

Mycobacterium Tuberculosis Evolutionary Pathogenesis and its Putative Impact on Drug Development

Fabien Le Chevalier, Alessandro Cascioferro, Laleh Majlessi, Jean Louis Herrmann, Roland Brosch


Abstract and Introduction

Abstract

Mycobacterium tuberculosis, the etiological agent of human TB, is the most important mycobacterial pathogen in terms of global patient numbers and gravity of disease. The molecular mechanisms by which M. tuberculosis causes disease are complex and the result of host–pathogen coevolution that might have started already in the time of its Mycobacterium canettii-like progenitors. Despite research progress, M. tuberculosis still holds many secrets of its successful strategy for circumventing host defences, persisting in the host and developing resistance, which makes anti-TB treatment regimens extremely long and often inefficient. Here, we discuss what we have learned from recent studies on the evolution of the pathogen and its putative new drug targets that are essential for mycobacterial growth under in vitro or in vivo conditions.

Introduction

Pathogenic mycobacteria are important causes of human and animal disease. Despite the availability of antibiotics and chemotherapeutic agents that show activity against certain mycobacteria, different drug-resistant forms are on the rise, which requires new efforts in the search for novel active compounds and treatment strategies. Phylogenetically, the genus Mycobacterium is contained within the phylum Actinobacteria, with the genera Corynebacterium and Streptomyces as close relatives. Most of the approximately 130 defined mycobacterial species contain harmless saprophytes, which is most evident for the large group of fast-growing mycobacteria that comprise only very few pathogenic species, such as the Mycobacterium abscessus–Mycobacterium massiliense–Mycobacterium bolletti complex. These latter mycobacteria represent emerging opportunistic pathogens that are more and more often being recognized as causative agents of acute and persistent lung infections in cystic fibrosis patients and are also virulent in animal infection models. They pose severe difficulties for treatment in humans due to their extensive level of drug resistance.

In contrast to the phylogenetically more diversified fast-growing mycobacteria, the slow-growing mycobacteria form a subcluster in the 16S rRNA-based phylogenetic tree and are well known to harbor major human pathogens, such as Mycobacterium leprae and Mycobacterium tuberculosis, and also a series of opportunistic human pathogens, such as Mycobacterium marinum or Mycobacterium kansasii, which are considered to be close relatives of M. tuberculosis. Overall, mycobacteria are considered to be high-GC content, Gram-positive bacteria, but in contrast to other Gram-positive bacteria, such as Staphylococcus or Bacillus species, which have no outer membrane, mycobacteria possess a lipid-rich cell envelope that contains a standard inner membrane and a particular outer membrane, named the mycomembrane, which is specific to mycobacteria and might fulfill a similar barrier function as the outer membrane of Gram-negative bacteria. In addition, the mycobacterial cell is also covered by a polysaccharide-based capsule. This overall envelope architecture is conserved between nonpathogenic and pathogenic mycobacteria and contributes to the enhanced ability of mycobacteria to resist and persist in different environments. The complex cell envelope necessitates efficient secretion...
systems that can ensure the transport of a range of biomolecules across this multilayer barrier. Despite the conservation of many core elements in the mycobacterial cell envelope (reviewed in[20]) (Figure 1), over the course of evolution, individual changes in the composition of cell wall components, such as particular lipids or specific secretion and transport systems, have emerged that contribute to the specific lifestyles of pathogenic mycobacterial species. Evolution towards pathogenicity has often been accompanied by a reduction of genome size that is compensated in part by gene acquisition, gene duplications and diversification (Figure 2A).[11,13,14,21–23] When *M. tuberculosis* is compared with *M. marinum*, which is characterized by an approximately 6.7-Mb genome and the more distantly related, fast-growing *Mycobacterium smegmatis* that harbors an approximately 7-Mb genome, a core of 2450 genes that are conserved among the three species can be defined (Figure 2B). This comparison also showed that approximately 600 genes are specific for *M. tuberculosis* and might contribute to the pathogenicity of *M. tuberculosis* relative to the other mycobacterial species.[13] When *M. tuberculosis* is compared within the group of tubercle bacilli (i.e., mycobacteria that cause TB in mammalian species), recent genome comparisons have shown that *Mycobacterium canettii* strains, which represent rare human isolates with smooth colony morphology from the region of the Horn of Africa, harbor somewhat larger genomes that differ from *M. tuberculosis* by 16,000–60,000 single-nucleotide polymorphisms and individual accessory genomes of up to 366 genes,[23] which argues for their ancestral status relative to *M. tuberculosis*. As these *M. canettii* strains were found to be less virulent and less persistent than *M. tuberculosis* sensu stricto in laboratory mouse models, their genomes provide very interesting gene repertoires for learning more about the ancestral gene content of the putative progenitor of *M. tuberculosis* and about possible gene transfers that might have contributed to the virulence gain of *M. tuberculosis* during evolution. Comparison of several *M. canettii* strains with *M. tuberculosis* revealed some 51 genes that were specifically present in *M. tuberculosis* and absent from *M. canettii*,[23] such as the gene *rv1818* encoding a specific member (PE_PGRS33) of a large mycobacterial protein family named after the Pro–Glu motif in the N-terminus and the presence of polymorphic GC-rich repetitive sequences in the central part or the C-terminus of the proteins.[11] On the other hand, nine genes were found that were specifically absent from *M. tuberculosis* relative to *M. canettii*, such as the *cobF* gene encoding an enzyme in the vitamin ref-12 metabolic pathway.[23,24] The loss of *cobF* from *M. tuberculosis* might be compensated by substrates provided by the host, which is presently under investigation. As such, the different genes and their gene products that differ between *M. canettii* and *M. tuberculosis* can now be further investigated for their impact on the different phenotypes of the concerned phylogenetically diverging tubercle bacilli.
Figure 1.

