Seminars in Immunology 39 (2018) 88–101

Contents lists available at ScienceDirect

Seminars in Immunology

journal homepage: www.elsevier.com/locate/ysmim

Genome wide approaches discover novel *Mycobacterium tuberculosis* antigens as correlates of infection, disease, immunity and targets for vaccination

Mariateresa Coppola*, Tom HM Ottenhoff

Dept. Infectious Diseases, LUMC, PO Box 9600, 2300RC Leiden, The Netherlands

A B S T R A C T

Every day approximately six thousand people die of Tuberculosis (TB). Its causative agent, *Mycobacterium tuberculosis* (*Mtb*), is an ancient pathogen that through its evolution developed complex mechanisms to evade immune surveillance and acquire the ability to establish persistent infection in its hosts. Currently, it is estimated that one-fourth of the human population is latently infected with *Mtb* and among those infected 3–10% are at risk of developing active TB disease during their lifetime. The currently available diagnostics are not able to detect this risk group for prophylactic treatment to prevent transmission. Anti-TB drugs are available but only as long regimens with considerable side effects, which could both be reduced if adequate tests were available to monitor the response of TB to treatment. New vaccines are also urgently needed to substitute or boost Bacille Calmette-Guérin (BCG), the only approved TB vaccine: although BCG prevents disseminated TB in infants, it fails to impact the incidence of pulmonary TB in adults, and therefore has little effect on TB transmission. To achieve TB eradication, the discovery of *Mtb* antigens that effectively correlate with the human response to infection, with the curative host response following TB treatment, and with natural as well as vaccine induced protection will be critical.

Over the last decade, many new *Mtb* antigens have been found and proposed as TB biomarkers and vaccine candidates, but only a very small number of these is being used in commercial diagnostic tests or is being assessed as candidate TB vaccine antigens in human clinical trials, aiming to prevent infection, disease or disease recurrence following treatment. Most of these antigens were discovered decades ago, before the complete *Mtb* genome sequence became available, and thus did not harness the latest insights from post-genomic antigen discovery strategies and genome wide approaches. These have, for example, revealed critical phase variation in *Mtb* replication and accompanying gene –and therefore antigen– expression patterns. In this review, we present a brief overview of past methodologies, and subsequently focus on the most important recent *Mtb* antigen discovery studies which have mined the *Mtb* antigenome through “unbiased” genome wide approaches. We compare the results for these approaches -as far as we know for the first time-, highlight *Mtb* antigens that have been identified independently by different strategies and present a comprehensive overview of the *Mtb* antigens thus discovered.

1. Introduction: classical antigens, latency antigens, rpfs, and vaccine potential

Tuberculosis (TB)is an ancient infectious disease [1] that has killed approximately one billion people during the last two centuries [2]. Every year more than 10 million people develop active TB disease, and almost 2 million die from TB, also known as the white plague. Despite the fact that global TB control efforts avert millions of cases and hundred thousands of deaths every year, TB remains the leading cause of mortality from a single infectious agent worldwide [3]. The protective efficacy of Bacille Calmette-Guérin (BCG), the only licensed TB vaccine, is variable and inconsistent, ranging from 0 to 80%, and even though it is effective in reducing the incidence of disseminated TB in children, it is clearly insufficient in preventing the onset of pulmonary TB in adolescents and adults [4]. The current diagnostics for *Mycobacterium tuberculosis* (*Mtb*) infection, the century old tuberculin skin test (TST) and the more recent interferon-gamma release assays (IGRAs), are indirect measures of infection as these tests determine previous host sensitization to *Mtb* antigens by detecting memory T cell responses. Moreover their diagnostic performance in HIV/TB-co-infected patients is poor when CD4+ T cell counts are diminished. These tests also cannot distinguish between active TB and latent TB infection (LTBI), and have

* Corresponding author.
E-mail address: m.coppola@lumc.nl (M. Coppola).

https://doi.org/10.1016/j.smim.2018.07.001
Received 2 July 2018; Accepted 2 July 2018
Available online 07 July 2018

virtually no predictive value in estimating the risk of progression from latent to active state of disease, which typically happens in 3–10% of those infected [5]. In the past decades, development of better diagnostics and vaccines has been a key goal in TB research, yet thwarted by the incomplete understanding of the complex and heterogeneous interplay between the human host and *Mtb* [6].

In order to provide better insights into this intricate host-pathogen interaction and to discover new *Mtb* antigens as vaccine candidates and TB biomarkers, searches for novel and relevant *Mtb* antigens have been intensified during recent years, particularly following the application of genome wide approaches. In this paper, after a short review of past methodologies, we summarize and discuss several important recent *Mtb* antigen discovery studies that have harnessed the unprecedented opportunities of "unbiased" genome wide approaches. We compare their results, discuss those *Mtb* antigens that were discovered independently by different approaches, and present a comprehensive overview of *Mtb* protein and peptide antigens.

1.1. Traditional Mtb antigen discovery approaches in the last century

Traditional *Mtb* antigen discovery approaches in the pre-genomics era relied on classical biochemical analyses of *in vitro* cultured *Mtb* from which fractions and secreted moieties were isolated and purified [7], based on much excellent work from many researchers. The antigenicity of the isolated components was typically tested using cells or sera from *Mtb* infected animals, including mice, rabbits, goats and cattle, as well as human beings exposed to mycobacteria such as LTBI, BCG vaccinees and active or treated TB patients. After the discovery of IFN-γ as an important cytokine in protective immunity to TB in animal models and humans [8,9], and the development of suitable assays to measure it, the antigenicity of *Mtb* antigens started to be determined increasingly by the ability to induce IFN-γ in mouse models and in blood cells of *Mtb* exposed individuals *in vitro*, assuming that IFN-γ could be a correlate of protection. These approaches in the 1980′s and 1990′s led to the identification of important *Mtb* antigens abundantly expressed *in vitro*, including Ag85C (fibronectin-binding protein C, Rv0129c), Ag85B (Rv1886), Ag85A (Rv3804c) [10–12], Mpb64 (Rv1980) [13], 38kDa (Rv0934) [14], ESAT-6 (Rv3875) [15], Mtb8.4 (Rv1174) [16], CFP-10 (Rv3874) [17], Mpt51 (Rv3803c) [18] and TB10.4 (Rv0288) [19]. Virtually all of these proteins were secreted *Mtb* proteins, and therefore amenable to identification in the supernatants of *in vitro* grown *Mtb*. In addition, also heat shock proteins became a prominent focus of attention as will be discussed below.

The completion of the sequencing and annotation of the *Mtb* (H37Rv) genome in 1998 [20] in the first decade of this century was key to the initiation of the first *Mtb* post-genomic approaches, and allowed the discovery of new classes of *Mtb* antigens using unbiased approaches. One example of such approaches is the discovery of so called *Mtb* "latency antigens", based on genome-wide expression profiling of *Mtb* bacteria *in vitro* exposed to hypoxic culture conditions.

1.2. Mtb latency antigens are important tools for designing new, and for improving current TB vaccines

By using *in vitro* conditions that mimic the host environment encountered by *Mtb* during *in vivo* infection, such as hypoxia, nutrient starvation or IFN-γ-activated macrophages, a genetically regulated metabolic shift down was discovered in *Mtb* bacilli exposed to such conditions. *Mtb* down regulated most of its genes as an adaptive response to stress, but this app eared to be accompanied by a remarkable upregulation of the expression of 48Mtbgenes in response to various environmental stress factors including hypoxia. These genes were found to be regulated by a response regulator termed the DosR protein (encoded by Rv3133c) and hence this regulon was called the dormancy or DosR regulon [21,22]. One of the genes most strongly upregulated was Rv2031c/α-crystallin, which is involved in cell wall thickening and stabilisation under hypoxia (further discussed below). The DosR regulon encoded gene products were designated "latency antigens" and were tested as recombinant proteins for cellular and humoral recognition in LTBI individuals and TB patients from different geographical cohorts in Europe, Africa, South America and India. In all cohorts, the *Mtb* latency antigens were recognized preferentially by LTBI donors compared to (active or treated) TB patients, in terms of both cellular and immunoglobulin responses [23–28]. The data suggested that several of these antigens such as Rv1733c, Rv2029c, Rv2626 and R2628 could be useful antigenic targets to discriminate LTBI from TB patients, which is not well possible by IGRA or TST (Supplementary Table 1). Moreover, the fact that the BCG genome contains homologues of the DosR genes [29] suggested that latency antigens might be interesting TB booster vaccine candidates, particularly since they have been demonstrated in several studies to have protective efficacy in different animal models [30–34]. However, a puzzling observation was that neither subcutaneous (sc) BCG vaccination in animals, nor intradermal (id) BCG vaccination in humans was found to induce T cell responses against latency antigens, suggesting that these routes of immunisation might not allow BCG to enter a state of latency [29,35].

We have hypothesized that the inability of BCG to induce immune responses to latency antigens may underlie its impaired ability to induce full protective immunity to *Mtb*. We have submitted the hypothesis that this should be repaired, either by recombinant overexpression or by subunit vaccine boosting, in order to induce improved protection. Our results [36] are in support of this concept, since an Rv1733c, Rv3407 and Rv2659 (representing latency and starvation antigens) expressing recombinant BCG strain (rBCGΔureC::hly) induced better protection against highly virulent Mtb(Beijing) in mice than did the regular recombinant BCG (rBCGΔureC::hly) lacking these antigens. In addition, our more recent data have demonstrated that Rv1733c vaccination following BCG significantly improved the protective efficacy of BCG [33].

Another important latency antigen discovered around the same time was the 28kDa heparin-binding protein (HBHA, Rv0475), an antigen not regulated by DosR. The HBHA adhesin, identified using heparine-Sepharose chromatography from culture supernatant and extracts of *Mtb* and M. bovis [37], is expressed on the surface of Mtband promotes its interaction with non-phagocytic cells, thereby facilitating *Mtb* extrapulmonary dissemination [38]. It has been shown that when coated with anti-HBHA antibodies *Mtb* has a reduced ability to disseminate outside the lung in mice, suggesting that anti-HBHA antibodies in TB patients might help containing *Mtb* infection by blocking HBHA [38]. Interestingly, stimulation of lymphocytes from LTBI but not TB patients with HBHA induced IFN-γ secretion [39,40] and perforin-producing CD8+ T cells cytotoxic against mycobacterium infected macrophages [41]. Based on these findings, HBHA induced T cell IFN-γ production has been proposed as a biomarker of LTBI [42]. In addition, HBHA has been shown to enhance BCG protective efficacy both in adult and newborn mice when administrated in prime-boost regimens [43,44]. These data, supported by the evidence that BCG induces specific HBHA multi cytokine responses [45], suggest that HBHA might be part of a promising TB subunit vaccine candidate to boost BCG.

In conclusion, latency antigens (DosR regulon encoded antigens, HBHA) are highly potent novel *Mtb* antigens, which could find application in the design of new TB vaccines (see also below) as well as in tests diagnosing LTBI.

1.3. Mtb resuscitation-promoting factors (Rpf) are antigens that are promising targets for diagnosis and vaccination

Next to latency antigens another class of *Mtb* antigens discovered were *Mtb's* Rpfs [46,47]. Rpfs are secreted bacterial proteins that have hormone like activity, and are able to promote the transition from a dormant into an active e replicating state of bacteria, including mycobacteria [48]. The five *Mtb* Rpfs, specifically RpfA (Rv0867c), RpfB

(Rv1009), RpfC (Rv1884c), RpfD (Rv2389c), and RpfE (Rv2450c) were screened for immunological reactivity in *Mtb* exposed individuals and found to be recognized *in vitro* by IFN-γ producing cells preferentially from LTBI [28,49,50]. Based on this property, these proteins could further help differentiating LTBI from TB, next to the above latency antigens.

1.4. New developments for TB vaccines with latency and Rpf antigens in NHP and mouse models

From the TB vaccine perspective, several of the above discussed *Mtb* antigens (Rpfs, latency antigens and early secreted proteins) have been evaluated recently as combinatorial, multistage vaccines in mice and rhesus macaques. Vaccination with a virally (*i.e.* modified RhCMV) vectored combination of 6 or 9Mtbantigens in Rhesus macaques strongly reduced mycobacterial loads in the lung of the *Mtb* infected animals, even to the unprecedented extent of sterile eradication in half of the animals [51–53]. The RhCMV vectored insert contained 6 or 9Mtbantigens, namely: Ag85A, Ag85B, ESAT-6, Rv1733c, Rv2626c, Rv3407, RpfA, RpfC, and RpfD [53]. Additionally, 5 of these antigens have been evaluated recently as part of another multi-antigenic vectored vaccine, MVATG18598, in this case in post-exposure BALB/c and C57BL/6 mouse models [54]. MVATG18598 specifically expressed Rv2626c, Ag85B, CFP-10, ESAT-6, TB10.4, Rv0287, RpfB, RpfD, Rv3407, and Rv1813c. Mice were challenged with *Mtb*, treated with a standard antibiotic regimen (rifampin, isoniazid and pyrazinamide) and then "therapeutically" vaccinated. Two different vaccination regimens (*i.e.* during and after chemotherapy), two delivery routes (s.c. or i.n.) and different numbers of injections (1x, 3x, 5x, and 7x) were tested to evaluate the effect of MVATG18598 in this post-challenge model. The s.c. administration (3x) of MVATG18598 during treatment showed the strongest reduction in the occurrence of post treatment TB relapses and *Mtb* burden in BALB/c mice. This finding was confirmed in C57BL/6 mice (in which only 1x injection of the vaccine was tested). In both mouse strains, the MVATG18598 vaccination triggered strong IFN-γ production and high levels of polyclonal antibodies against specific components of the multi-antigenic vaccine.

These preclinical TB vaccination data from NHP and mouse confirm and extend the unexpectedly strong vaccine potential of the newly discovered latency, early secreted and Rpf antigens, and hold promise for application in humans in the nearby future, in both preventive and therapeutic fashion.

1.5. Results from latest clinical trials with new TB subunit antigen based vaccines and BCG. Implications for developing new vaccines

To date, there are only three clinical studies that have reported human efficacy data following vaccination with *Mtb* antigen-based subunit vaccines. One was from a large phase IIb trial conducted with MVA85A expressing the early secreted protein Ag85A. Unfortunately, this candidate failed to induce additional protective efficacy against developing TB (in a PoD (prevention of disease) trial design) following initial standard BCG vaccination in children [55]. In a more recent, prevention of infection (PoI) clinical study in the same area, H4 (a fusion protein of secreted antigens Ag85B/TB10.4) was administered in Th1 inducing IC31 adjuvant to previously BCG vaccinated adolescents. Although H4:IC31 failed to prevent initial or sustained *Mtb* infection significantly, a clear trend towards reduced sustained infection (defined as three consecutive positive IGRA tests) was distinguishable in this PoI trial, providing a first "signal" for a TB subunit vaccine. Unexpectedly, and encouragingly however, in this study BCG revaccination gave a significant reduction in sustained infection as determined by three consecutive positive QFG/IGRA tests [56]. This could be interpreted as suggesting that immunity can be induced by vaccination that can control or even eliminate *Mtb*, and calls for follow up studies.

