos, S.E.G., Suffys, P., Koch, A., Wilkinson, R., Gail-Bekker, L., Malla, B., Ley, S.D., Beck, H.-P., de Jong, B.C., Toit, K., Sanchez-Padilla, E., Bonnet, M., Gil-Brusola, A., Frank, M., Penlap Beng, V.N., Eisenach, K., Alani, I., Wangui Ndung'u, P., Revathi, G., Gehre, F., Akter, S., Ntoumi, F., Stewart-Isherwood, L., Ntinginya, N.E., Rachow, A., Hoelscher, M., Cirillo, D.M., Skenders, G., Hoffner, S., Bakonyte, D., Stakenas, P., Diel, R., Crudu, V., Moldovan, O., Al-Hajaj, S., Otero, L., Barletta, F., Jane Carter, E., Diero, L., Supply, P., Comas, I., Niemann, S., Gagneux, S., 2016. *Mycobacterium tuberculosis* lineage 4 comprises globally distributed and geographically restricted sublineages. *Nat. Genet.* 48, 1535–1543. <https://doi.org/10.1038/ng.3704>.
- Supply, P., Marceau, M., Mangenot, S., Roche, D., Rouanet, C., Khanna, V., Majlessi, L., Criscuolo, A., Tap, J., Pawlik, A., Fiette, L., Orgeur, M., Fabre, M., Parmentier, C., Frigui, W., Simeone, R., Boritsch, E.C., Debie, A.-S., Willery, E., Walker, D., Quail, M.A., Ma, L., Bouchier, C., Salvignol, G., Sayes, F., Cascioferro, A., Seemann, T., Barbe, V., Locht, C., Gutierrez, M.-C., Leclerc, C., Bentley, S.D., Stinear, T.P., Brisse, S., Médigue, C., Parkhill, J., Cruveiller, S., Brosch, R., 2013. Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of *Mycobacterium tuberculosis*. *Nat. Genet.* 45, 172–179. <https://doi.org/10.1038/ng.2517>.
- Tortoli, E., Fedrizzi, T., Meehan, C.J., Trovato, A., Grottole, A., Giacobazzi, E., Serpini, G.F., Tagliazucchi, S., Fabio, A., Bettua, C., Bertorelli, R., Frascaro, F., De Sanctis, V., Pecorari, M., Jousson, O., Segata, N., Cirillo, D.M., 2017. The new phylogeny of the genus *Mycobacterium*: the old and the news. *Infect. Genet. Evol.* 56, 19–25. <https://doi.org/10.1016/J.MEEGID.2017.10.013>.
- Turenne, C.Y., 2019. Nontuberculous mycobacteria: insights on taxonomy and evolution. *Infect. Genet. Evol.* <https://doi.org/10.1016/J.MEEGID.2019.01.017>.
- van Helden, P.D., Warren, R.M., Victor, T.C., van der Spuy, G., Richardson, M., Hoal-van Helden, E., 2002. Strain families of *Mycobacterium tuberculosis*. *Trends Microbiol.* 10, 167–168 (author reply 168).
- van Ingen, J., 2017. Mycobacteria. *Infect. Dis. (Auckl.)* e2, 1645–1659. <https://doi.org/10.1016/B978-0-7020-6285-8.00185-4>.