**Organization of the mycobacterial cell envelope.** (A) Model of the cell envelope of *Mycobacterium tuberculosis* and other tubercle bacilli. (B) Cell envelope of *Mycobacterium bovis* BCG by cryoelectron microscopy of vitreous sections.

(A) Adapted with permission from [15]; (B) Reproduced with permission from [16].
Phylogenetic organization and gene content of members of the genus *Mycobacterium*. (A) Phylogenetic tree based on the consensus of the 230 most parsimonious trees of the 16S rRNA DNA sequences of 80 species of the genus *Mycobacterium* with the sequence of the species *Gordonia aichiensis* as the outgroup. Sequenced genomes are highlighted in yellow and underlined species are considered to be pathogens. The division between fast- and slow-growing species is indicated by a dotted line. Genome sizes are indicated according to information from the Gold Genome online database [25]. (B) Venn diagram showing orthologous coding sequences among three mycobacterial species as determined by BLASTClust analysis. Numbers in parentheses include paralogous coding sequences.

† *Mycobacterium farcinogenes* a slow-growing *Mycobacterium*.

(A) Adapted with permission from [22]; (B) Adapted with permission from [13].

As mentioned above, genome comparisons between fast-growing mycobacterial saprophytes and slow-growing mycobacterial pathogens show an overall trend of genome size reduction. A possible explanation for the genome downsizing of pathogenic mycobacteria could be that adaptation to a pathogenic lifestyle that includes the exploitation of resources of a host organism might make certain gene functions redundant and subject to gene loss. This feature is most visible in the only 3.2-Mb genome of the obligate intracellular pathogen *M. leprae* that has undergone a dramatic reductive evolution due to gene decay and gene loss, which is responsible for the fact that *M. leprae*, despite its discovery by Armauer Hansen more than 140 years ago, can still not be cultured on axenic growth media.[10]

The situation of genome reduction is less pronounced for *M. tuberculosis*, which harbors a 4.4-Mb genome and has retained the ability to grow on axenic culture media,[11] potentially reflecting the extraordinary faculty of this human pathogen to adapt to numerous intra- and extra-cellular environments encountered during its interplay with the host. The particularity of *M. tuberculosis* in comparison with many facultative mycobacterial pathogens is that it is an obligatory pathogen whose transmission to new hosts occurs from patients who have developed active disease, implicating lung tissue necrosis, cavity formation and coughing-up of infectious bacilli into the immediate environment as tiny droplets. This infection strategy seems to be highly efficient as it is estimated that approximately a third of the human population is infected with *M. tuberculosis*, from which 5–10% develop active disease. *M. tuberculosis* is responsible for 9 million cases of active TB and 1.3 million deaths per year.[26] This situation suggests that by selection, the pathogen has developed a balance for using the human host for its own proliferation and global distribution during a long-lasting coevolution.[23,27,28] This situation poses enormous challenges to scientists in order to unravel the factors that are responsible for the evolutionary success of *M. tuberculosis*. Knowledge on the molecular features that contribute to pathogenicity, persistence and efficient spreading of *M. tuberculosis* among humans will provide important input for the development of alternative strategies in TB control.

Genomics & Identification of the Essential Genes of *M. Tuberculosis*

Within the more than 130 years since the discovery of *M. tuberculosis* as the causative microorganism responsible for human TB by Robert Koch, numerous scientific advances have been made that help to cope with this major pathogen. However, despite this progress, *M. tuberculosis* still holds many
unresolved secrets and much work remains to be done in order to translate the basic findings from recent research into novel strategies against the pathogen. After several anti-TB drugs were developed in the mid-20th century, there has been a long gap in which no new molecules to be used in the treatment of TB were developed. One of the reasons for that could have been that TB was believed to be a disease of the past, which would disappear in modern societies due to improved social conditions, vaccination and already-existing treatment regimens. However, it is now clear that this was not the case for developing countries and also certain population groups in industrialized countries. The synergy of TB with the HIV epidemic, the phenomenon of multidrug resistance, massive global population growth and a worsening of social conditions in many countries dramatically changed the situation, making TB a continuing global health threat even at the beginning of the 21st century. This fact has also lead to a strongly renewed research interest that is likely to be continued in the future. In this context, the accomplishment and publication of the genome sequence of the widely used M. tuberculosis H37Rv strain was one key step that allowed the scientific community to gain deeper insight into the overall organization of this pathogen and to design numerous follow-up strategies in a wide range of postgenomic applications.

For scientists interested in increasing the arsenal of new compounds and drugs against TB and its causative organism, the knowledge of the complete gene pool of the pathogen has helped considerably with the identification and investigation of potentially vulnerable targets of M. tuberculosis. The genome information served as a starting point for identifying which among the 4000 genes in the M. tuberculosis genome were essential for the growth and/or survival of the pathogen. The genome-wide use of high-density transposon mutagenesis in combination with the microarray-based identification of insertion sites revealed a minimal gene set of 614 genes required for optimal in vitro growth of M. tuberculosis in culture broth. This technique, named transposon site hybridization (TraSH), was subsequently also adapted to other biological settings, such as infection models in mice, where a set of 194 genes implicated in growth under in vivo conditions was identified. Among them, numerous genes encoding proteins involved in the pathogenicity of the tubercle bacilli can be found that are also discussed in a later section. Most recently, saturation transposon mutagenesis of M. tuberculosis grown under different in vitro conditions was combined with next-generation sequencing (NGS), which allowed a set of 774 genes to be identified, of which 451 overlapped with the initial TraSH-based screen. In addition, this work permitted differences in the essential gene set to be linked to the use of cholesterol instead of glycerol as a carbon source. Identification of the bacterial genes required for cholesterol utilization in M. tuberculosis allowed the widespread metabolic changes to be predicted that are associated with the adaptation of the bacterium to this carbon source and refined the understanding of bacterial physiology.