Both MVA85A and H4:IC31 vaccines were designed principally to boost prior BCG induced CD4+ Th1 immunity. The lack of significant efficacy against the primary clinical endpoints could be interpreted as indicating that, while CD4+ Th1 cells are clearly essential they may not be sufficient to induce adequate protection. This agrees with the finding that frequencies of IFN-γ-expressing cells did not correlate with BCG induced protection in a South African population of infants [57]. Alternative explanations obviously are possible but a detailed discussion of that topic falls outside the scope of this review.

Finally, the third and most recent clinical data from a TB subunit PoD trial come from the M72 (GSK) trial, which administered a fusion protein called M72 (consisting of rearranged antigen fragments from Rv1196 and Rv0125) adjuvanted in the Th1 promoting AS01E. The results of this seminal study will become available mid-2018.

Thus, despite some initially disappointing results there is hope for new and properly delivered TB subunit vaccines. Also clear signals from live mycobacterial "whole cell" vaccines are emerging, including recombinant BCG strains [58] and attenuated *Mtb* strains (no clinical efficacy data are yet available, although the MTBVAC vaccine was reportedly safe) [59]. This will not be discussed further here.

2. Broadening the discovery of the *Mtb* antigenome using genome wide approaches

Advances in DNA technologies, genome sequencing and bioinformatics tools now offer unprecedented opportunities to rapidly and exhaustively mine the potential *Mtb* antigenome, at least at the proteome level. Here, we will focus chiefly on the most representative recent studies and discuss their results.

The first whole genome based *Mtb* antigenome search was innovative in that it generated the first *Mtb* genome expression libraries using newly developed recombinant DNA technologies. This *Mtb* genome expression library was built by shearing *Mtb* Erdman strain's DNA, that was subsequently linked to EcoRI linkers and inserted into λgt11 vectors, allowing access to *Escherichia coli's* transcriptional and translational machinery to produce *Mtb* protein fragments [60]. The potential to detect "all" (note: the library almost certainly was incomplete) *Mtb* proteins, including those not always expressed by *Mtb* depending on its phase, constituted an important novelty. By screening this library with murine monoclonal antibodies [61–63] and polyclonal sera [64,65] several new *Mtb* proteins such as HspX/ α-crystallin (Rv2031c) were identified and further validated using lymphocytes or T cell clones from TB patients [66]. In addition, also other heat shock proteins such as hsp60 (Rv0440) and hsp70 (Rv0350) [62,67,68] were extensively studied, in part due to their high immunogenicity and in part due to their sequence similarities with mammalian hsp, raising the interesting possibility of being involved in inducing cross-reactivity and the subsequent precipitation of auto-immunity [69–71].

Expression cloning technologies used in subsequent studies, including those employing DNA from virulent *Mtb* strains [72], resulted in the discovery of additional immunogenic proteins comprising Rv1510 [72], Mtb39a (Rv1196) [73], Mtb32a (Rv0125) [74], Mtb9.9a (Rv1793c) [75], Mtb9.8 (Rv0287) [76] and Mtb41 (Rv0915c) (Supplementary Table 1) [77–82]. Obviously, these methods are biased towards protein antigens and do not offer the possibility to evaluate non-protein antigens like polysaccharides, lipopolysaccharides, and glycolipids. In addition, the gene products synthesized by recombinant expression in heterologous hosts does often not replicate post-translational modifications such as lipidation and glycosylation, which are naturally occurring in mycobacteria. Whether and how those changes impact protein immunogenicity has been studied to some extent, and it was recently shown that glycoconjugate derivatives of Ag85B, obtained after glycosylation of its lysine residues, were less recognised by BCG vaccinees and TB patients compared to the unmodified variant [83]. This finding suggests that recombinant products lacking post-translational modifications could lead to misinterpret the immunological activity of antigens compared to the variants present during natural *Mtb*

infection. The use of other vectors phylogenetically closer to *Mtb*, such as M. smegmatishas been proposed as an alternative host to overcome such post-translational issues [84].

Although allowing the identification of the *Mtb* antigens Rv1196 and Rv0125 which are now included in an advanced TB vaccine candidate, M72/AS01E (see above) [85, 86], the above discussed *Mtb* genome expression libraries have nevertheless unlocked only a small portion of the *Mtb* antigenome (Supplementary Table1). This is probably due to issues of low expression in heterologous systems and the use of small DNA fragments, precluding the expression of conformationally more complex structures such as antibody epitopes.

This scenario was revolutionized by the completion of the *Mtb* genome sequence that now permitted access to the entire *Mtb* protein antigenome [20]. The *Mtb* genome sequence made it possible to identify novel candidate antigens without the need for, and limitations of having to cultivate *Mtb* [7]. A breakthrough was also that all genome information was made available *via* open-source internet genomic and proteomic databases such as TubercuList (http://svitsrv8.epfl.ch/tuberculist/) and later TB Database (www.tbdb.org) [87]. In addition, the increasing accessibility of highly powerful bioinformatics tools now also allowedin silicogenomic analyses and comparative evaluation (*e.g.*: BLAST, https://blast.ncbi.nlm.nih.gov/Blast.cgi), and to predict new *Mtb* T cell epitopes based on predictive HLA binding motifs (*e.g.*: ProPred, http://crdd.osdd.net/raghava/propred/, HLA_BIND, https://www-bimas.cit.nih.gov/molbio/hla_bind/, EpiMatrix, http://i-cubed.org/tools/ivax/ivax-tool-kit/epimatrix/, DTU prediction servers, http://www.cbs.dtu.dk/services/). These developments and technologies have revolutionized the field of antigen discovery in general, and that of *Mtb* in particular (*Mtb* is a complex pathogen to work with due to its slow growth, its phase variation and its restriction to BSL3 safety level laboratories).

3. Mining the complete *Mtb* antigenome for epitope and antigen discovery

3.1. In search of new Mtb Epitopes using peptide libraries

Experimental methods to identify new *Mtb* (peptide) epitopes have been classically based on the screening of overlapping peptides spanning the full sequence of the relevant target proteins, followed by assaying antibody and/or T cell immune responses in *in vitro* assays and/or *in vivo* models [88–94]. This approach has also been applied to interrogate novel interesting candidate antigens selected from the whole *Mtb* genome sequence. Some of the most studied proteins have been those encoded by the so-called regions of differences (RDs). RDs were identified by comparative genome analysis and consist of 16 genomic segments that are present in *Mtb* (H37Rv strain) but absent from most BCG vaccines and non-tuberculous mycobacterial strains [95–97] (Supplementary Table1). Although those genomic regions together encode 129 predicted open reading frames (ORFs), only a few of these have been studied in detail -at least according to accessible published information. Moreover, those that were translated into approved TB diagnostic tests had already been discovered through conventional approaches described above [98]. Very recently only, a new overlapping peptide cocktail covering the RD7 encoded esx member antigen Rv2348 in combination with CFP10 and two other secreted proteins (EspC (Rv3615s) and EspF (Rv3865)) was successfully evaluated using this approach, and subsequently developed into an ESAT-6 free diagnostic IGRA test [99]. The ESAT-6 family comprises 23 small proteins (from EsxA to EsxW) that are mostly secreted in pairs (with the only exception of EsxQ). Not only ESAT-6 or CFP-10 but also other members of the ESAT-6 family are highly antigenic as demonstrated across several independent post-genomic studies (Supplementary Table 1, Figure 1). Furthermore, several ESAT-6 family members were found to have strong vaccine efficacy in mouse models: immunization with the fusion protein H65 (containing the ESX dimer substrates EsxD-EsxC, ExsG-EsxH, and ExsW-EsxV) had comparable protective efficacy with BCG [100].

3.1.1. Conventional HLA class I restricted CD8+ T cell targeting peptide libraries and predictive algorithms: complementary approaches successfully identify Mtb epitopes for Mtb specific CD8+T cells

Other new candidate *Mtb* antigens were studied for antigenicity using conventional peptide arrays. These candidates were selected based on specific criteria such as MHC binding algorithms that were used to mine public genomic databases (mainly TubercuList) or published proteomic/transcriptomic data obtained from *in vitro Mtb* intracellular cultures. One of the first libraries built on such an approach consisted of 15-mer peptides covering the sequences of 389 proteins [101], and aimed to identify CD8+ T cell stimulating epitopes by using a small panel of human CD8+ T cell indicator clones restricted to the classical HLA class Ia alleles B*5701, B*3905 and B*3514. The proteins included were selected based on three parameters: (i) gene products described in TubercuList as “PPE/PE”, “cell wall and cell processes”, “virulence, detoxification, adaptation” and “secreted”, (ii) genes highly expressed during *in vitro Mtb* intracellular infection, and (iii) genes not expressed by BCG strains. This resulted in the synthesis of 39,499 peptides, which were then pooled and tested for T cell recognition by IFN-ɣ ELISPOT. The peptide pools were screened in the presence of up to nine different CD8+ T cell clones from two different donors (one TB patient and one LTBI), autologous dendritic cells and IL-2. In this way, new CD8+ epitopes encoded by the *Mtb* antigens EsxJ (Rv1038c), PE9 (Rv1088) and PE_PGRS42 (Rv2487c) could be identified (Supplementary Table1). When this same 15-mers library was screened with a panel of more than 30 non-classically HLA class I (that is: non-HLA class Ia A,B or C) restricted *Mtb*-reactive CD8+ T cell clones, no epitopes presented by the HLA class Ib molecules MR-1 or HLA-E were found.

Recently, the same peptide library has been tested again in IFN-ɣ ELISPOT assays usingex-vivoperipheral blood CD8+ T cells from 20 ethnically diverse individuals including5 TB patients and 15 LTBI donors [102]. That study did not show differences in the magnitude of response between TB and LTBI, but validated 17 know CD8+ T cell epitopes and identified several new ones scattered across 58 antigens (Supplementary Table 1). Peptide pools were considered immunodominant when able to induce positive IFN-ɣ ELISPOT responses in 3 or more *Mtb* exposed individuals. Four antigens, specifically PPE15 (Rv1039c), PPE51 (Rv3136), PE12 (Rv1172c) and PE3 (Rv0159c), discovered in this CD8+ T cell library and three of them independently by other two post-genomic approaches (discussed below) (Fig. 1) [103,104] have been recently evaluated for protective efficacy in C57BL/6 and BALB/c mice models using a ChAdOx vector platform [105]. Two of them, PPE15 and PPE51 were able to reduce the *Mtb* pulmonary load as stand-alone vaccines, but only PPE15 boosted BCG protective efficacy in the C57BL/6 (but not in BALB/c) mice. The reason why antigens selected using similar criteria (based *e.g.* on protein category, antigenic or immunogenic properties) have different protective effects *in vivo* is a common finding but still not fully understood.

3.1.2. HLA restricted CD8+ and CD4+ T cell targeting peptide libraries based on in silico predictive algorithms successfully identified epitopes for Mtb specific CD4+ and CD8+ T cells

In silico epitope prediction programs, even when unable to consider the effects of 3D peptide structures, post-translation modifications or alternative peptides sequences generated by proteasome-catalyzed splicing [106,107], can be used to select putative antigenic peptides without the need to screen the entire amino acid (aa) sequence of target proteins. Using these bioinformatics tools, novel CD8+ or CD4+ T cell epitopes of known *Mtb* proteins were found, which were validated by MHC binding assays and reactogenicity of lymphocytes from *Mtb* exposed individuals, and/or used to generate *Mtb* epitope specific induced T cell clones [108–112]. These pioneering studies, although limited to

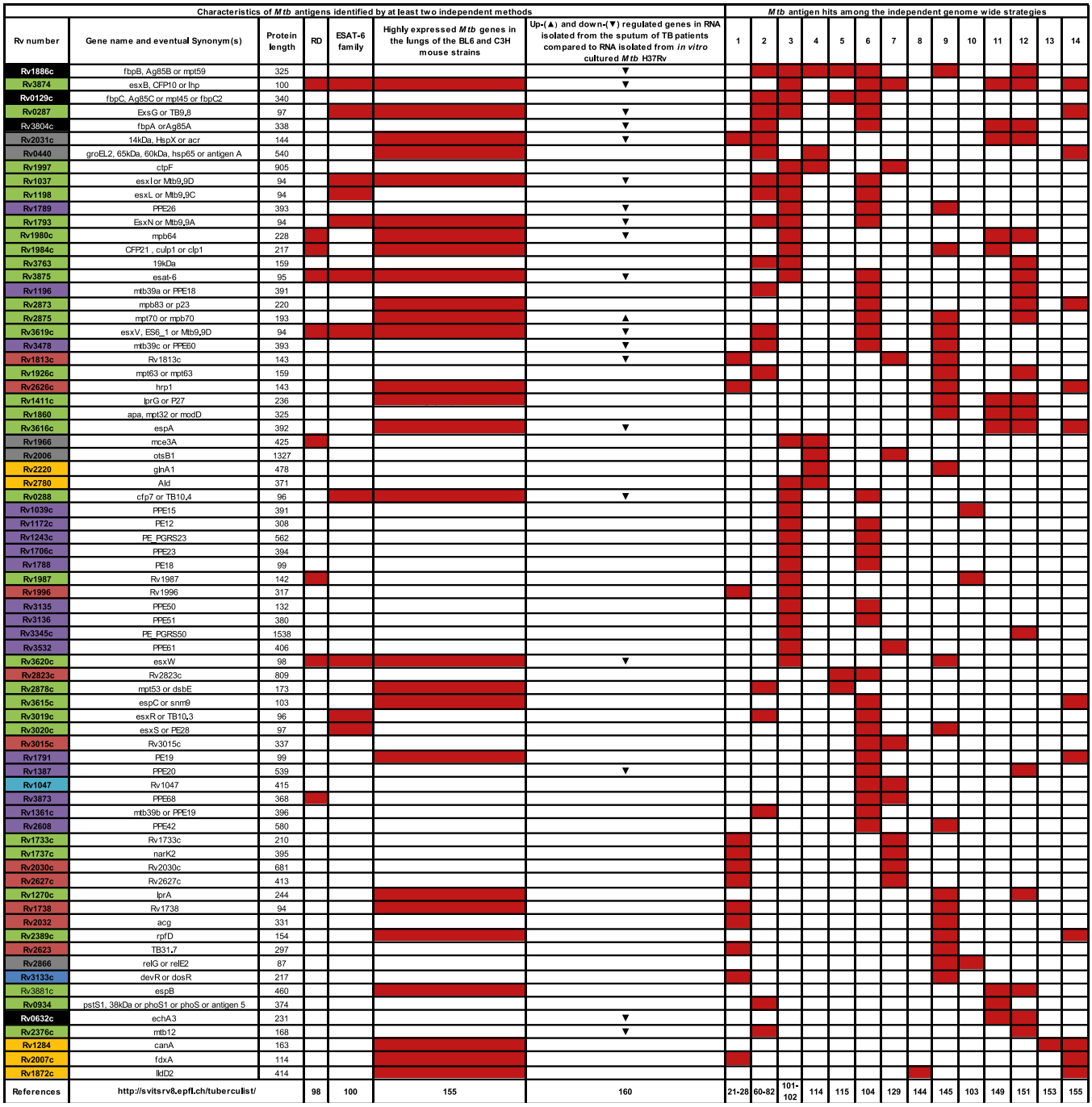

Fig. 1. Overlap among Mtb antigens identified by independent genome wide strategies.
Mtb antigens identified by at least two independent genome wide strategies (discussed in this review) are shown in this heat map. The methods included are indicated with number from 1 to 14 (the respective references are specified in the last row and in the legend). The colours assigned to each Rv number follows the division of the functional categories described in the TubercuList database (details in the legend). Additional information regarding these proteins (gene names, protein lengths, being encoded by genes belonging to the Regions of Differences (RD) or to the ESAT-6 family and the expression profile of the encoding genes identified by RNA isolated from the lung of Mtb infected mice and from the sputum of TB patients in relation to *in vitro* cultured Mtb H37Rv) are also accessible in this Figure.

previously studied *Mtb* proteins and to the most common HLA alleles, provided the proof of concept for later application ofin silico epitope screenings to the full *Mtb* antigenome. In search of new potential CD4+ T cell epitopes, 73 secreted proteins were directly selected from the *Mtb* genome and screened for HLA-II binding motifs [113]. The positive hits were then ranked and the top 17 novel peptides synthesized. One of these epitopes (belonging to the Rv2223c protein) was promiscuously and broadly recognized by IFN-γ producing peripheral blood cells (PBMC) from LTBI donors (n = 11) (Supplementary Table 1) and proved to be immunogenic in HLA-DR B*0101 transgenic mice when administered in a DNA plasmid vector co-expressing other 23 *Mtb* epitopes previously identified.