Potential Roles of Essential Genes as Drug Targets

The identification of a set of essential genes for M. tuberculosis has greatly enhanced the ability to define new, potentially vulnerable drug targets of the pathogen. In combination with the advances in crystallography and the increasing availability of structural information on mycobacterial proteins, in silico drug screening should be a promising way to find new active molecules against M. tuberculosis. However, this approach must also be taken with caution, as binding of a given compound to an essential target protein under experimental condition often does not translate to actual growth inhibition of the bacterium in subsequent MIC screens. PknB, a serine–threonine protein kinase of M. tuberculosis, represents a typical example of such a situation. This protein, which is part of the 11 eukaryotic-like serine–threonine protein kinases of M. tuberculosis, is listed as an essential protein in the M. tuberculosis TraSH and NGS datasets. The essentiality of this protein has also been suggested by independent genetic studies, in which the knockout of the original pknB gene only became possible once a second intact pknB copy was integrated into the M. tuberculosis genome. Similarly, the essentiality of
PknB was also demonstrated by gene knockdown experiments, which showed a dramatic impairment of bacterial viability in liquid culture upon depletion of the kinase.\(^{(32)}\) Nevertheless, potential inhibitors that showed affinity to PknB were active only in the higher micromolar MIC range,\(^{(34,35)}\) which makes their further development into potential anti-TB drugs rather difficult and/or unlikely. For proteins involved in the regulation of essential cell processes, very small amounts might already be sufficient, so that it is very difficult for an inhibitor to fully block the function of the protein.

Apart from potential quantitative particularities, many other factors may impact on the final inhibition activity of a compound on the bacterium. For \textit{M. tuberculosis}, the complex, lipid-rich cell wall composition, which in addition is protected by a polysaccharide-based capsule,\(^{(15,18,19)}\) certainly complicates permeability. In addition, the potential effect of efflux pumps and protein binding in the assay media may also play a role in the evaluation of the final antimicrobial activity of a compound. Thus, according to most recent trends in anti-TB drug development, whole-cell screening of compound libraries has re-emerged as a frequently used technique to find new active compounds against \textit{M. tuberculosis} for which the drug target can then be identified in follow-up studies. This approach is of great importance for enlarging the arsenal of presently used first-line (isoniazid, rifampicin, ethambutol and pyrazinamide) and second-line (streptomycin, fluoroquinolones, para-aminosalicylic acid and the injectable agents amikacin, kanamycin or capreomycin) anti-TB drug regimens (\&)\(^{(36)}\) that may all be subject to resistance mechanisms in extensively and/or totally resistant \textit{M. tuberculosis} strains.\(^{(37)}\)

Table 1. First- and second-line anti-TB drugs used in standard treatment regimens.

<table>
<thead>
<tr>
<th>Anti-TB drug</th>
<th>Principle target(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isoniazid</td>
<td>InhA; KatG (activation)</td>
</tr>
<tr>
<td>Rifampicin</td>
<td>RpoB</td>
</tr>
<tr>
<td>Ethambutol</td>
<td>EmbB</td>
</tr>
<tr>
<td>Pyrazinamide</td>
<td>RpsA; PncA (activation)</td>
</tr>
<tr>
<td>Streptomycin</td>
<td>RpsL</td>
</tr>
<tr>
<td>Fluoroquinolones</td>
<td>GyrA</td>
</tr>
<tr>
<td>Para-aminosalicylic acid</td>
<td>Dihydropteroate synthase; folate metabolism</td>
</tr>
</tbody>
</table>

\textbf{Injectable agents}

| Amikacin, kanamycin or capreomycin | 30S ribosomal subunit or interbridge |
| Ethionamide                      | InhA; EthA (activation)              |
| Cycloserine                      | D-alanine racemase and ligase        |

Data taken from [36,38].

Table 2. Recommended anti-TB drug regimens; including details of the initial phase and continuation phase.
| New smear- or culture-positive cases | (months) | (months) | Isoniazid, rifampicin, pyrazinamide and ethambutol | 4 | Isoniazid and rifampicin |
| New culture-negative cases | 4 | 2 | Isoniazid, rifampicin, pyrazinamide and ethambutol | 2 | Isoniazid and rifampicin |
| Resistance (or intolerance) to isoniazid | 6 | 6 | rifampicin, pethambutol |  |  |
| Resistance to isoniazid and rifampicin | 12–18 | 12–18 | Pyrazinamide, ethambutol, fluoroquinolones and streptomycin (or another injectable agent) |  |  |
| Resistance to all first-line drugs | 24 | 24 | One injectable agent + three drugs from the following group: ethionamide, cycloserine, fluoroquinolones or para-aminosalicylic acid |  |  |

Data taken from [36,38].