Using a similar approach, which instead was CD8+ T cell epitope oriented, 235Mtbproteins were screened using genome-based bioinformatics for potential binding peptide epitopes to HLA-A2, -A3, and -B7 as HLA class Ia superfamily members [114]. The *Mtb* proteins included were: (i) proteins already used in TB vaccine trials, (ii) proteins with already known CD8+ T cell epitopes, (iii) proteins with the

Functional categories (TubercuList)
cell wall and cell processes
conserved hypotheticals
information pathways
insertion seqs and phages
intermediary metabolism and respiration
lipid metabolism
PE/PPE
regulatory proteins
virulence, detoxification, adaptation

Method	Starting point	Leading Screening Criteria	Readouts	Types of sample used	Study subjects	References
1	genome-wide expression profiling of *Mtb in vitro* exposed to hypoxic culture conditions (DosR regulon)	*Mtb* antigens expressed as an adaptive response to stress	IFN-γ/Flow cytometry	T cell lines/PBMC/whole blood assay	LTBI/TB patients/household contacts	21-28
2	*Mtb* genome expression libraries	isolation of novel *Mtb* antigens including those not always expressed by *Mtb*	immunoscreen/ T cell proliferation/IFN-γ	(murine) monoclonal antibodies /polyclonal sera/PBMC	TB patiens/PPD+ donors	60-82
3	389 *Mtb* proteins selected via arbitrary criteria	15-mers library (HLA-B*5701/B*3905 /B*3514)	IFN-γ	CD8+ T cell clones/ex-vivo peripheral blood CD8+ T cells	TB patiens/LTBI donors	101-102
4	235 *Mtb* proteins selected via arbitrary criteria	9-mers library (HLA-A*0201/A*0301/B*0702)	CD8+ T cell proliferation assays/HLA-A*0201-binding/polyfunctional (IFN-γ+/IL-2+/TNF-α+) CD8+ T cells	PBMC	PPD+ donors/cured TB patients/healthy controls	114
5	3924 predicted ORFs of *Mtb* H37Rv	9-mers library (HLA-B*3501)	IFN-γ	expanded CD8+ T cells	BCG vaccenees	115
6	all predicted protein sequences of 21 *Mtb* strains	15-mers library (22 different HLA-DR, -DP and -DQ)	IFN-γ/Flow cytometry	PBMC	LTBI	104
7	ORFs corresponding to all predicted proteins of *Mtb* H37Rv	9-mers predicted to bind HLA-E	peptide binding assay/T cell proliferation/Flow cytometry/Cytotoxicity and Suppression assays	PBMC/HLA-E expressing cells lines	PPD+ donors/healthy controls/BCG vaccinated infants/cord blood samples	129
8	3,774 coding regions from the genome of 216 *Mtb* strains	hypervariable genes	IFN-γ	diluted whole blood samples	TB patients	144
9	94 *Mtb* genes predicted to encode secreted proteins	EsX or PE/PPE /expressed in *Mtb* infected macrophages/up- or down-regulated during hypoxia or starvation	IFN-γ	PBMC	PPD- and PPD+ donors	145
10	3480 *Mtb* proteins	antigens reactive to pooled sera	high-throughput proteome microarray technology	pooled sera	TB patients	103
11	4099 *Mtb* proteins	antigens reactive to sera	high-throughput proteome microarray technology	sera	TB patients	149
12	164 *Mtb* proteins identified through serological screening	novel *Mtb* CD4+ T cell antigens	IFN-γ	polyclonal *Mtb* -reactive CD4+ T cell lines	LTBI/healthy controls	151
13	2170 *Mtb* genes in the lung of 4 mouse strains(6 and 9 weeks post-infection)	IVE-TB genes associated with different TB outcomes	IFN-γ/Flow cytometry	whole blood samples/PBMC	LTBI/TB patients	153
14	2068 *Mtb* IVE-TB genes in the lungs of B6 and C3H mouse strains (0-2-4-6-9-12 weeks post-infection)	highly expressed IVE-TB genes combined with *in silico* criteria	IFN-γ,TNF-β, IL-17, IL-13, IP-10, and GM-CSF/T cell proliferation	diluted whole blood samples/PBMC	LTBI	155

Fig. 1. (*continued*)

highest HLA class I binding prediction values, (iv) conserved proteins, (v) proteins encoded by the *Mtb* DosR regulon, (vi) proteins with known B cell epitopes, (vii) secreted proteins, and (viii) proteins predicted to be secreted [114]. Out of 432Mtb9-mer peptides tested, 70 (of which 58 were unknown epitopes) were found to induce proliferative responses in CD8+ T cells from two or more individuals out of the 41 purified protein derivative responder (PPD+) healthy donors screened. Based on the frequency and the magnitude of CD8+ T cell responses, 18 out of these 70 epitopes, mainly including peptides from secreted antigens, were then selected for further testing in ten HLA-A*0201 positive TB patients and an equal number of HLA-A*0201 positive controls, using *Mtb* specific petide/HLA-A*0201 tetramers and functional peptide stimulation assays. All 18 selected epitopes were confirmed to be antigenic in TB patients (but not negative controls) as visualised by specific tetramer staining, and to induce specific poly functional (IFN-γ+/IL-2+/TNF-α+) CD8+ T cells (Supplementary Table 1). Of note, half of the 18 proteins encoding these epitopes had also been screened in the 15-mer CD8+ peptide library mentioned above, but only three antigens (Rv1966, Rv1997 and Rv2780) were found in both approaches to contain *Mtb* antigenic peptides stimulating CD8+ T cells (Fig. 1). This significant yet limited overlap could be explained by differences in the design of the two studies: the criteria applied to determine potential *Mtb* antigens (this resulted in only 56 putative antigens shared between the ones selected in both studies), the length of the peptides generated from these proteins (9-mers *vs*. 15-mers), the way in which the peptides were evaluated (single peptides *vs*. peptide pools), and the read-outs chosen to evaluate the immunodominance of the peptides (proliferation / CD8+ poly functionality / HLA class Ia tetramer staining in the second *vs*. the use of selected T cell clones / IFN-γ ELISPOT in the first study).

The first *in silico* genome-wide screening which did not apply an arbitrary criteria driven selection of *Mtb* proteins was performed in search of putative HLA-B*3501 T cell stimulating *Mtb* epitopes [115]. This allowed the identification of both known as well as novel antigens (Rv0670, Rv1280c, Rv1464, Rv1641, Rv2182c, Rv2476, Rv2823c, Rv3378c and Rv3689) encoding new epitopes recognised by CD8+ T cells from BCG vaccinated healthy donors in the context of one single HLA allele (Supplementary Table 1). Another study, applying a similar methodology but restricted rather to HLA-A*0201, discovered two other new antigens (Rv1490 and Rv1614) that contained CD8+ T cell stimulating epitopes, and that were recognized by CD8+ T cells from LTBI and TB patients (Supplementary Table 1) [116].

3.1.3. Additional across genome epitope predictions identify many novel Mtb epitopes for CD4+ T cells

*With the better definition of HLA-I supertypes and corresponding alleles [117], the increased knowledge of the global distribution of the most common HLA-II molecules [118], and the availability of additional *Mtb* genome sequences, the comprehensiveness and sizes of HLA allele based *in silico* epitope predictions for *Mtb* sharply increased. One

of the largest *in silico* genome-wide screening efforts used an approach parsing all predicted protein sequences of 21 *Mtb* strains into all their possible 15-mers and subsequently selecting the most conserved and best-predicted promiscuous binders for 22 different HLA-DR, -DP and -DQ class II alleles [104]. A total of 20,610 peptides was then selected for synthesis and tested in pools of 20 peptides to assess the *in vitro* IFN-ɣ response elicited in PBMC of 28 LTBI donors by ELISPOT. Based on peptide pools recognised by three or more LTBI donors, 369 individual *Mtb* epitopes, covering 82 different *Mtb* antigens, were identified (Supplementary Table 1). Of note, it was reported that 47% of those peptides accounted for 90% of the total responses. By depleting PBMC of either CD4+ or CD8+ T cells it could be confirmed that 97% of the epitopes were recognized exclusively by CD4+ T cells. It is not fully clear why so few CD8+ T cell epitopes were found, but suboptimal peptide length and assay criteria might have contributed to favouring CD4+ T cell epitope identification. In any case, out of the 82 *Mtb* antigens, 34 were recognized as CD4+ T cell targets for the first time. The majority of the *Mtb* antigens identified were associated with most categories indicated in TubercuList, but with a relative overrepresentation of antigens from the PE/PPE family, a large and unique family of genes representing over 10% of the entire *Mtb* genome [20]. Interestingly, the authors introduced the concept of antigenic clusters which they defined as groups of at least 4 proteins recognized by at least two LTBI donors and encoded within a 5 gene-interval. Based on this approach, they then identified three antigenic islands that mainly included known components of the type VII secretion systems (T7SS) Esx-1 and Esx-3, and ten novel antigens were reported for the first time. The researchers then further investigated the antigenicity of those new peptide pools by cytokine induction (IFN-γ/TNF-α/IL-2) and by assessing T cell memory phenotypes. No differences in cytokine production were observed between the proteins identified [104]. Of particular interest, the authors found a new memory CD4+ T cell subset to be involved in the recognition of *Mtb* antigens in LTBI donors. This Th1* cell subset expressed a unique, CD4+CXCR3+CCR6+CCR4- phenotype, and had a lineage-specific transcriptional signature shared with both Th1 (such as T-bet (TBX21), granzymes A and K, perforin (PRF1), and the transcription factor EOMES) and Th17 (RORC, DAM12, PTPN13, and IL17RE, the receptor for IL-17C) cells [119]. Furthermore, the CD4+CXCR3+CCR6+CCR4− T cells selectively expressed genes involved in cytokine/receptor interactions (CCR2, IL12RB2, IL23R, KIT [CD117, c-Kit], BAFF [CD257, TNFSF13B]), cell persistency and proliferation (*i.e.*, BAFF, MDR1 (ABCB1), and KIT) as well as genes previously reported associated with TB susceptibility (CCR2 and IL12RB2). In addition to IFNG, TNF and IL2, the activated CD4+CXCR3+CCR6+CCR4− T cells expressed cytokine transcripts (CSF1/2, CCL3/4, GZMB, IL6/17A/22, CXCL9, and VEGFA, CSF1 (M-CSF) and GM-CSF) many of which have been shown to play a role in TB containment previously. This makes these CD4+ T cell subsets also interesting as potential *Mtb* specific correlates of protective immunity.

Screening large cohorts of *Mtb* exposed individuals in different geographic settings obviously is essential to corroborate the antigenicity of *Mtb* peptides whose immunodominance has been based on positive IFN-γ responses limited to few donors in a single site. To accomplish this, 15-mer peptide pools (20 per protein per pool) of 25 novel antigens identified in the last study described were further examined by ELISPOT in 128 LTBI from nine different geographical locations. Importantly the results validated the antigenicity of the selected epitopes [120]. It is interesting to note that peptide pools from four novel antigens (Rv1172c, Rv1788, Rv1791, and Rv3135) that were shown to be recognised by (presumably CD4 + T cells from) most of the LTBI cohorts, were among those subsequently described to contain also CD8 + T cell epitopes in another study (Fig. 1) [102]. Recently, IFN-γ based recognition of 40 *Mtb* antigens identified in the previously described CD4+ library together with 20 *Mtb* antigens selected from other studies was evaluated in a diluted whole blood assay in Atlanta and Kenya [121]. The peptide pools of proteins previously described as antigenic (such as Rv1172c and Rv1872c among others) did not always induce high IFN-γ responses in the cohorts of LTBI, TB household contacts and TB patients included in this study. This difference could be explained by the different composition of peptide pools used (18-mer based), the read-out implemented (1:4 diluted whole blood stimulated 7 days with 1ug/ml), environmental or genetic (HLA) differences between the populations, amongst others.

The *in vitro* reactogenicity of human T cells to *Mtb* peptides restricted to diverse HLA molecules can indeed be highly heterogeneous due to the extensive HLA polymorphisms [122]. To solve those limitations, a recent study tested a peptide pool of 300 *Mtb* epitopes and showed that this comprehensive megapool, consisting of peptides from 52 *Mtb* antigens (including 45 antigens discovered in the above described CD4+ peptide library, of which 7 were also evaluated as peptide pools in the 9 different cohorts of LTBI discussed above), accounted for 80% of *Mtb*-specific T cell responses in LTBI, including adults and adolescents [123]. The study cohort included LTBI with and without a self-reported history of active TB. Interestingly, the *Mtb* epitope reactivity, based on IFN-γ ELISPOT and intracellular cytokine staining, was around 10-fold lower among subjects that had experienced past TB disease in the last 6 years, for a specific set of epitopes, which was designated “type 2” or “post TB sensitive” *Mtb* epitopes [124]. No such difference was found for “type 1” or “persistent” epitopes, which were recognized equally well by LTBI with or without previous TB. Confounding factors such as an active TB disease process (which was evaluated by comparing responses between type 1 and type 2 epitopes in 16 TB patients), differences in T cell memory phenotypes, or differences in gene expression in responding T cells could be excluded. Also numbers of *in silico* predicted HLA-II binding motifs did not differ among the different classes of epitopes. The only difference found was that *Mtb* epitopes that were less recognised by LTBI with past-TB were those with the highest sequence homologies to proteins from the human microbiota or from nontuberculous mycobacteria (NTM). Based on their data, the authors concluded that the diversity in type 2 epitope reactogenicity among LTBI with or without past TB could be related to changes induced in the microbiota by TB treatment, although some mycobacterial specific proteins, such as ESAT-6 or CFP-10, still contained several type 2 peptides. Data on the microbiota compositions of TB patients followed pre-, during and after treatment will, however, be needed to validate this hypothesis. It would also be important to determine the general homology between other proposed *Mtb* antigens and the human microbiota. The fact that NTM homologous *Mtb* epitopes were found to be less reactogenic in the post-TB donor group is also quite intriguing especially considering previous reports showing a strong reactogenicity to such epitopes in LTBI and in healthy individuals from non-endemic areas [125]. Moreover, the low recognition of NTM homologue epitopes in the post-TB LTBI donors is in line with the finding that other NTM homologues, such as the DosR regulon encoded proteins [126], are less recognized in treated TB patients than in LTBI [23]. A role for NTM exposure in inhibiting BCG vaccine induced protective immunity has been repeatedly proposed [4] and experimentally proven in mouse models [127]. The relation to the phenomenon of type 2 epitope remains to be clarified.