Screening and target identification studies allowed the diarylquinoline drug TMC207 to be discovered, for example, for which resistant mutants showed single-nucleotide polymorphisms in the gene atpE encoding the C chain of the ATP synthase of *M. tuberculosis*. The activity of TMC207 on multidrug-resistant *M. tuberculosis* was confirmed in clinical trials, and in December 2012, the drug (also known as bedaquiline or Sirturo™ [Janssen Therapeutics, NJ, USA]) was approved by the US FDA as a new treatment for multidrug-resistant TB that can be used as an alternative when other drugs fail. Other recently discovered active compounds against *M. tuberculosis* with potential for further development are the benzothiazinones, which are extremely efficient at blocking the growth of *M. tuberculosis* under *in vitro* conditions and are also active *in vivo* in mouse models. Benzothiazinones target DprE1, which is involved in the synthesis of arabinans, essential components of the mycobacterial cell wall. DprE1 seems to be a highly vulnerable target, as recently shown for a range of compounds with different chemical scaffolds. For example, DprE1 was found as the target of dinitrobenzamide derivatives (DNref-1), a class of compounds that were identified by high-content screening to interfere with the replication of *M. tuberculosis* within macrophages. The employed screening technique, which is based on automated,
simultaneous imaging of fluorescent bacteria and host macrophages exposed to different representatives of chemical compound libraries, is a particularly attractive and powerful method for identifying molecules that show activity against intracellular *M. tuberculosis* but little or no toxicity towards host cells. The same high-content screening approach was also the starting point for the discovery of Q203, a potent candidate drug that targets *qcrB*, encoding for a ubiquinol–cytochrome C reductase of the respiration chain of *M. tuberculosis*.

Another vulnerable drug target that is hit by different classes of small-molecule inhibitors is MmpL3, an essential mycobacterial membrane protein that harbors 11 transmembrane domains and is crucial for the transport of cell wall constituents. Probably the most advanced compound in terms of clinical development that was found to target MmpL3 is the diamine SQ109, a drug candidate that is in Phase II of clinical testing. Another molecule that is in clinical studies is PA-824, a nitroimidazole compound, which targets a deazaflavin-dependent nitroreductase.

Most recently, thiophene compounds were identified that kill *M. tuberculosis* by the previously uncharacterized mechanism of inhibition of the polyketide synthase Pks13, which is an essential enzyme that catalyzes the last condensation step of mycolic acid biosynthesis. In addition, a series of fluoroquinolones were studied in preclinical and clinical settings, with a particular focus on gatifloxacin and moxifloxacin, in order to evaluate whether they can improve the activity of the standard drug regimen when substituted for ethambutol and thereby might help to shorten the duration of treatment for fully drug-susceptible TB. As another example of the clinical evaluation of new drugs, the addition of linezolid into the regimen against extensively drug-resistant forms led to strongly improved prognosis of the treatment outcomes and culture conversion in patients with severe, chronic, extensively drug-resistant TB, which gives hope that other oxazolidinone-based drugs currently under development might also be effective.

Recent research has also elucidated the targets and putative modes of action of some anti-TB drugs that have been known for a long time. One striking example is the identification of the enoyl reductase InhA as the drug target of the natural compound pyridomycin, which bridges the NADH- and substrate-binding pockets of the enzyme and thus inhibits InhA in a different way than the first-line drug isoniazid, which is a prodrug and needs activation by KatG in order to form the active isonicotinic acyl–NADH complex. As a result, pyridomycin is also active on the most frequently encountered isoniazid-resistant *M. tuberculosis* strains, which opens new perspectives for drug development. Another old drug that targets InhA is ethionamide, a thiocarbamide-containing compound that is activated by the mycobacterial monooxygenase EthA, whose production is controlled by the transcriptional repressor EthR. A recently applied strategy that is based on inhibiting EthR shows promise for improving the therapeutic index of thiocarbamide derivatives and for overcoming resistance. Finally, for two other 'old' drugs, pyrazinamide and para-aminosalicylic acid, new mechanisms of action were recently identified. For pyrazinamide, activity directed against persistent forms of *M. tuberculosis* was discerned, which involves the inhibition of the process of trans-translation by ribosomal protein S1, which is essential for freeing scarce ribosomes in nonreplicating organisms. For para-aminosalicylic acid, it was found that it acted as a prodrug, which poisons folate-dependent pathways not only by serving as a replacement substrate for dihydropteroate synthase, but also by the generation of toxic byproducts created by the enzymes of these pathways.

---

### Mycobacterial Secretion Systems & Their Potential Utility as Drug Targets

Screening of different classes of compounds with antitubercular activity combined with sequence analysis and recombineering has revealed some new potential drug targets of *M. tuberculosis*, such as the membrane protein EccB3, which represents a novel candidate not targeted by currently existing anti-TB
The EccB₂ protein is a conserved component of the ESX-3 secretion system, which represents one of the five type VII secretion systems of *M. tuberculosis*. In contrast to ESX-1, which is not essential for *in vitro* growth of *M. tuberculosis* but is necessary for multiplication and survival of *M. tuberculosis* in the host cell, ESX-3 is a secretion system whose genes are essential for the *in vitro* growth of *M. tuberculosis* in standard culture media due to their role in mycobactin-mediated iron acquisition and zinc uptake. The aforementioned finding that the ESX-3 system can be targeted by small-compound inhibitors is an encouragement for further research and the development of inhibitors that might inactivate components of the secretion systems of *M. tuberculosis*. Apart from the ESX-3 system, among the type VII secretion systems, there is also the ESX-5 system, which has recently been shown to contain genes (e.g., *eccC₅*) that are essential for the *in vitro* growth of *M. tuberculosis* in addition to genes that are involved in the virulence and export of PE/PPE antigens (Figure 3).
Figure 3.
Working model of two selected secretion pathways (twin-arginine translocation and ESX) that are thought to be involved in the pathogenicity of *Mycobacterium tuberculosis*. (A) The genomic loci of two of the five ESX systems (ESX-1 and ESX-5), as well as the components of the TAT system. The gene nomenclature corresponds to that of the reference strain *Mycobacterium tuberculosis* H37Rv. (B) A model of the cell envelope of *M. tuberculosis* with the putative locations of selected proteins from the TAT and the ESX systems.