3.1.4. Unconventional HLA class Ib restricted CD8+T cells: identification of Mtb epitopes for CD8+T cells genetically restricted by HLA-E, with an unusual phenotype and function

Alternatively to conventional HLA-I and HLA-II binding motifs, peptides presented by non-classical HLA molecules have been proposed as interesting *Mtb* antigenic targets. For HLA-E only 2 coding variant molecules are known, which differ in a single amino acid localized outside the peptide binding groove. Thus, HLA-E molecules can essentially be seen as virtually monomorphic antigen presentation molecules, suggesting the interesting possibility that a relatively small number of *Mtb* peptides can be used for presentation *via* HLA-E for *e.g.* vaccination purposes. Based on initial evidence that HLA-E molecules could present

Mtb derived peptides to two human T cell clones [128], the *Mtb* (H37Rv) genome was screened *in silico* for candidate epitopes predicted to bind to HLA-E molecules [129]. This effort resulted in the selection of 69 potential HLA-E binding peptides, which were tested for HLA-E binding and CD8 + T cell recognition (by T cell proliferation and cytokine induction) in PPD + donors, PPD- healthy controls, BCG vaccinated infants and cord blood samples (Supplementary Table 1). Unexpectedly almost all predicted epitopes could be validated and some were recognized by as many as 40% of the PPD + responsive donors. Several epitopes were also recognized by the BCG vaccinated infants from South Africa, but not by cells from cord blood, suggesting vaccine induced T cell memory. Subsequent investigations focussed on the most frequently recognized immunodominant HLA-E presented *Mtb* peptides using T cell clones. This analysis revealed that these HLA-E restricted CD8+ *Mtb* specific T cells had a Th2 like phenotype and function. This was confirmed in the circulation of TB patients and LTBI [130–132]. Although numbers were small, there was a tendency towards higher frequencies of HLA-E restricted *Mtb* specific T cells in TB/HIV coinfected individuals. Using a new expansion protocol these results could be corroborated using HLA-E tetramers combined with specific functional phenotyping in the blood of TB patients and LTBI [132]. The low polymorphism of the HLA-E locus and the stable expression of the HLA-E molecule even upon infection with HIV may render HLA-E peptides to be attractive antigenic targets [133].

In this context it is relevant to mention that several of the *Mtb* peptide specific HLA-E restricted T cell clones inhibited *Mtb* growth in human macrophages, in line with a possible protective function. Furthermore, several studies in animals now suggest that HLA-E like Qa-1 restricted CD8 + T cells contribute to protective immunity *in vivo*. The murine homologue of HLA-E, Qa-1, was found capable of presenting the same HLA-E binding *Mtb* peptides to $CD8^+$ T cells, and Qa-1 knockout mouse were more susceptible to *Mtb* and died earlier with higher bacterial burdens [134]. This was not due to engagement of NK receptors, as demonstrated in genetic ablation experiments. In addition, as mentioned above, NHP vaccination studies with genetically modified RhCMV vectors expressing *Mtb* antigens revealed that MHC-E restricted T cells can be induced and mediate part of the strongly protective response against TB I the NHP model [53]. Clearly HLA-E restricted T cell immunity needs to be investigated in more detail in the context of TB vaccination.

Besides HLA-E restricted T cells there are also other groups of so-called donor unrestricted T cells or DURT T cells [135]. These cells are mostly CD8 + T cells that are genetically restricted by highly conserved, non-polymorphic presentation molecules: MR1 (presenting metabolites to mucosal associated invariant T cells or MAITs); CD1a, b and c molecules presenting mostly lipid antigens; NKT cells restricted by CD1d; TCR$\gamma\delta$ cells, some which recognize phospho-antigens in the context of butyrophilin 3A1. However, none of these cells recognize peptide antigens and therefore are not discussed in detail in this review, which focuses on postgenomic *Mtb* protein and peptide antigen identification.

3.1.5. Hyperconserved or variable Mtb epitopes

Most of the studies discussed so far have aimed at identifying immunodominant *Mtb* peptides and protein antigens for T cells. However, the focusing of T cell responses exclusively on conserved immunodominant epitopes has been under discussion lately after human T cell *Mtb* epitopes were found to be evolutionary more conserved than essential *Mtb* genes as a whole, or than the non-epitope encoding sequences of the same *Mtb* antigens [136] (in this work 491 experimentally verified peptides were studied covering 78 *Mtb* antigens). This discussion impacts also TB vaccine development strategies since it was proposed that the observed T cell epitope hyper conservation suggested that natural T cell immunity might be beneficial for *Mtb* transmission in some critical step of the infection cycle in the human host [137]. Although highly speculative, this phenomenon might play a role in TB cavitation, which is considered to be due to excessive inflammation and plays a key role in transmitting *Mtb* bacteria to new susceptible hosts.

Another possibility may be that *Mtb* has evolved to drive T cell responses preferentially to highly expressed *Mtb* epitopes, which would likely skew reactive T cells from a central memory towards a terminally exhausted phenotype, as suggested by observations in mice [138], thereby limiting the formation of a pool of long term memory cells. Under such a scenario an alternative approach towards vaccine design would then be to mobilize subdominant epitopes that are less recognized during natural infection, thus not only circumventing the induction of exhaustion but also to broaden the *Mtb* antigenic repertoire that can be recognized beyond the natural infection induced repertoire. A first study to assess the protective efficacy of subdominant *Mtb* peptides was performed in B6CBAF1 mice using ESAT-6 peptides [139]. It was demonstrated that *Mtb* peptides not highly recognized after natural *Mtb* infection could induce better protection when used as vaccine antigens than the naturally immunodominant epitopes. These results have been subsequently extended to other mouse models (CB6F1; C57BL/6 and BALB/c) [140,141] and other antigens [142]. Although the protective efficacy differed among mouse strains, it was shown that indeed these ESAT-6 subdominant epitopes led to less terminally differentiated T cell profiles, with a greater ability to sustain polyfunctional cytokine responses and proliferative capacity over time [141].

Alternative interpretations to the role of hyper conservation of T cell epitopes, however, exist and hyper conservation could also be due to a strong evolutionary selection pressure on *Mtb* protein domains due to indispensable functions rather than T cell immune pressure. It has been argued that if T cell recognition was driving epitope conservation, this should have been restricted mainly to HLA molecules associated with genetic TB disease susceptibility. However, in clinical *Mtb* strains isolated from TB endemic areas, *Mtb* T cell epitopes rarely recognised by TB patients or restricted to infrequent HLA molecules are not less conserved than those binding to the most common HLA molecules and more strongly recognised by TB patients [143].

Recent results have confirmed the low rate of *Mtb* epitope sequence variation and found only an small number of variable regions in the *Mtb* genome [144]. From the encoded sequences of 7 such variable genes (Rv0001c, RimJ (Rv0995), Rv0012, LldD2 (Rv1872c), Rv0990c, Rv2719c and TB7.3 (Rv3221c)) (Supplementary Table 1) HLA-I and HLA-II binding motifs were predicted *in silico*. Comparative analyses showed aa changes in predicted CD8+ but not CD4 + T cell epitopes. Interestingly, those substitutions had an effect on the antigenicity of epitopes *in vitro* as demonstrated by measuring the INF-γ production induced in dilute whole blood of 82 TB patients. Hence, variant *Mtb* epitopes, generated by aa substitutions, could be either less or more efficiently recognized than the ancestral peptide. Whether in a vaccine setting the hypervariable *Mtb* epitopes would be more efficacious than conserved *Mtb* epitopes still needs to be determined. Furthermore, additional studies, including other regions such as Asia will be needed to lend further support for hyperconservation and variation of *Mtb* epitopes.

3.2. Searching novel Mtb antigens

In the prior section, we reviewed different approaches used to search for novel *Mtb* peptide epitopes. However, other studies have used complementary data driven approaches to interrogate the *Mtb* genome seeking new *Mtb* antigens. In several of these approaches arbitrary criteria were used to mine genomic and proteomic databases, which in one study resulted in the selection of 94 *Mtb* genes predicted to encode secreted proteins, in particular members of the EsX or PE/PPE families, expressed in *Mtb* infected macrophages and up- or down-regulated in response to hypoxia or carbon starvation [145]. IFN-γ production in response to these recombinant proteins was tested in PBMC from PPD− and PPD+ individuals and 48 proteins were found to be highly reactogenic (Supplementary Table 1). It is interesting to

note that ten of those antigens were confirmed to contain CD4+ or CD8 + T cell epitopes in subsequent studies (Rv1789, Rv1813c, Rv1886c, Rv1984c, Rv2220, Rv2608, Rv2875, Rv3020c, Rv3478, Rv3619c, and Rv3620c) (Fig. 1) [91,102,104,114,129]. The 48 reactogenic antigens were then adjuvanted with CpG and each individually evaluated for protective efficacy in *Mtb* infected C57BL/6 mice. None of the antigens, although eliciting both antibodies (IgG1 and IgG2) and cytokines (IFN-γ and TNF-α), however, resulted in better reduction in lung CFU than BCG. Nevertheless, four were selected (Rv1813c, Rv2608, Rv3619c and Rv3620c) and fused to form the ID93 vaccine, which is now under clinical evaluation in various trials and trial designs (PoD as well as Prevention of Recurrence (PoR)) in combination with the glucopyranosyl lipid adjuvant (GLA) formulated in SE [146]. The ID93/GLA-SE fusion protein vaccine [147] was demonstrated to have similar -though not superior- protective efficacy as BCG as assessed by CFU in the lungs of *Mtb* infected mice. It induced pro-inflammatory cytokines (MCP-1, IL-8, IL-6, IL-5, TNF-α, and GM-CSF) in cynomolgus macaques and decreased mortality and lung pathology in guinea pigs, when used in a BCG prime-boost strategy [148].

Additional broad screening of the *Mtb* antigenome was performed by using high-throughput proteome microarray technology. In one study, the screening of 3480 *Mtb* proteins with pooled TB patients' sera reported significant immunoreactivity to 249 proteins; these included proteins belonging to the PE/PPE, DosR and RD families, as well as known "conventional" antigens [103]. From these screens, three novel proteins (Rv1987, Rv3807c, and Rv3887c) were found to be particularly highly reactogenic and to provide better diagnostic sensitivity and accuracy than commercial serological tests (Supplementary Table 1). Through a similar approach, the recognition of 4099 *Mtb* proteins was tested with sera from 561 TB suspects. This study identified 484 *Mtb* proteins that were recognized by the sera of at least one TB patient, and recognition of 13 proteins was significantly associated with active TB (Rv3881c, Rv3804c, Rv3874, Rv1860, Rv1411c, Rv2031c, Rv0934, Rv3616c, Rv3864, Rv1980c, Rv0632c, Rv1984c, and Rv2773) (Supplementary Table 1) [149]. There was no overlap with the proteins identified in the just mentioned previous study. Although both screenings identified novel antigens reactive with TB patients' sera, current global recommendations from WHO and others do not encourage the use of serological tests for TB diagnosis, since the high test variability leads to high rates of false positive and false negative results [150].

In another study, which was based on the assumption that CD4 + T cells are required for the maturation of long-lived plasma cells, 164 *Mtb* proteins that had been identified through serological screening [149], were interrogated in a search for novel *Mtb* CD4 + T cell antigens [151]. Polyclonal *Mtb*-reactive CD4 + T cell lines, generated from PBMC of 12 LTBI (three from USA and nine from India) and two healthy controls (from USA) exposed to *Mtb* lysate, were used to test this proteome set. The immunological responses were measured by IFN-γ release after three days of culturing the T cell lines with irradiated antigen presenting cells and the respective unpurified antigen. Forty-three proteins were recognized by CD4 + T cells from at least one LTBI or healthy control, resulting in a total of 27 newly characterized CD4 + T cell *Mtb* antigens (Supplementary Table 1). Although performed in a small group of individuals, this study proposed a scalable system that could be used as a workflow to screen *Mtb* antigens directly in endemic areas, and proved that antigens recognized by immunoglobulins often also react with CD4 + T cells.

3.2.1. Integrating Mtb antigen discovery with Mtb infection biology and in vivo expression in the infected lung

The various approaches described so far have expanded our knowledge of the *Mtb* anti genome for T cells and antibodies from LTBI or TB patients, but did not examine whether and how those antigens were expressed in the primary TB target organ, the lung. This characteristic might be not essential for TB biomarkers, but it could be crucial for antigens proposed as potential vaccine candidates. A minimal prerequisite for a vaccine antigen is that it is expressed by infected cells in the infected target organ, in this case the lung. Immune responses directed against *Mtb* antigens expressed in the lung could restrain the *Mtb* immunological life cycle at an early stage and prevent the onset of TB [152]. Based on this hypothesis, we studied a new class of *Mtb* antigens, which we designated IVE-TB (*in vivo* expressed) antigens [153]. Our first IVE-TB antigen set was based on the analysis of RNA expression patterns of 2170 *Mtb* genes in the lung of four mouse strains at 6 and 9 weeks post-infection. Based on the distinct TB susceptibility phenotypes of the mouse strains examined (relatively resistant C57BL6 *vs*. super susceptible C3H/FeJ as polar extremes) *Mtb* genes were then selected to represent genes expressed: (i) independently of host genetic background; (ii) in association with necrosis; (iii) in association with severe necrotic infection or susceptibility (expressed in the C3H but not B6, C3H.B6-sst1, or B6.C3H-sst1); (iv) in association with dense granuloma development; (v) in association with diffuse granuloma development; (vi) in association with resistance; (vii) in association with low inflammation; (viii) inflammation; and (ix) relapse. This resulted in the first selection of 16 IVE-TB genes, which were tested as recombinant proteins *in vitro* and *in vivo*. IFN-γ responses to seven IVE-TB antigens were observed *in vitro* by screening 133 TST + donors (Supplementary Table 1). A further in depth study using PBMC from six LTBI revealed that the most pronounced T cell subsets recognizing IVE-TB antigens were IFN-γ+/TNF-α+ CD8 + T cells and TNF-α+/IL-2+ CD154+CD4 + T (central memory) cells. Further experiments with one of those IVE-TB antigens, Rv2034, confirmed its *in vivo* immunogenicity in HLA-DR transgenic mice by strong induction of T cells and antibodies [154]. Moreover, immunization with Rv2034 or the hybrid-protein Ag85B-ESAT6-Rv2034 adjuvanted with CpG or CAF09 induced over one log reduction in lung CFU compared to unvaccinated controls both in *Mtb* challenged HLA-DR3 transgenic mice and guinea pigs [154]. Together, these data suggest the potential use of this novel class of antigens for future TB vaccination.