† The exact shape and mechanism by which the proteins are transported through the outer membrane is still unknown.

RR: Twin arginine; TAT: Twin-arginine translocation.

(A) Data taken from [11,60,64,67]; (B) Data taken from [60,67].

Besides the type VII secretion systems, which represent specialized secretion systems of mycobacteria and related actinobacteria, *M. tuberculosis* also uses various other secretion pathways for protein export (i.e., the SecA1-mediated general secretory pathway, the alternative SecA2-operated pathway and a twin-arginine translocation (TAT) system), which might also serve as targets for new drugs. The mycobacterial TAT system, for example, is constituted by a protein complex that includes TatA (Rv2094c) and TatC (Rv2093c), which are essential for the *in vitro* growth of *M. tuberculosis* and are thought to be localized in the plasma membrane (Figure 3). A third protein called TatB (Rv1224), which is encoded elsewhere in the genome, also seems to be essential for the function of the TAT apparatus in *M. tuberculosis*, although in TAT systems of other bacterial species, this does not always seem to be the case. Finally, based on sequence homology, a fourth protein, TatD (Rv1008), a nonessential protein with supposed DNase activity, is also a putative component of the TAT protein export system in *M. tuberculosis*. The proteins secreted by the TAT system are characterized by the presence of the twin arginine (RR) motif in the N-terminal signal sequence. In contrast to other secretion pathways of *M. tuberculosis*, the TAT system exports the proteins in their folded state. Among the various proteins with a TAT signal sequence motif, some have been described as being implicated in pathogenesis, such as the phospholipase C proteins (i.e., PlcA–D) and resistance to β-lactam antibiotics, such as BlaC, representing a class A β-lactamase.

Overall, it is clear that many of the proteins secreted by the various secretion pathways of *M. tuberculosis* play important roles in the survival and growth of the bacterium inside the host. Hence, specific protein export is a very important feature for pathogenic bacteria, as it enables the pathogen to withstand the defence mechanisms mounted by the host cells and/or exploit cellular functions for the benefit of the pathogen. Although the secreted proteins are often not essential for the *in vitro* growth of the organisms in culture broth, their *in vivo* growth essentiality might be exploited as an *in vivo* drug target for chemical compounds that could target these proteins during infection. As such, research on the potential virulence factors of pathogenic mycobacteria combined with elucidating their mechanisms of action might uncover new, unconventional targets for the development of new classes of antimycobacterial compounds that might act under *in vivo* conditions in synergy with more conventional bactericidal drugs.

---

**Selected Virulence Factors of *M. Tuberculosis* & Their Potential Role for Alternative Intervention Strategies**
As a result of the significant advances in mycobacterial molecular genetics and genomics in recent years, a number of *M. tuberculosis* genes that intervene in host–pathogen interactions have been identified. The first such mycobacterial virulence genes were identified by the use of pioneering gene knockout techniques\[76,77\] that were later combined with signature-tagged mutagenesis \[78,79\] resulting in a first list of candidate genes involved in the pathogenicity of *M. tuberculosis*. Among the genes identified by this approach were several with insertions in a 70-kb chromosomal segment that encodes proteins involved in the synthesis and transport of phthiocerol dimycocerosate, which corresponds to specific, extractable lipids in the outer layers of the mycobacterial cell envelope that play an important role in virulence. The number of potential virulence genes was later extended by the use of the TraSH technique adapted to mouse infection, which predicts that, in addition to the 15% of genes that are essential for *in vitro* growth, another 5% are involved in *in vivo* growth.\[30\] Among the approximately 200 listed genes are also several that encode proteins that belong to macromolecular components or secreted proteins of the aforementioned major secretion systems of *M. tuberculosis*, which might be particularly interesting as potential drug targets, as their encoded proteins are supposed to be localized in the cell envelope and thus more easily accessible by putative inhibitors. As a first example, the gene secA2, which codes for a main constituent of the alternative SecA2-operated pathway, is part of this list. This pathway was shown to export SodA and KatG, which are both involved in the detoxification of reactive oxygen intermediates produced by the host cell as part of the oxidative attack within the phagosome.\[69,80\] Small-molecule inhibitors that would inhibit these oxidative stress defense mechanisms of *M. tuberculosis* might strongly impact on the fitness and transmissibility of the pathogen, as suggested by virulence studies on KatG-deletion mutants in the context of isoniazid resistance.\[81,82\]

The list of mutants with *in vivo* growth defects also points to multiple genes implicated in the ESX-1 type VII secretion pathway.\[30\] The ESX-1 secretion system was identified as a primary factor determining the host–pathogen interactions of tubercle bacilli by several independent and complementary approaches (reviewed in\[61\]). Indeed, deletion of part of the ESX-1-encoding chromosomal region, named region of difference 1 (RD1), from *Mycobacterium bovis* BCG, which still represents the only anti-TB vaccine currently used on a global scale, is one of the main reasons for the loss of virulence of this attenuated live vaccine. Similarly, the ESX-1 system is also truncated in *Mycobacterium microti* strains, of which several were used as live-attenuated vaccines in the 1960s.\[81\] Moreover, the presence of the 6-kDa early secreted antigenic target ESAT-6 and its protein partner CFP-10, which are both secreted by ESX-1 in *M. tuberculosis*, but are absent from *M. bovis* BCG and *M. microti*-based vaccines, represents the basis for the differential potential of the next-generation IFN-γ release assays that have substantially refined TB diagnosis.\[83\]

ESX-1 might thus serve as a potential target for the development of molecules that could interfere with the process of ESX-1-mediated host–pathogen interactions. Despite intense research within the last decade since its discovery, the exact function of the ESX-1 system is only partially known. However, recent pathogenicity-related research has shown that the intracellular behavior of wild-type *M. tuberculosis* is profoundly different from ESX-1-deficient variants of *M. tuberculosis* (ΔRD1) or other ESX-1-deleted tubercle bacilli, such as BCG. The main differences observed during the infection of macrophages are linked to the finding that ESX-1-proficient strains are able to rupture the phagosomal membrane and obtain access to the host cell cytosol at later stages of the infection, whereas strains carrying a partial or interrupted ESX-1 system remain locked in the phagovacuole.\[84,85\] Hence, the ability of *M. tuberculosis* to break the phagosomal membrane and access the cytosol seems to be tightly linked to secretion ESAT-6 and CFP-10. The observed activity seems to be specific for ESAT-6 of *M. tuberculosis*[86,87] that, under acidic pH conditions, may undergo significant conformational changes, which is not the case for orthologous proteins from the nonpathogenic *M. smegmatis* species.\[88\] The ESX-1-
dependent ability of inducing phagosomal breakage seems to be a key factor during the infection of host cells by *M. tuberculosis*, as only strains with intact ESX-1 secretion can cause cell death and have enhanced cell-to-cell spread. The cytosolic contact of ESX-1-proficient strains also determines the autophagic flux and may promote cross-presentation of mycobacterial antigens by the cytosolic MHC class I-processing machinery (i.e., the proteasome). The differences in cytosolic contact and induction of host cell necrosis between ESX-1-containing *M. tuberculosis* and naturally ESX-1-deleted BCG are also relevant for NLRP3 inflammasome activation, which leads to IL-1β secretion, the generation of type I interferons and impacts induction of CD8+ T-cell responses. The identification of small-molecule inhibitors specifically targeting the ESX-1 secretion machinery or its secreted proteins may represent a suitable alternative way to neutralize the virulence factors of *M. tuberculosis* Without being bactericidal themselves, these virulence inhibitors could accentuate the innate and adaptive immune responses of the host on bacteria that have become devoid of their ESX-1-linked intracellular defense strategies. Such a strategy, which might work best in combination with conventional, bactericidal, anti-TB drugs, could also help with avoiding the generation of drug resistance due to the action of the immune system on potential escape mutants.

Inhibition of the ESX-1 system might also be achieved by targeting the regulation process of ESAT-6/CFP10 secretion, which is linked to the ESX-1-associated proteins EspA, EspC and EspD. The *espACD* locus, located in a different region of the *M. tuberculosis* genome than the ESX-1 core components, is important for ESX-1 functions because EspA, EspC and EspD are secreted in an ESAT-6/CFP-10-codependent fashion. The expression of EspACD is regulated by a mechanism that involves the virulence regulator EspR, which was recently described as a nucleoid-associated protein, and the two-component regulators MprAB and PhoPR. PhoPR is well known for its implication in the virulence of *M. tuberculosis* and its large regulon that accounts for more than 40 positively and 70 negatively regulated genes. Although the exact regulation cascade of PhoPR, MprAB and EspR in association with EspACD remains unclear, the strong ESX-1-mediated implication of the *espACD* operon in the virulence of *M. tuberculosis* makes these regulators potential drug targets.

In analogy with the ESX-1 system, the ESX-5 secretion system might also represent a potential drug target, as it is simultaneously involved with the *in vitro* viability and *in vivo* replication of *M. tuberculosis*. It was recently shown that disruption of certain ESX-5 core components, such as the predicted transmembrane channel protein EccD, or the membrane-bound ATPase EccC, affects the growth of *M. tuberculosis* on solid medium, yielding small colony morphotypes of mutant colonies in comparison with the wild-type strain. ESX-5 was also shown to be essential for mycobacterial cell wall integrity, which is further confirmed by the enhanced detergent and hydrophilic antibiotic sensitivities of ESX-5-knockout strains. The function of the ESX-5 system in *M. tuberculosis* appears to be mainly linked to the secretion of Esx and PE/PPE proteins, with the latter corresponding to two large protein families in *M. tuberculosis* that are named and classified after a characteristic Pro–Glu or Pro–Glu–Glu motif at the N-terminus class, as well as amino acid sequence similarity. The functions of these immunogenic proteins remain largely unknown, but it is clear now that some of them play a role in pathogenicity. For example, upon the deletion of the genomic section encoding PPE25-PE19 within the ESX-5 locus, the virulence of the resulting *M. tuberculosis* mutant was strongly diminished in immunocompetent and immunodeficient mice, whereas the complemented strain regained virulence. Some of these proteins contain potent T-cell epitopes with or without cross-reactivity with other PE/PPE proteins and might thus show good potential for inclusion into subunit vaccine combinations that might be used in combination with specific drug therapies as a form of immunotherapy, which is analogous to what has been tested in other bacterial species.
Conclusion