Building upon these results, we recently selected a second, much more extensive set of IVE-TB antigens [155]. Data included the relative gene copy number of 2068 *Mtb* genes expressed in the lungs of the same B6 *vs*. C3H mouse strains, but now at multiple different times post-infection (0-2-4-6-9-12 weeks). A total of 194 genes was found to be consistently up-regulated independent of the time of infection or the host genetic background. Bioinformatics was then applied to further select the most promising genes for functional and immunological evaluation. These analyses included: (i) the top 15% genes up-regulated at a late stage of infection; (ii) highly conserved genes with wide HLA coverage and/or with the highest numbers of predicted HLA class I and II binders; and (iii) genes with a high homology with *M. leprae*, the second human mycobacterial pathogen which is the cause of leprosy. A total of 50 IVE-TB proteins was selected and analysed, for the first time using an extensive cytokine screening panel, since IFN-γ is not a correlate of protection in TB. We therefore included also TNF-α, IL-17, IL-13, IP-10, and GM-CSF. Twenty-nine IVE-TB antigens were strongly recognized by multi-cytokine production by blood cells from *Mtb* exposed individuals (n = 12) and LTBI (n = 25) (Supplementary Table 1). To the best of our knowledge, 17 out of those were described for the first time as *Mtb* antigens. Importantly, almost half of the antigens were recognized by cells producing cytokines other than IFN-γ, including IP-10 that we recently showed to be mechanistically involved in controlling *in vitro* mycobacterial growth in MGIA assays in recently *Mtb* infected individuals [156]; and GM-CSF which also has been associated with protective immunity [157–159]. This data suggest that IFN-γ based screening approaches may have significantly underestimated as well biased *Mtb* antigen discovery studies. These findings are currently being validated in an independent cohort of LTBI and TB patients and protection studies are ongoing in different mouse strains to assess the protective efficacy of the most interesting IVE-TB antigens (personal communication).

In both our IVE-TB studies, IVE-TB genes were selected based on the *Mtb* transcriptomes conditioned by the interaction between the pathogen and the murine pulmonary (immune) environment. Although we demonstrated that IVE-TB antigens are strongly recognized by peripheral cells of LTBI, we could not exclude that the *Mtb* gene expression profile in the lung of human hosts might differ from the ones found in mice. Recently, the analysis of *Mtb* RNA isolated from the sputum of untreated TB patients (n = 7) offered the possibility to closely look at the *Mtb* transcriptional profile in humans (Supplementary Table 1) [160]. However, since this study mainly defined up- and down-regulated genes in RNA isolated from the sputum of TB patients to RNA isolated from *in vitro* cultured *Mtb* H37Rv, a direct comparison with our findings is not possible at this stage. Nevertheless, since the necrotic phenotype of TB lesions in the C3H model closely resembles key features of human TB lesions (caseating necrotic granulomas accompanied by hyper-susceptibility), as discussed by Kramnik et al. [161] previously, we contend our model is an important novel tool in the discovery of highly expressed antigens in affected target organs, not only for TB but also other diseases were similar approaches can be adopted.

4. Summary and implications

The discovery of *Mtb* antigens that correlate with infection, protection and vaccine immunogenicity is a complex process that has evolved over decades, and now is yielding important new results. The first goal is to identify the proteins and peptides that are expressed by the pathogen and can be recognized by the host immune system. In the early stages of antigen discovery, there were significant limitations in the resolution of biochemical technologies used to isolate and characterize proteins from *Mtb* from *in vitro* cultures. These limitations were overcome first by the availability of *Mtb* genome wide expressing libraries and subsequently by the availability of the whole *Mtb* genome sequence. This allowed informational and experimental access to the entire *Mtb* antigenome, with its approximately 4000 ORFs. These have been probed now extensively for their potential antigenicity, using several genome wide strategies. This has led to the identification of 3282 *Mtb* (ID1773) epitopes and more than 500 *Mtb* antigens (IEDB, www.iedb.org, March 2017). However, as described in the past [113], the majority of the *Mtb* epitopes included in IEDB disproportionally (54%) belongs to a relatively small proportion of proteins (n = 26).

The novel approaches discussed in this review have typically defined the antigenicity of recombinant peptides and proteins by *in vitro* measuring of IFN-γ and proliferative T cell responses using PBMC, T cell lines or whole blood from LTBI donors. We contend that additional parameters need to be included as well, based in part on our own observations that many new *Mtb* antigens were recognized by cells producing cytokines other than IFN-γ, and often no IFN-γ at all [155]. This suggests that IFN-γ based screening approaches may not have captured the *Mtb* antigenome adequately. This also is evident when examining alternative T cell responses such as those restricted by HLA-E, which often release Th2 rather than Th1 cytokines [130].

Current diagnostic tests, including TST and IGRA, have poor prognostic capacity in predicting which *Mtb* infected individuals will progress towards TB, that would allow rapid preventive treatment of these subjects to decrease the risk of *Mtb* transmission. Differences in *Mtb* antigen specific IFN-γ production and in polyfunctionality of T cell responses, such as to Rpf or DosR regulon antigens, have been found repeatedly between LTBI and TB patients [28,162]. However, not many of such antigens have been assessed in longitudinal follow-up studies of TB household contacts to examine whether they could predict TB progression. The novel epitopes or antigens identified by wide genome screenings as discussed above have been evaluated sporadically in multiple TB cohorts. When analysed, very few differences in the magnitude and frequency of responses, which were mostly IFN-γ centred, were found between TB patients and LTBI [102,114,124,131,153]. However, the number of subjects included in those studies was generally quite low and future studies would need to screen larger cohorts including follow-up analyses to capture their disease -or protection- association. Such studies could also be interesting as they might elucidate how the immune response repertoire against antigens/epitopes is shaped during the natural course of infection. That is important considering that antigens highly expressed at early stage of infection can lead to T cell exhaustion and dysfunction [163,164] while others not evoking exhaustion could induce long term memory. Moreover, most of the novel *Mtb* antigens/epitopes identified have high homologies to antigens from NTM or other bacteria [124,125], including those present in human microbiota. How this impacts their reactogenicity in TB needs to be clarified. Additionally, the currently used rather narrowly focused immunological read-outs (mostly IFN-γ or polyfunctional CD4 + T cell centred) are unlikely to detect immunological changes in other domains of immunity, which widely occur in *Mtb* carrying hosts as they transit from a stage of controlling *Mtb* infection to a process culminating in active TB disease [165–167]. Identifying such changes and defining the corresponding biomarkers of TB risk would greatly facilitate early TB diagnosis and prediction of TB onset at an early stage. Innovative animal models, like cynomolgus macaques that can display the entire human TB clinical spectrum [168,169], and which can also recognize CD4+ epitope pools defined in humans [170] would be of great value to help identifying such markers in translational studies.

As evident from the above, most *Mtb* genome wide antigen discoveries have relied on samples from LTBI donors. Those individuals are interesting from a vaccine development point of view, since LTBI results in an almost 80% lower risk of developing active TB than non-LTBI subjects upon re-infection [171]. However, the underlying biological mechanisms and immune correlates remain unknown. Most studies today still follow IFN-γ oriented approaches although we know that the presence of activated [172] and polyfunctional (IFN-γ+/IL-2+/TNF-α+) T cells are not correlates or sufficient mediators of protection [57,143]. A recent study in mice demonstrated that CD4 + T cells activated by systemic peptides administration was able to reach the lung parenchyma but, critically, failed to act directly with *Mtb* infected cells [173]. *Mtb* infected cells have the ability to decoy immune cells through different mechanisms such as suboptimal antigen presentation, exporting antigens to bystander uninfected cells reducing the recognition of those cells containing the mycobacteria, the release of inhibitory cytokines or the induction of inhibitory mechanisms such as regulatory T cells [9,174–176]. It would be interesting to study whether the decoy activity is restricted to certain antigens, such as secreted *Mtb* antigens. If so, vaccine strategies might need to be focused on non-secreted antigens, which would be contrary to most current thinking.

To advance novel antigens into the TB vaccine pipeline it will be necessary to prove their immunogenicity and especially their protective efficacy in preclinical animal models of increasing complexity and relevance to human TB. To our knowledge, from all antigens identified by recent genome wide strategies, only few antigens (http://www.tbvi.eu/what-we-do/pipeline-of-vaccines/) have been tested *in vivo*. Mice are generally used as first line *in vivo* model and usually the protective effects of adjuvanted/vectored proteins are tested alone and compared to BCG. This might not be the best strategy since most protein based subunit vaccine candidates aim to boost BCG vaccination, and there could be antigens able to improve the protective efficacy of BCG but not as much reduce the bacterial load as stand-alone vaccines. Moreover, the diversity in mouse strains, regimens, adjuvants, infection challenges and doses used differ widely and impede a comparison between different vaccine candidates. In that regard, a head-to-head comparison of vaccine candidates in the same models and experiments should be strongly promoted to provide more solid and consistent data in the preclinical stage of vaccine development. TBVI is one of the first organisations that has been promoting such a TB vaccine selection process during the past decade [177].

In conclusion, genome wide strategies have discovered a wealth of

new *Mtb* antigens and epitopes that have escaped detection by previous classical methodologies, with interesting overlaps among those identified by independent approaches (Supplementary Table 1) (Fig. 1). Unlocking their potential as vaccine targets as well as TB biomarker antigens, *e.g.* for diagnostic or prognostic purposes, will be the next fascinating challenge.

Acknowledgements

We gratefully acknowledge funding by EC HORIZON2020 TBVAC2020 (Grant Agreement No. 643381); EC ITN FP7 VACTRAIN project; EC FP7 EURIPRED (FP7-INFRA-2012 Grant Agreement No. 312661); EC-FP7 Infrastructure Project TRANSVAC2: Infrastructural project on systems analyses (Grant Agreement No. 730964); EC EDCTP through the SCREENTB project (Grant Agreement no. IP DRIA2014-311)(the text represents the authors' views and does not necessarily represent a position of the Commission who will not be liable for the use made of such information); The Netherlands Organization for Scientific Research (NWO-TOP Grant Agreement No. 91214038).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.smim.2018.07.001.