Mycobacteria are very widespread in different environments, including soil and aqueous microbiota, where contact with other microorganisms is dense. From an evolutionary perspective, it seems plausible that factors that contribute today to mycobacterial pathogenesis might have evolved during the long interplay of mycobacteria with other microorganisms, including protozoa. Today's virulence factors might constitute ancestral invasion and defense mechanisms that were transformed by pathogenic mycobacteria into tools that are now used for intracellular survival in cells of vertebrate hosts. These genes involved in pathogenicity, together with the core genes involved in the essential functions of mycobacteria, may constitute promising new targets for the development of specific antimycobacterial molecules and compounds that can act in synergy with other available strategies. Besides the treatment regimens already in place or in development, high-content screening approaches of compounds and/or cellular particularities\[43,114\] that can simultaneously evaluate the viability of host phagocytes and the bacterial growth characteristics thus show great promise for the identification of appropriate candidates. Comparative genome information from various mycobacterial species and insights into the evolutionary traits of the concerned genes and gene products will be of further help for validation purposes.

Future Perspective

In order to predict future trends, it might be useful to look back some years and evaluate whether the current situation of research and development could have been predicted. Research within the last 20 years since TB was declared a global emergency by the WHO in 1993\[26\] has contributed to the enormous gain of knowledge regarding the causative agent of TB and its interactions with its host. Apart from the fact that some of these recent advances of mycobacterial research and technology have already been translated into applications with direct impact on diagnosis (e.g., IFN-\(\gamma\) release assays) or treatment (e.g., bedaquiline), the situation concerning TB research has profoundly changed in recent years concerning our knowledge regarding the evolution, genomics and genetics of \(M.\) \(tuberculosis\), the disease-causing processes related to host–pathogen interactions and the discovery of new molecules, immunogens and drug targets. For example, today it is very difficult to imagine biological and medical research without access to genome data and all of the data produced by follow-up strategies, such as genome-wide transposon insertion screens, transcriptome analyses and structural biology approaches, among others. The broad impact of the genome sequence for the mycobacterial research community is reflected in the fact that the publication presenting the first complete genome sequence of \(M.\) \(tuberculosis\) (H37Rv) \[11\] represents the most cited article in the field of TB research at more than 4000 citations. The use of genome sequencing has been refined in recent years, as NGS approaches have allowed the cost of genome sequencing to be dramatically reduced. The reduced costs and technical advances opened the NGS technologies up for use in the sequencing of the genomes of many different \(M.\) \(tuberculosis\) isolates, thereby allowing the phylogenetic reconstruction of the global \(M.\) \(tuberculosis\) strains to be refined.\[28\] NGS also showed that epidemiological tracing of \(M.\) \(tuberculosis\)-caused outbreaks could be carried out in much greater detail than with the previously used molecular typing methods,\[115,116\] a finding that could revolutionize epidemiological studies of \(M.\) \(tuberculosis\) outbreaks in the near future.

Recent research on virulence factors and the fate of \(M.\) \(tuberculosis\) in the phagosome has elucidated many of the key features of mycobacterial host–pathogen interactions. While \(M.\) \(tuberculosis\) was considered until very recently to be a pathogen that, during the infection of phagocytic cells, resides exclusively in the phagovacuole, technical advances in cryoelectron microscopy and single-cell assays based on fluorescence imaging suggested that \(M.\) \(tuberculosis\) breaks the phagosomal membrane and accesses the cytosol of the host cell at later stages of infection.\[84,85,89\] Future research will certainly further
refine these observations and link them to the survival strategy of *M. tuberculosis*, with central importance being given for the induction of host-cell death and bacterial spread. Future therapeutic strategies that would prevent the rupture of *M. tuberculosis*-containing phagovacuoles could have a strong impact on the outcome of infection. In order to identify new molecules with such activity, phenotypic cell-based assays, which use automated confocal fluorescence microscopy for the high-throughput screening of chemicals that interfere with the replication of *M. tuberculosis* within macrophages, could be used. An adaptation of this method, which was successfully employed to identify acyl-trehalose-containing glycolipids involved in phagosome remodeling might well be used for finding factors that might interfere with the induction of phagosomal rupture and related cellular processes. As mentioned above, this method has the great advantage of simultaneously observing the behavior of the bacterial invader and the defense strategies of the host cell, which is thus a very promising technology for future applications.

Looking at the advances in biomedical research of the last few years, which would have been difficult to foresee 20 years ago, the outlook for the future may only be very vague. In order to cope with the global problem of TB, combination strategies might be most advantageous. The trend in human medicine is to individualize treatment, and this might also apply for the prevention and treatment of TB cases. TB is a highly complex disease, with the complexity arising from the varied immunological and nutritional status of the individual, the type of *M. tuberculosis* strains circulating in the community, the presence of drug resistance and the social environment, among other factors. In this context, it might be best to combine treatment with immunotherapies (i.e., immunization strategies that would add an additional layer of immune defense to the action of conventional or novel anti-TB drugs). However, care has to be taken in order to avoid inducing a so-called Koch phenomenon, which corresponds to the development of necrotic lesions as a result of hypersensitivity to products of the tubercle bacillus. In addition, the use of molecules that would act against the pathogenicity factors of *M. tuberculosis* might be a very useful addition to such combinatorial treatment strategies. The last 20 years of TB research have unravelled numerous new details regarding the pathogen and its interactions with the host that are beginning to find practical applications for patients. Apart from increasing the social standards of the populations that are most affected by high infection rates with *M. tuberculosis*, it is to be hoped that the next 20 years of TB research will reveal even more secrets of the TB pathogen that can be used in the continued fight against TB. We should not make the mistake of the 1970s again, which would be to underestimate the adaptability of this pathogen, whose evolution is directly linked with human evolution. High-quality TB research will have to continue to cope with today’s situation and find novel innovative solutions for future trends.