References

[1] I. Comas, M. Coscolla, T. Luo, S. Borrell, K.E. Holt, M. Kato-Maeda, et al., Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans, Nat. Genet. 45 (10) (2013) 1176–1182.
[2] K. Dheda, C.E. Barry 3rd, G. Maartens, Tuberculosis Lancet 387 (10024) (2016) 1211–1226.
[3] WHO, Global Tuberculosis Report 2017, World Health Organization, Geneva, 2017 Licence: CC BY-NCSA, 3.0 IGO.; 2017.
[4] H.M. Dockrell, S.G. Smith, What have we learnt about BCG vaccination in the last 20 years? Front. Immunol. 8 (2017) 1134.
[5] WHO, Implementing Tuberculosis Diagnostics, (2015).
[6] A.M. Cadena, S.M. Fortune, J.L. Flynn, Heterogeneity in tuberculosis, Nat. Rev. Immunol. 17 (11) (2017) 691–702.
[7] R. Rappuoli, M.J. Bottomley, U. D'Oro, O. Finco, E. De Gregorio, Reverse vaccinology 2.0: human immunology instructs vaccine antigen design, J. Exp. Med. 213 (4) (2016) 469–481.
[8] J.L. Flynn, J. Chan, K.J. Triebold, D.K. Dalton, T.A. Stewart, B.R. Bloom, An essential role for interferon gamma in resistance to Mycobacterium tuberculosis infection, J. Exp. Med. 178 (6) (1993) 2249–2254.
[9] T.H. Ottenhoff, F.A. Verreck, E.G. Lichtenauer-Kaligis, M.A. Hoeve, O. Sanal, J.T. van Dissel, Genetics, cytokines and human infectious disease: lessons from weakly pathogenic mycobacteria and salmonellae, Nat. Genet. 32 (1) (2002) 97–105.
[10] T.M. Daniel, L.E. Ferguson, Purification and Characterization of Two Proteins from Culture Filtrates of Mycobacterium tuberculosis H(37)Ra Strain.Infection and immunity, Infect. Immun. 1 (2) (1970) 164–168.
[11] H.G. Wiker, M. Harboe, The antigen 85 complex: a major secretion product of Mycobacterium tuberculosis, Microbiol. Rev. 56 (4) (1992) 648–661.
[12] Y. Fukui, T. Hirai, T. Uchida, M. Yoneda, Extracellular proteins of tubercle bacilli. IV. Alpha and beta antigens as major extracellular protein products and as cellular components of a strain (H37Rv) of Mycobacterium tuberculosis, Biken J. 8 (4) (1965) 189–199.
[13] M. Harboe, S. Nagai, M.E. Patarroyo, M.L. Torres, C. Ramirez, N. Cruz, Properties of proteins MPB64, MPB70, and MPB80 of Mycobacterium bovis BCG, Infect. Immun. 52 (1) (1986) 293–302.
[14] A.B. Andersen, Z.L. Yuan, K. Haslov, B. Vergmann, J. Bennedsen, Interspecies reactivity of five monoclonal antibodies to Mycobacterium tuberculosis as examined by immunoblotting and enzyme-linked immunosorbent assay, J. Clin. Microbiol. 23 (3) (1986) 446–451.
[15] P. Andersen, A.B. Andersen, A.L. Sorensen, S. Nagai, Recall of long-lived immunity to Mycobacterium tuberculosis infection in mice, J. Immunol. 154 (7) (1995) 3359–3372.
[16] R.N. Coler, Y.A. Skeiky, T. Vedvick, T. Bement, P. Ovendale, A. Campos-Neto, et al., Molecular cloning and immunologic reactivity of a novel low molecular mass antigen of Mycobacterium tuberculosis, J. Immunol. 161 (5) (1998) 2356–2364.
[17] F.X. Berthet, P.B. Rasmussen, I. Rosenkrands, P. Andersen, B. Gicquel, A Mycobacterium tuberculosis operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10), Microbiology 144 (Pt 11) (1998) 3195–3203.
[18] P.R. Jungblut, U.E. Schaible, H.J. Mollenkopf, U. Zimny-Arndt, B. Raupach, J. Mattow, et al., Comparative proteome analysis of Mycobacterium tuberculosis and Mycobacterium bovis BCG strains: towards functional genomics of microbial pathogens, Mol. Microbiol. 33 (6) (1999) 1103–1117.
[19] R.L. Skjot, T. Oettinger, I. Rosenkrands, P. Ravn, I. Brock, S. Jacobsen, et al., Comparative evaluation of low-molecular-mass proteins from Mycobacterium tuberculosis identifies members of the ESAT-6 family as immunodominant T-cell antigens, Infect. Immun. 68 (1) (2000) 214–220.
[20] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, et al., Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence, Nature 393 (6685) (1998) 537–544.
[21] M.I. Voskuil, D. Schnappinger, K.C. Visconti, M.I. Harrell, G.M. Dolganov, D.R. Sherman, et al., Inhibition of respiration by nitric oxide induces a Mycobacterium tuberculosis dormancy program, J. Exp. Med. 198 (5) (2003) 705–713.
[22] D. Schnappinger, S. Ehrt, M.I. Voskuil, Y. Liu, J.A. Mangan, I.M. Monahan, et al., Transcriptional adaptation of mycobacterium tuberculosis within macrophages: insights into the phagosomal environment, J. Exp. Med. 198 (5) (2003) 693–704.
[23] E.M. Leyten, M.Y. Lin, K.L. Franken, A.H. Friggen, C. Prins, K.E. van Meijgaarden, et al., Human T-cell responses to 25 novel antigens encoded by genes of the dormancy regulon of Mycobacterium tuberculosis, Microbes Infect. 8 (8) (2006) 2052–2060.
[24] G.F. Black, B.A. Thiel, M.O. Ota, S.K. Parida, R. Adegbola, W.H. Boom, et al., Immunogenicity of novel DosR regulon-encoded candidate antigens of Mycobacterium tuberculosis in three high-burden populations in Africa, Clin. Vaccine Immunol. 16 (8) (2009) 1203–1212.
[25] D. Goletti, O. Butera, V. Vanini, F.N. Lauria, C. Lange, K.L. Franken, et al., Response to Rv2628 latency antigen associates with cured tuberculosis and remote infection, Eur. Respir. J. 36 (1) (2010) 135–142.
[26] S. Rakshit, V. Adiga, S. Nayak, P.N. Sahoo, P.K. Sharma, K.E. van Meijgaarden, et al., Circulating Mycobacterium tuberculosis DosR latency antigen-specific, polyfunctional, regulatory IL10(+) Th17 CD4 T-cells differentiate latent from active tuberculosis, Sci. Rep. 7 (1) (2017) 11948.
[27] M. Coppola, L. Arroyo, K.E. van Meijgaarden, K.L. Franken, A. Geluk, L.F. Barrera, et al., Differences in IgG responses against infection phase related Mycobacterium tuberculosis (Mtb) specific antigens in individuals exposed or not to Mtb correlate with control of TB infection and progression, Tuberculosis 106 (2017) 25–32.
[28] L. Arroyo, D. Marin, K. Franken, T.H.M. Ottenhoff, L.F. Barrera, Potential of DosR and Rpf antigens from Mycobacterium tuberculosis to discriminate between latent and active tuberculosis in a tuberculosis endemic population of Medellin Colombia, BMC Infect. Dis. 18 (1) (2018) 26.
[29] M.Y. Lin, A. Geluk, S.G. Smith, A.L. Stewart, A.H. Friggen, K.L. Franken, et al., Lack of immune responses to Mycobacterium tuberculosis DosR regulon proteins following Mycobacterium bovis BCG vaccination, Infect. Immun. 75 (7) (2007) 3523–3530.
[30] V. Roupie, M. Romano, L. Zhang, H. Korf, M.Y. Lin, K.L. Franken, et al., Immunogenicity of eight dormancy regulon-encoded proteins of Mycobacterium tuberculosis in DNA-vaccinated and tuberculosis-infected mice, Infect. Immun. 75 (2) (2007) 941–949.
[31] M. Bivas-Benita, M.Y. Lin, S.M. Bal, K.E. van Meijgaarden, K.L. Franken, A.H. Friggen, et al., Pulmonary delivery of DNA encoding Mycobacterium tuberculosis latency antigen Rv1733c associated to PLGA-PEI nanoparticles enhances T cell responses in a DNA prime/protein boost vaccination regimen in mice, Vaccine 27 (30) (2009) 4010–4017.
[32] C. Shi, L. Chen, Z. Chen, Y. Zhang, Z. Zhou, J. Lu, et al., Enhanced protection against tuberculosis by vaccination with recombinant BCG over-expressing HspX protein, Vaccine 28 (32) (2010) 5237–5244.
[33] M. Coppola, S.J. van den Eeden, L. Wilson, K.L. Franken, T.H. Ottenhoff, A. Geluk, Synthetic long peptide derived from mycobacterium tuberculosis latency antigen Rv1733c protects against tuberculosis, Clin. Vaccine Immunol. 22 (9) (2015) 1060–1069.
[34] K.W. Kwon, W.S. Kim, H. Kim, S.J. Han, M.Y. Hahn, J.S. Lee, et al., Novel vaccine potential of Rv3131, a DosR regulon-encoded putative nitroreductase, against hyper-virulent Mycobacterium tuberculosis strain K, Sci. Rep. 7 (2017) 44151.
[35] A. Geluk, M.Y. Lin, K.E. van Meijgaarden, E.M. Leyten, K.L. Franken, T.H. Ottenhoff, et al., T-cell recognition of the HspX protein of Mycobacterium tuberculosis correlates with latent M. tuberculosis infection but not with M. bovis BCG vaccination, Infect. Immun. 75 (6) (2007) 2914–2921.
[36] S.T. Reece, A. Nasser-Eddine, J. Dietrich, M. Stein, U. Zedler, S. Schommer-Leitner, et al., Improved long-term protection against Mycobacterium tuberculosis Beijing/W in mice after intra-dermal inoculation of recombinant BCG expressing latency associated antigens, Vaccine 29 (47) (2011) 8740–8744.
[37] F.D. Menozzi, J.H. Rouse, M. Alavi, M. Laude-Sharp, J. Muller, R. Bischoff, et al., Identification of a heparin-binding hemagglutinin present in mycobacteria, J. Exp. Med. 184 (3) (1996) 993–1001.
[38] K. Pethe, S. Alonso, F. Biet, G. Delogu, M.J. Brennan, C. Locht, et al., The heparin-binding haemagglutinin of M. tuberculosis is required for extrapulmonary dissemination, Nature 412 (6843) (2001) 190–194.
[39] C. Masungi, S. Temmerman, J.P. Van Vooren, A. Drowart, K. Pethe, F.D. Menozzi, et al., Differential T and B cell responses against Mycobacterium tuberculosis heparin-binding hemagglutinin adhesin in infected healthy individuals and patients with tuberculosis, J. Infect. Dis. 185 (4) (2002) 513–520.
[40] G. Delogu, V. Vanini, G. Cuzzi, T. Chiacchio, F. De Maio, B. Battah, et al., Lack of response to HBHA in HIV-infected patients with latent tuberculosis infection, Scand. J. Immunol. 84 (6) (2016) 344–352.
[41] S.T. Temmerman, S. Place, A.S. Debrie, C. Locht, F. Mascart, Effector functions of heparin-binding hemagglutinin-specific CD8+ T lymphocytes in latent human

tuberculosis, J. Infect. Dis. 192 (2) (2005) 226–232.
[42] J.M. Hougardy, K. Schepers, S. Place, A. Drowart, V. Lechevin, V. Verscheure, et al., Heparin-binding-hemagglutinin-induced IFN-gamma release as a diagnostic tool for latent tuberculosis, PLoS ONE 2 (10) (2007) e926.
[43] C. Rouanet, A.S. Debrie, S. Lecher, C. Locht, Subcutaneous boosting with heparin binding haemagglutinin increases BCG-induced protection against tuberculosis, Microbes Infect. 11 (13) (2009) 995–1001.
[44] G.G. Guerrero, A.S. Debrie, C. Locht, Boosting with mycobacterial heparin-binding haemagglutinin enhances protection of Mycobacterium bovis BCG-vaccinated newborn mice against M. tuberculosis, Vaccine 28 (27) (2010) 4340–4347.
[45] S.G. Smith, S. Lecher, R. Blitz, C. Locht, H.M. Dockrell, Broad heparin-binding haemagglutinin-specific cytokine and chemokine response in infants following Mycobacterium bovis BCG vaccination, Eur. J. Immunol. 42 (9) (2012) 2511–2522.
[46] G.V. Mukamolova, A.S. Kaprelyants, D.I. Young, M. Young, D.B. Kell, A bacterial cytokine, PNAS 95 (15) (1998) 8916–8921.
[47] G.V. Mukamolova, O.A. Turapov, D.I. Young, A.S. Kaprelyants, D.B. Kell, M. Young, A family of autocrine growth factors in Mycobacterium tuberculosis, Mol. Microbiol. 46 (3) (2002) 623–635.
[48] A.P. Davies, A.P. Dhillon, M. Young, B. Henderson, T.D. McHugh, S.H. Gillespie, Resuscitation-promoting factors are expressed in Mycobacterium tuberculosis-infected human tissue, Tuberculosis 88 (5) (2008) 462–468.
[49] S.D. Schuck, H. Mueller, F. Kunitz, A. Neher, H. Hoffmann, K.L. Franken, et al., Identification of T-cell antigens specific for latent mycobacterium tuberculosis infection, PLoS ONE 4 (5) (2009) e5590.
[50] A. Geluk, K.E. van Meijgaarden, S.A. Joosten, S. Commandeur, T.H. Ottenhoff, Innovative Strategies to Identify M. tuberculosis Antigens and Epitopes Using Genome-Wide Analyses.Frontiers in immunology, Front. Immunol. 5 (2014) 256.
[51] C. Aagaard, T. Hoang, J. Dietrich, P.J. Cardona, A. Izzo, G. Dolganov, et al., A multistage tuberculosis vaccine that confers efficient protection before and after exposure, Nat. Med. 17 (2) (2011) 189–194.
[52] A. Geluk, S.J. van den Eeden, K.E. van Meijgaarden, K. Dijkman, K.L. Franken, T.H. Ottenhoff, A multistage-polyepitope vaccine protects against Mycobacterium tuberculosis infection in HLA-DR3 transgenic mice, Vaccine 30 (52) (2012) 7513–7521.
[53] S.G. Hansen, D.E. Zak, G. Xu, J.C. Ford, E.E. Marshall, D. Malouli, et al., Prevention of tuberculosis in rhesus macaques by a cytomegalovirus-based vaccine, Nat. Med. 24 (2) (2018) 130–143.
[54] S. Leung-Theung-Long, C.A. Coupet, M. Gouanvic, D. Schmitt, A. Ray, C. Hoffmann, et al., A multi-antigenic MVA vaccine increases efficacy of combination chemotherapy against Mycobacterium tuberculosis, PLoS ONE 13 (5) (2018) e0196815.
[55] M.D. Tameris, M. Hatherill, B.S. Landry, T.J. Scriba, M.A. Snowden, S. Lockhart, et al., Safety and efficacy of MVA85A, a new tuberculosis vaccine, in infants previously vaccinated with BCG: a randomised, placebo-controlled phase 2b trial, Lancet 381 (9871) (2013) 1021–1028.
[56] E. Nemes, H. Geldenhuys, V. Rozot, K.T. Rutkowski, F. Ratangee, N. Bilek, et al., Prevention of M. tuberculosis Infection with H4: IC31 Vaccine or BCG Revaccination, N. Engl. J. Med. 379 (2) (2018) 138–149.
[57] B.M. Kagina, B. Abel, T.J. Scriba, E.J. Hughes, A. Keyser, A. Soares, et al., Specific T cell frequency and cytokine expression profile do not correlate with protection against tuberculosis after bacillus Calmette-Guerin vaccination of newborns, Am. J. Respir. Crit. Care Med. 182 (8) (2010) 1073–1079.
[58] N.E. Nieuwenhuizen, S.H.E. Kaufmann, Next-generation vaccines based on bacille Calmette-Guerin, Front. Immunol. 9 (2018) 121.
[59] F. Spertini, R. Audran, R. Chakour, O. Karoui, V. Steiner-Monard, A.C. Thierry, et al., Safety of human immunisation with a live-attenuated Mycobacterium tuberculosis vaccine: a randomised, double-blind, controlled phase I trial, Lancet Respir. Med. 3 (12) (2015) 953–962.
[60] R.A. Young, B.R. Bloom, C.M. Grosskinsky, J. Ivanyi, D. Thomas, R.W. Davis, Dissection of Mycobacterium tuberculosis antigens using recombinant DNA, PNAS 82 (9) (1985) 2583–2587.
[61] R.A. Young, V. Mehra, D. Sweetser, T. Buchanan, J. Clark-Curtiss, R.W. Davis, et al., Genes for the major protein antigens of the leprosy parasite Mycobacterium leprae, Nature 316 (6027) (1985) 450–452.
[62] T.M. Shinnick, C. Krat, S. Schadow, Isolation and restriction site maps of the genes encoding five Mycobacterium tuberculosis proteins, Infect. Immun. 55 (7) (1987) 1718–1721.
[63] V. Colizzi, D. Vismara, C. D'Urso, M.F. Mezzopreti, G. Lombardi, E. Piccolella, et al., Analysis of Mycobacterium tuberculosis genoma and production of a recombinant protein containing specific B and T cell antigenic determinants–new approaches to second generation antituberculosis vaccines, Z. Erkr. Atmungsorgane 172 (1) (1989) 40–52.
[64] D.B. Young, L. Kent, R.A. Young, Screening of a recombinant mycobacterial DNA library with polyclonal antiserum and molecular weight analysis of expressed antigens, Infect. Immun. 55 (6) (1987) 1421–1425.
[65] D. Vismara, M.F. Mezzopreti, M.S. Gilardini Montani, P. Del Porto, G. Lombardi, E. Piccolella, et al., Identification of a 35-kilodalton Mycobacterium tuberculosis protein containing B- and T-cell epitopes, Infect. Immun. 58 (1) (1990) 245–251.
[66] F. Oftung, A.S. Mustafa, R. Husson, R.A. Young, T. Godal, Human T cell clones recognize two abundant Mycobacterium tuberculosis protein antigens expressed in *Escherichia coli*, J. Immunol. 138 (3) (1987) 927–931.
[67] T.H. Ottenhoff, B.K. Ab, J.D. Van Embden, J.E. Thole, R. Kiessling, The recombinant 65-kD heat shock protein of Mycobacterium bovis Bacillus Calmette-Guerin/M. tuberculosis is a target molecule for CD4+ cytotoxic T lymphocytes that lyse human monocytes, J. Exp. Med. 168 (5) (1988) 1947–1952.
[68] B. Kaleab, T. Ottenoff, P. Converse, E. Halapi, G. Tadesse, M. Rottenberg, et al., Mycobacterial-induced cytotoxic T cells as well as nonspecific killer cells derived from healthy individuals and leprosy patients, Eur. J. Immunol. 20 (12) (1990) 2651–2659.
[69] W. van Eden, E.J. Hogervorst, R. van der Zee, J.D. van Embden, E.J. Hensen, I.R. Cohen, The mycobacterial 65 kD heat-shock protein and autoimmune arthritis, Rheumatol. Int. 9 (3-5) (1989) 187–191.
[70] T.H. Ottenhoff, J.B. Haanen, A. Geluk, T. Mutis, B.K. Ab, J.E. Thole, et al., Regulation of mycobacterial heat-shock protein-reactive T cells by HLA class II molecules: lessons from leprosy, Immunol. Rev. 121 (1991) 171–191.
[71] U. Zugel, B. Schoel, S. Yamamoto, H. Hengel, B. Morein, S.H. Kaufmann, Crossrecognition by CD8 T cell receptor alpha beta cytotoxic T lymphocytes of peptides in the self and the mycobacterial hsp60 which share intermediate sequence homology, Eur. J. Immunol. 25 (2) (1995) 451–458.
[72] R.R. Amara, V. Satchidanandam, Analysis of a genomic DNA expression library of Mycobacterium tuberculosis using tuberculosis patient sera: evidence for modulation of host immune response, Infect. Immun. 64 (9) (1996) 3765–3771.
[73] D.C. Dillon, M.R. Alderson, C.H. Day, D.M. Lewinsohn, R. Coler, T. Bement, et al., Molecular characterization and human T-cell responses to a member of a novel Mycobacterium tuberculosis mtb39 gene family, Infect. Immun. 67 (6) (1999) 2941–2950.
[74] Y.A. Skeiky, M.J. Lodes, J.A. Guderian, R. Mohamath, T. Bement, M.R. Alderson, et al., Cloning, expression, and immunological evaluation of two putative secreted serine protease antigens of Mycobacterium tuberculosis, Infect. Immun. 67 (8) (1999) 3998–4007.
[75] M.R. Alderson, T. Bement, C.H. Day, L. Zhu, D. Molesh, Y.A. Skeiky, et al., Expression cloning of an immunodominant family of Mycobacterium tuberculosis antigens using human CD4(+) T cells, J. Exp. Med. 191 (3) (2000) 551–560.
[76] R.N. Coler, D.C. Dillon, Y.A. Skeiky, M. Kahn, I.M. Orme, Y. Lobet, et al., Identification of Mycobacterium tuberculosis vaccine candidates using human CD4+ T-cells expression cloning, Vaccine 27 (2) (2009) 223–233.
[77] N.I. Carlin, S. Lofdahl, M. Magnusson, Monoclonal antibodies specific for elongation factor Tu and complete nucleotide sequence of the tuf gene in Mycobacterium tuberculosis, Infect. Immun. 60 (8) (1992) 3136–3142.
[78] C. Manca, K. Lyashchenko, R. Colangeli, M.L. Gennaro, MTC28, a novel 28-kilodalton proline-rich secreted antigen specific for the Mycobacterium tuberculosis complex, Infect. Immun. 65 (12) (1997) 4951–4957.
[79] J.R. Webb, T.S. Vedvick, M.R. Alderson, J.A. Guderian, S.S. Jen, P.J. Ovendale, et al., Molecular cloning, expression, and immunogenicity of MTB12, a novel low-molecular-weight antigen secreted by Mycobacterium tuberculosis, Infect. Immun. 66 (9) (1998) 4208–4214.
[80] H.G. Wiker, S.L. Michell, R.G. Hewinson, E. Spierings, S. Nagai, M. Harboe, Cloning, expression and significance of MPT53 for identification of secreted proteins of Mycobacterium tuberculosis, Microb. Pathog. 26 (4) (1999) 207–219.
[81] Y.A. Skeiky, P.J. Ovendale, S. Jen, M.R. Alderson, D.C. Dillon, S. Smith, et al., T cell expression cloning of a Mycobacterium tuberculosis gene encoding a protective antigen associated with the early control of infection, J. Immunol. 165 (12) (2000) 7140–7149.
[82] C. Peng, L. Zhang, D. Liu, W. Xu, K. Hao, Z. Cui, et al., Mtb9.9 protein family: an immunodominant antigen family of Mycobacterium tuberculosis induces humoral and cellular immune responses in mice, Hum. Vaccin. Immunother. 8 (4) (2012) 435–442.
[83] T. Bavaro, S. Tengattini, L. Piubelli, F. Mangione, R. Bernardini, V. Monzillo, et al., Glycosylation of recombinant antigenic proteins from Mycobacterium tuberculosis: in silico prediction of protein epitopes and ex vivo biological evaluation of new semi-synthetic glycoconjugates, Molecules 22 (7) (2017).
[84] H.A. Amoudy, H.A. Safar, A.S. Mustafa, Development of Escherichia coli and Mycobacterium smegmatis recombinants expressing major Mycobacterium tuberculosis-specific antigenic proteins, Int. J. Mycobacteriol. 5 (Suppl 1) (2016) S84-s5.
[85] S.G. Reed, R.N. Coler, W. Dalemans, E.V. Tan, E.C. DeLa Cruz, R.J. Basaraba, et al., Defined tuberculosis vaccine, Mtb72F/AS02A, evidence of protection in cynomolgus monkeys, PNAS 106 (7) (2009) 2301–2306.
[86] K. Von Eschen, R. Morrison, M. Braun, O. Ofori-Anyinam, E. De Kock, P. Pavithran, et al., The candidate tuberculosis vaccine Mtb72F/AS02A: tolerability and immunogenicity in humans, Hum. Vaccin. 5 (7) (2009) 475–482.
[87] J.M. Lew, C. Mao, M. Shukla, A. Warren, R. Will, D. Kuznetsov, et al., Database resources for the tuberculosis community, Tuberculosis 93 (1) (2013) 12–17.
[88] F. Oftung, A. Geluk, K.E. Lundin, R.H. Meloen, J.E. Thole, A.S. Mustafa, et al., Mapping of multiple HLA class II-restricted T-cell epitopes of the mycobacterial 70-kilodalton heat shock protein, Infect. Immun. 62 (12) (1994) 5411–5418.
[89] A. Lalvani, R. Brookes, R.J. Wilkinson, A.S. Malin, A.A. Pathan, P. Andersen, et al., Human cytolytic and interferon gamma-secreting CD8+ T lymphocytes specific for Recombinant Antigenic Proteins from Mycobacterium tuberculosis, PNAS 95 (1) (1998) 270–275.
[90] N. Mohagheghpour, D. Gammon, L.M. Kawamura, A. van Vollenhoven, C.J. Benike, E.G. Engleman, CTL response to Mycobacterium tuberculosis: identification of an immunogenic epitope in the 19-kDa lipoprotein, J. Immunol. 161 (5) (1998) 2400–2406.
[91] A. Geluk, K.E. van Meijgaarden, K.L. Franken, J.W. Drijfhout, S. D'Souza, A. Necker, et al., Identification of major epitopes of Mycobacterium tuberculosis AG85B that are recognized by HLA-A*0201-restricted CD8+ T cells in HLA-transgenic mice and humans, J. Immunol. 165 (11) (2000) 6463–6471.
[92] D.M. Lewinsohn, L. Zhu, V.J. Madison, D.C. Dillon, S.P. Fling, S.G. Reed, et al., Classically restricted human CD8+ T lymphocytes derived from Mycobacterium tuberculosis-infected cells: definition of antigenic specificity, J. Immunol. 166 (1)

(2001) 439–446.
[93] D.A. Lewinsohn, E. Winata, G.M. Swarbrick, K.E. Tanner, M.S. Cook, M.D. Null, et al., Immunodominant tuberculosis CD8 antigens preferentially restricted by HLA-B, PLoS Pathog. 3 (9) (2007) 1240–1249.
[94] A.S. Mustafa, Th1 cell reactivity and HLA-DR binding prediction for promiscuous recognition of MPT63 (Rv1926c), a major secreted protein of Mycobacterium tuberculosis, Scand. J. Immunol. 69 (3) (2009) 213–222.
[95] M.A. Behr, M.A. Wilson, W.P. Gill, H. Salamon, G.K. Schoolnik, S. Rane, et al., Comparative genomics of BCG vaccines by whole-genome DNA microarray, Science 284 (5419) (1999) 1520–1523.
[96] I. Brock, K. Weldingh, E.M. Leyten, S.M. Arend, P. Ravn, P. Andersen, Specific T-cell epitopes for immunoassay-based diagnosis of Mycobacterium tuberculosis infection, J. Clin. Microbiol. 42 (6) (2004) 2379–2387.
[97] A.S. Mustafa, R. Al-Attiyah, S.N. Hanif, F.A. Shaban, Efficient testing of large pools of Mycobacterium tuberculosis RD1 peptides and identification of major antigens and immunodominant peptides recognized by human Th1 cells, Clin. Vaccine Immunol. 15 (6) (2008) 916–924.
[98] O. Parkash, B.P. Singh, M. Pai, Regions of differences encoded antigens as targets for immunodiagnosis of tuberculosis in humans, Scand. J. Immunol. 70 (4) (2009) 345–357.
[99] M. Ruhwald, L. de Thurah, D. Kuchaka, M.R. Zaher, A.M. Salman, A.R. Abdel-Ghaffar, et al., Introducing the ESAT-6 free IGRA, a companion diagnostic for TB vaccines based on ESAT-6, Sci. Rep. 7 (2017) 45969.
[100] N.P. Knudsen, S. Norskov-Lauritsen, G.M. Dolganov, G.K. Schoolnik, T. Lindenstrom, P. Andersen, et al., Tuberculosis vaccine with high predicted population coverage and compatibility with modern diagnostics, PNAS 111 (3) (2014) 1096–1101.
[101] D.M. Lewinsohn, G.M. Swarbrick, M.E. Cansler, M.D. Null, V. Rajaraman, M.M. Frieder, et al., Human Mycobacterium tuberculosis CD8 T cell Antigens/Epitopes identified by a proteomic peptide library, PLoS ONE 8 (6) (2013) e67016.
[102] D.A. Lewinsohn, G.M. Swarbrick, B. Park, M.E. Cansler, M.D. Null, K.G. Toren, et al., Comprehensive definition of human immunodominant CD8 antigens in tuberculosis, NPJ Vaccines 2 (2017).
[103] Y. Li, J. Zeng, J. Shi, M. Wang, M. Rao, C. Xue, et al., A proteome-scale identification of novel antigenic proteins in Mycobacterium tuberculosis toward diagnostic and vaccine development, J. Proteome Res. 9 (9) (2010) 4812–4822.
[104] C.S. Lindestam Arlehamn, A. Gerasimova, F. Mele, R. Henderson, J. Swann, J.A. Greenbaum, et al., Memory T cells in latent Mycobacterium tuberculosis infection are directed against three antigenic islands and largely contained in a CXCR3+CCR6+ Th1 subset, PLoS Pathog. 9 (1) (2013) e1003130.
[105] E. Stylianou, R. Harrington-Kandt, J. Beglov, N. Bull, N. Pinpathomrat, G.M. Swarbrick, et al., Identification and evaluation of novel protective antigens for the development of a candidate TB subunit vaccine, Infect. Immun. (2018).
[106] A. Sette, R. Rappuoli, Reverse vaccinology: developing vaccines in the era of genomics, Immunity 33 (4) (2010) 530–541.
[107] J. Liepe, F. Marino, J. Sidney, A. Jeko, D.E. Bunting, A. Sette, et al., A large fraction of HLA class I ligands are proteasome-generated spliced peptides, Science 354 (6310) (2016) 354–358.
[108] M.R. Klein, S.M. Smith, A.S. Hammond, G.S. Ogg, A.S. King, J. Vekemans, et al., HLA-B*35-restricted CD8 T cell epitopes in the antigen 85 complex of Mycobacterium tuberculosis, J. Infect. Dis. 183 (6) (2001) 928–934.
[109] M.R. Klein, A.S. Hammond, S.M. Smith, A. Jaye, P.T. Lukey, K.P. McAdam, HLA-B*35-restricted CD8(+)-T-cell epitope in Mycobacterium tuberculosis Rv2903c, Infect. Immun. 70 (2) (2002) 981–984.
[110] M. Vordermeier, A.O. Whelan, R.G. Hewinson, Recognition of mycobacterial epitopes by T cells across mammalian species and use of a program that predicts human HLA-DR binding peptides to predict bovine epitopes, Infect. Immun. 71 (4) (2003) 1980–1987.
[111] A.S. Mustafa, A.T. Abal, F. Shaban, A.M. El-Shamy, H.A. Amoudy, HLA-DR binding prediction and experimental evaluation of T-cell epitopes of mycolyl transferase 85B (Ag85B), a major secreted antigen of Mycobacterium tuberculosis, Medical principles and practice 14 (3) (2005) 140–146.
[112] M. Wang, S.T. Tang, A. Stryhn, S. Justesen, M.V. Larsen, M.H. Dziegiel, et al., Identification of MHC class II restricted T-cell-mediated reactivity against MHC class I binding Mycobacterium tuberculosis peptides, Immunology 132 (4) (2011) 482–491.
[113] A.S. De Groot, J. McMurry, L. Marcon, J. Franco, D. Rivera, M. Kutzler, et al., Developing an epitope-driven tuberculosis (TB) vaccine, Vaccine 23 (17–18) (2005) 2121–2131.
[114] S.T. Tang, K.E. van Meijgaarden, N. Caccamo, G. Guggino, M.R. Klein, P. van Weeren, et al., Genome-based in silico identification of new Mycobacterium tuberculosis antigens activating polyfunctional CD8+ T cells in human tuberculosis, J. Immunol. 186 (2) (2011) 1068–1080.
[115] A.S. Hammond, M.R. Klein, T. Corrah, A. Fox, A. Jaye, K.P. McAdam, et al., Mycobacterium tuberculosis genome-wide screen exposes multiple CD8 T cell epitopes, Clin. Exp. Immunol. 140 (1) (2005) 109–116.
[116] N. Caccamo, G. Guggino, S. Meraviglia, G. Gelsomino, P. Di Carlo, L. Titone, et al., Analysis of Mycobacterium tuberculosis-specific CD8 T-cells in patients with active tuberculosis and in individuals with latent infection, PloS one 4 (5) (2009) e5528.
[117] A. Sette, J. Sidney, HLA supertypes and supermotifs: a functional perspective on HLA polymorphism, Curr. Opin. Immunol. 10 (4) (1998) 478–482.
[118] J. Greenbaum, J. Sidney, J. Chung, C. Brander, B. Peters, A. Sette, Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes, Immunogenetics 63 (6) (2011) 325–335.
[119] C.L. Arlehamn, G. Seumois, A. Gerasimova, C. Huang, Z. Fu, X. Yue, et al., Transcriptional profile of tuberculosis antigen-specific T cells reveals novel multifunctional features, J. Immunol. 193 (6) (2014) 2931–2940.
[120] C. Carpenter, J. Sidney, R. Kolla, K. Nayak, H. Tomiyama, C. Tomiyama, et al., A side-by-side comparison of T cell reactivity to fifty-nine Mycobacterium tuberculosis antigens in diverse populations from five continents, Tuberculosis 95 (6) (2015) 713–721.
[121] W.E. Whatney, N.R. Gandhi, C.S. Lindestam Arlehamn, A. Nizam, H. Wu, M.J. Quezada, et al., A High throughput whole blood assay for analysis of multiple antigen-specific T cell responses in human Mycobacterium tuberculosis infection, J. Immunol. 200 (8) (2018) 3008–3019.
[122] C.S. Lindestam Arlehamn, D. Lewinsohn, A. Sette, D. Lewinsohn, Antigens for CD4 and CD8 T cells in tuberculosis, Cold Spring Harb. Perspect. Med. 4 (7) (2014) a018465.
[123] C.S. Lindestam Arlehamn, D.M. McKinney, C. Carpenter, S. Paul, V. Rozot, E. Makgotlho, et al., A quantitative analysis of complexity of human pathogen-specific CD4 T cell responses in healthy M. Tuberculosis infected South Africans, PLoS Pathog. 12 (7) (2016) e1005760.
[124] T.J. Scriba, C. Carpenter, S.C. Pro, J. Sidney, M. Musvosvi, V. Rozot, et al., Differential recognition of Mycobacterium tuberculosis-specific epitopes as a function of tuberculosis disease history, Am. J. Respir. Crit. Care Med. 196 (6) (2017) 772–781.
[125] C.S. Lindestam Arlehamn, S. Paul, F. Mele, C. Huang, J.A. Greenbaum, R. Vita, et al., Immunological consequences of intragenus conservation of Mycobacterium tuberculosis T-cell epitopes, PNAS 112 (2) (2015) E147–55.
[126] M.Y. Lin, T.B. Reddy, S.M. Arend, A.H. Friggen, K.L. Franken, K.E. van Meijgaarden, et al., Cross-reactive immunity to Mycobacterium tuberculosis DosR regulon-encoded antigens in individuals infected with environmental, non-tuberculous mycobacteria, Infect. Immun. 77 (11) (2009) 5071–5079.
[127] L. Brandt, Cunha J. Feino, Olsen A. Weinreich, B. Chilima, P. Hirsch, R. Appelberg, et al., Failure of the Mycobacterium bovis BCG vaccine: some species of environmental mycobacteria block multiplication of BCG and induction of protective immunity to tuberculosis, Infect. Immun. 70 (2) (2002) 672–678.
[128] A.S. Heinzel, J.E. Grotzke, R.A. Lines, D.A. Lewinsohn, A.L. McNabb, D.N. Streblow, et al., HLA-E-dependent presentation of Mtb-derived antigen to human CD8+ T cells, J. Exp. Med. 196 (11) (2002) 1473–1481.
[129] S.A. Joosten, K.E. van Meijgaarden, P.C. van Weeren, F. Kazi, A. Geluk, N.D. Savage, et al., Mycobacterium tuberculosis peptides presented by HLA-E molecules are targets for human CD8 T-cells with cytotoxic as well as regulatory activity, PLoS Pathog. 6 (2) (2010) e1000782.
[130] K.E. van Meijgaarden, M.C. Haks, N. Caccamo, F. Dieli, T.H. Ottenhoff, S.A. Joosten, Human CD8+ T-cells recognizing peptides from Mycobacterium tuberculosis (Mtb) presented by HLA-E have an unorthodox Th2-like, multifunctional, Mtb inhibitory phenotype and represent a novel human T-cell subset, PLoS Pathog. 11 (3) (2015) e1004671.
[131] N. Caccamo, G. Pietra, L.C. Sullivan, A.G. Brooks, T. Prezzemolo, M.P. La Manna, et al., Human CD8 T lymphocytes recognize Mycobacterium tuberculosis antigens presented by HLA-E during active tuberculosis and express type 2 cytokines, Eur. J. Immunol. 45 (4) (2015) 1069–1081.
[132] T. Prezzemolo, K.E. van Meijgaarden, K. Franken, N. Caccamo, F. Dieli, T.H.M. Ottenhoff, et al., Detailed characterization of human Mycobacterium tuberculosis specific HLA-E restricted CD8(+) T cells, Eur. J. Immunol. 48 (2) (2018) 293–305.
[133] S.A. Joosten, L.C. Sullivan, T.H. Ottenhoff, Characteristics of HLA-E restricted T-cell responses and their role in infectious diseases, J. Immunol. Res. 2016 (2016) 2695396.
[134] Y. Bian, S. Shang, S. Siddiqui, J. Zhao, S.A. Joosten, T.H.M. Ottenhoff, et al., MHC Ib molecule Qa-1 presents Mycobacterium tuberculosis peptide antigens to CD8+ T cells and contributes to protection against infection, PLoS Pathog. 13 (5) (2017) e1006384.
[135] D.I. Godfrey, A.P. Uldrich, J. McCluskey, J. Rossjohn, D.B. Moody, The burgeoning family of unconventional T cells, Nat. Immunol. 16 (11) (2015) 1114–1123.
[136] I. Comas, J. Chakravartti, P.M. Small, J. Galagan, S. Niemann, K. Kremer, et al., Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved, Nat. Genet. 42 (6) (2010) 498–503.
[137] J.S. Woodworth, P. Andersen, Reprogramming the T cell response to tuberculosis, Trends Immunol. 37 (2) (2016) 81–83.
[138] I.M. Orme, R.T. Robinson, A.M. Cooper, The balance between protective and pathogenic immune responses in the TB-infected lung, Nat. Immunol. 16 (1) (2015) 57–63.
[139] A.W. Olsen, P.R. Hansen, A. Holm, P. Andersen, Efficient protection against Mycobacterium tuberculosis by vaccination with a single subdominant epitope from the ESAT-6 antigen, Eur. J. Immunol. 30 (6) (2000) 1724–1732.
[140] C.S. Aagaard, T.T. Hoang, C. Vingsbo-Lundberg, J. Dietrich, P. Andersen, Quality and vaccine efficacy of CD4+ T cell responses directed to dominant and subdominant epitopes in ESAT-6 from Mycobacterium tuberculosis, J. Immunol. 183 (4) (2009) 2659–2668.
[141] J.S. Woodworth, C.S. Aagaard, P.R. Hansen, J.P. Cassidy, E.M. Agger, P. Andersen, Protective CD4 T cells targeting cryptic epitopes of Mycobacterium tuberculosis resist infection-driven terminal differentiation, J. Immunol. 192 (7) (2014) 3247–3258.
[142] M.T. Orr, G.C. Ireton, E.A. Beebe, P.W. Huang, V.A. Reese, D. Argilla, et al., Immune subdominant antigens as vaccine candidates against mycobacterium tuberculosis, J. Immunol. 193 (6) (2014) 2911–2918.
[143] K.B. Urdahl, Understanding and overcoming the barriers to T cell-mediated immunity against tuberculosis, Semin. Immunol. 26 (6) (2014) 578–587.