Sidebar
Executive Summary

**Evolution of mycobacterial pathogenicity**

- *Mycobacterium tuberculosis* is characterized by a cell envelope that contains an inner membrane, an outer membrane (mycomebrane) and a capsule. Despite being categorized as a high-GC-content, Gram-positive bacterium, the mycobacterial cell wall functionally resembles that of Gram-negative bacteria.

- *M. tuberculosis* shares a core genome of 2500 genes with nontuberculous mycobacteria. More than 600 genes seem to be specific to the tubercle bacillus.
Comparison of *M. tuberculosis* with *Mycobacterium canettii*, known as the nonclonal tubercle bacilli, shows that *M. tuberculosis* has gained some genes after its evolutionary branching.

**Potential roles of essential genes as drug targets**

- High-density transposon analysis showed that approximately 15% of the genes of *M. tuberculosis* are essential for optimal growth in broth culture. The gene products of these essential genes represent potential drug targets.
- Several of these genes have been recently identified by phenotypic compound screening and target identification as vulnerable targets of *M. tuberculosis*.

**Mycobacterial secretion systems & their potential utility as drug targets**

- *M. tuberculosis* encodes major secretion systems that allow the transport of biomolecules across its complex cell envelope.
- Some of these systems, such as the SecA1, the twin-arginine translocation system and the ESX-3 and ESX-5 systems, are essential for *in vitro* growth. The SecA2, ESX-1 and ESX-5 systems were shown to be required for the pathogenicity of *M. tuberculosis* and represent potential virulence drug targets.

**Selected virulence factors of *M. tuberculosis* & their potential roles in alternative intervention strategies**

- Analogously to the genes that are essential for *in vitro* growth, approximately 5% of *M. tuberculosis* genes were identified as being essential for *in vivo* growth (survival in the mouse model) by saturation transposon screening.
- In addition, several genes that code for proteins involved in pathogenicity have been predicted by comparative genomics of virulent and attenuated tubercle bacilli and gene-knockout studies.

**Conclusion & future perspective**

- Predictions for the next 20 years, based on the advances of mycobacterial research achieved within the last two decades, promise we might be able to identify the Achilles' heel of *M. tuberculosis* and thereby allow new treatment strategies to be implemented.
- Targeting virulence genes could become a strong alternative strategy of treatment that relies on the focused action of the immune system.
- It is probable that treatments against the pathogenicity factors of *M. tuberculosis* might have to be combined with drugs that act against essential targets.

**References**


* Reviews the evolutionary processes that could have led to the evolution of pathogenicity of slow-growing mycobacteria, including *Mycobacterium tuberculosis*.


* Describes the different components of the mycobacterial cell wall, taking into account the advances of numerous research projects.


* Reports and discusses the ancestral gene pool of tubercle bacilli (Mycobacterium canettii strains) from which M. tuberculosis has emerged and pinpoints genes that seem to have been acquired by M. tuberculosis through horizontal transfer. It remains to be determined whether these genes have contributed to the evolutionary success of M. tuberculosis as a very widespread key pathogen.


* Key study that refines the previous transposon site hybridization-based analyses to the level of essential regions of proteins.


* Landmark study presenting the first entirely new anti-TB compound after many years of quiescence, which is now in clinical use for treatment against multiresistant TB.


* Seminal study presenting new, highly active compounds and a vulnerable target that had not been found before. Compounds of the same chemical scaffold are being developed for clinical use.


* Describes the clinical use of a relatively new drug in conjunction with other treatment regimens and the life-saving treatment advantages gained for patients.


* Very carefully executed study on the mechanisms of action of an almost-forgotten anti-TB drug.


* Presents a list of new potential drug targets that have recently been identified.


* Important study that evaluated the impact of the ESX-5 secretion system of *M. tuberculosis* on protein transport and pathogenicity.

* Important study that identifies the immunogenic epitopes of PE and PPE proteins within and outside of the ESX-5 secretion system.


* Landmark study that revises a long-kept scheme of mycobacterial pathogenicity.


* Important study that confirms and largely extends the hypothesis of phagosomal breakage due to ESX-1-proficient *M. tuberculosis* and recombinant BCG strains.


* Important study that reports the cell biological consequences of ESX-1-mediated cytosolic contact of *M. tuberculosis*.


* Important study that shows the consequences of mycobacterial DNA release into the cytosol.


* Very interesting study that describes the mechanism of the virulence regulator EspR, which regulates the secretion of several important antigens and virulence factors of *M. tuberculosis*.


* Highly interesting study that focuses on the copurification of core proteins of type VII secretion machineries.


* Highly interesting study that shows how high-throughput cell-based screening may be used to identify potential pathogenicity factors. The same technology might be useful for identifying compounds that can interfere with important host–pathogen interaction steps.


**Financial & competing interests disclosure**

We acknowledge the support by the Fondation pour la Recherche Médicale FRM grant no. DEQ20130326471, the European Community’s (EC) Framework Programme 7 grant MM4TB 260872 and grant 09 BLAN 0400 01 from the Agence Nationale pour la Recherche (France). F Le Chevalier is supported by the French Region Ile-de-France (Domaine d'Intérêt Majeur Maladies Infectieuses et Emergentes) PhD program. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.