[144] M. Coscolla, R. Copin, J. Sutherland, F. Gehre, B. de Jong, O. Owolabi, et al., M. Tuberculosis T cell epitope analysis reveals paucity of antigenic variation and identifies rare variable TB antigens, Cell Host Microbe 18 (5) (2015) 538–548.
[145] S. Bertholet, G.C. Ireton, M. Kahn, J. Guderian, R. Mohamath, N. Stride, et al., Identification of human T cell antigens for the development of vaccines against Mycobacterium tuberculosis, J. Immunol. 181 (11) (2008) 7948–7957.
[146] A. Penn-Nicholson, M. Tameris, E. Smit, T.A. Day, M. Musvosvi, L. Jayashankar, et al., Safety and immunogenicity of the novel tuberculosis vaccine ID93 + GLA-SE in BCG-vaccinated healthy adults in South Africa: a randomised, double-blind, placebo-controlled phase 1 trial, Lancet Respir. Med. 6 (4) (2018) 287–298.
[147] M. Coppola, S.J.F. van den Eeden, N. Robbins, L. Wilson, K. Franken, L.B. Adams, et al., Vaccines for leprosy and tuberculosis: opportunities for shared research, development, and application, Front. Immunol. 9 (2018) 308.
[148] S. Bertholet, G.C. Ireton, D.J. Ordway, H.P. Windish, S.O. Pine, M. Kahn, et al., A defined tuberculosis vaccine candidate boosts BCG and protects against multidrug-resistant Mycobacterium tuberculosis, Sci. Transl. Med. 2 (53) (2010) 53ra74.
[149] S. Kunnath-Velayudhan, H. Salamon, H.Y. Wang, A.L. Davidow, D.M. Molina, V.T. Huynh, et al., Dynamic antibody responses to the Mycobacterium tuberculosis proteome, PNAS 107 (33) (2010) 14703–14708.
[150] WHO, WHO Warns Against the Use of Inaccurate Blood Tests for Active Tuberculosis, (2011).
[151] K. Nayak, L. Jing, R.M. Russell, D.H. Davies, G. Hermanson, D.M. Molina, et al., Identification of novel Mycobacterium tuberculosis CD4 T-cell antigens via high throughput proteome screening, Tuberculosis 95 (3) (2015) 275–287.
[152] J.D. Ernst, The immunological life cycle of tuberculosis, Nat. Rev. Immunol. 12 (8) (2012) 581–591.
[153] S. Commandeur, K.E. van Meijgaarden, C. Prins, A.V. Pichugin, K. Dijkman, S.J. van den Eeden, et al., An unbiased genome-wide Mycobacterium tuberculosis gene expression approach to discover antigens targeted by human T cells expressed during pulmonary infection, J. Immunol. 190 (4) (2013) 1659–1671.
[154] S. Commandeur, S.J. van den Eeden, K. Dijkman, S.O. Clark, K.E. van Meijgaarden, L. Wilson, et al., The in vivo expressed Mycobacterium tuberculosis (IVE-TB) antigen Rv2034 induces CD4(+) T-cells that protect against pulmonary infection in HLA-DR transgenic mice and guinea pigs, Vaccine 32 (29) (2014) 3580–3588.
[155] M. Coppola, K.E. van Meijgaarden, K.L. Franken, S. Commandeur, G. Dolganov, I. Kramnik, et al., New genome-wide algorithm identifies novel in-vivo expressed Mycobacterium tuberculosis antigens inducing human T-cell responses with classical and unconventional cytokine profiles, Sci. Rep. 6 (2016) 37793.
[156] S.A. Joosten, K.E. van Meijgaarden, S.M. Arend, C. Prins, F. Oftung, G.E. Korsvold, et al., Mycobacterial growth inhibition is associated with trained innate immunity, J. Clin. Invest. 128 (5) (2018) 1837–1851.
[157] M. Gonzalez-Juarrero, J.M. Hattle, A. Izzo, A.P. Junqueira-Kipnis, T.S. Shim, B.C. Trapnell, et al., Disruption of granulocyte macrophage-colony stimulating factor production in the lungs severely affects the ability of mice to control Mycobacterium tuberculosis infection, J. Leukoc. Biol. 77 (6) (2005) 914–922.
[158] J. Dou, Q. Tang, F. Yu, H. Yang, F. Zhao, W. Xu, et al., Investigation of immunogenic effect of the BCG priming and Ag85A- GM-CSF boosting in Balb/c mice model, Immunobiology 215 (2) (2010) 133–142.
[159] A.C. Rothchild, P. Jayaraman, C. Nunes-Alves, S.M. Behar, iNKT cell production of GM-CSF controls Mycobacterium tuberculosis, PLoS Pathog. 10 (1) (2014) e1003805.
[160] S. Sharma, M.B. Ryndak, A.N. Aggarwal, R. Yadav, S. Sethi, S. Masih, et al., Transcriptome analysis of mycobacteria in sputum samples of pulmonary tuberculosis patients, PloS one 12 (3) (2017) e0173508.
[161] I. Kramnik, G. Beamer, Mouse models of human TB pathology: roles in the analysis of necrosis and the development of host-directed therapies, Semin. Immunopathol. 38 (2) (2016) 221–237.
[162] M.M. Serra-Vidal, I. Latorre, K.L. Franken, J. Diaz, M.L. de Souza-Galvao, I. Casas, et al., Immunogenicity of 60 novel latency-related antigens of Mycobacterium tuberculosis, Front. Microbiol. 5 (2014) 517.
[163] A.O. Moguche, M. Musvosvi, A. Penn-Nicholson, C.R. Plumlee, H. Mearns, H. Geldenhuys, et al., Antigen availability shapes T cell differentiation and function during tuberculosis, Cell Host Microbe 21 (6) (2017) 695–706.e5.
[164] T. Lindenstrom, A. Moguche, M. Damborg, E.M. Agger, K. Urdahl, P. Andersen, T cells primed by live mycobacteria versus a tuberculosis subunit vaccine exhibit distinct functional properties, EBioMedicine 27 (2018) 27–39.
[165] M. Pai, M.A. Behr, D. Dowdy, K. Dheda, M. Divangahi, C.C. Boehme, et al., Tuberculosis. Nature reviews Disease primers. 2 (2016) 16076.
[166] D.E. Zak, A. Penn-Nicholson, T.J. Scriba, E. Thompson, S. Suliman, L.M. Amon, et al., A blood RNA signature for tuberculosis disease risk: a prospective cohort study, Lancet 387 (10035) (2016) 2312–2322.
[167] S. Suliman, E. Thompson, J. Sutherland, Rd J. Weiner, M.O.C. Ota, S. Shankar, et al., Four-gene Pan-African blood signature predicts progression to tuberculosis, Am. J. Respir. Crit. Care Med. (2018).
[168] J.L. Flynn, H.P. Gideon, J.T. Mattila, P.L. Lin, Immunology studies in non-human primate models of tuberculosis, Immunol. Rev. 264 (1) (2015) 60–73.
[169] E.G. Thompson, S. Shankar, H.P. Gideon, J. Braun, J. Valvo, J.A. Skinner, et al., Prospective discrimination of controllers from progressors early after Low-dose Mycobacterium tuberculosis infection of cynomolgus macaques using blood RNA signatures, J. Infect. Dis. 217 (8) (2018) 1318–1322.
[170] B.R. Mothe, C.S. Lindestam Arlehamn, C. Dow, M.B.C. Dillon, R.W. Wiseman, P. Bohn, et al., The TB-specific CD4(+) T cell immune repertoire in both cynomolgus and rhesus macaques largely overlap with humans, Tuberculosis 95 (6) (2015) 722–735.
[171] J.R. Andrews, F. Noubary, R.P. Walensky, R. Cerda, E. Losina, C.R. Horsburgh, Risk of progression to active tuberculosis following reinfection with Mycobacterium tuberculosis, Clin. Infect. Dis. 54 (6) (2012) 784–791.
[172] H.A. Fletcher, M.A. Snowden, B. Landry, W. Rida, I. Satti, S.A. Harris, et al., T-cell activation is an immune correlate of risk in BCG vaccinated infants, Nat. Commun. 7 (2016) 11290.
[173] J.D. Ernst, A. Cornelius, L. Desvignes, J. Tavs, B.A. Norris, Limited anti-mycobacterial efficacy of epitope peptide administration despite enhanced antigen-specific CD4 T cell activation, J. Infect. Dis. (2018).
[174] T.H. Ottenhoff, New pathways of protective and pathological host defense to mycobacteria, Trends Microbiol. 20 (9) (2012) 419–428.
[175] S. Srivastava, P.S. Grace, J.D. Ernst, antigen export reduces Antigen presentation and limits T cell control of M. Tuberculosis, Cell Host Microbe 19 (1) (2016) 44–54.
[176] P.S. Grace, J.D. Ernst, Suboptimal antigen presentation contributes to virulence of Mycobacterium tuberculosis in vivo, J. Immunol. 196 (1) (2016) 357–364.
[177] S.H.E. Kaufmann, H.M. Dockrell, N. Drager, M.M. Ho, H. McShane, O. Neyrolles, et al., TBVAC2020: advancing tuberculosis vaccines from discovery to clinical development, Front. Immunol. 8 (2017) 1203.