An Unbiased Genome-Wide *Mycobacterium tuberculosis* Gene Expression Approach To Discover Antigens Targeted by Human T Cells Expressed during Pulmonary Infection


*J Immunol* published online 14 January 2013
http://www.jimmunol.org/content/early/2013/01/12/jimmunol.1201593

---

**Supplementary Material**
http://www.jimmunol.org/content/suppl/2013/01/15/jimmunol.1201593.DC1

**Subscription**
Information about subscribing to *The Journal of Immunology* is online at:
http://jimmunol.org/subscription

**Permissions**
Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**
Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
An Unbiased Genome-Wide *Mycobacterium tuberculosis* Gene Expression Approach To Discover Antigens Targeted by Human T Cells Expressed during Pulmonary Infection

Susanna Commandeur,* Krista E. van Meijgaarden,* Corine Prins,* Alexander V. Pichugin,†,§, Karin Dijkman,* Susan J. F. van den Eeden,* Annemieke H. Friggen,* Kees L. M. C. Franken,* Gregory Dolganov,‡ Igor Kramnik,†,2 Gary K. Schoolnik,‡ Fredrik Oftung,§ Gro Ellen Korsvold,§ Annemieke Geluk,* and Tom H. M. Ottenhoff*  

*Mycobacterium tuberculosis* is responsible for almost 2 million deaths annually. *Mycobacterium bovis* bacillus Calmette-Guérin, the only vaccine available against tuberculosis (TB), induces highly variable protection against TB, and better TB vaccines are urgently needed. A prerequisite for candidate vaccine Ags is that they are immunogenic and expressed by *M. tuberculosis* during infection of the primary target organ, that is, the lungs of susceptible individuals. In search of new TB vaccine candidate Ags, we have used a genome-wide, unbiased Ag discovery approach to investigate the in vivo expression of 2170 *M. tuberculosis* genes during *M. tuberculosis* infection in the lungs of mice. Four genetically related but distinct mouse strains were studied, representing a spectrum of TB susceptibility controlled by the supersusceptibility to TB 1 locus. We used stringent selection approaches to select in vivo–expressed *M. tuberculosis* (IVE-TB) genes and analyzed their expression patterns in distinct disease phenotypes such as necrosis and granuloma formation. To study the vaccine potential of these proteins, we analyzed their immunogenicity. Several *M. tuberculosis* proteins were recognized by immune cells from tuberculin skin test-positive, ESAT6/CFP10–responsive individuals, indicating that these Ags are presented during natural *M. tuberculosis* infection. Furthermore, TB patients also showed responses toward IVE-TB Ags, albeit lower than tuberculin skin test-positive, ESAT6/CFP10–responsive individuals. Finally, IVE-TB Ags induced strong IFN-γ+/TNF-α+/CD8+ and TNF-α+/IL-2+ CD154+/CD4+ T cell responses in PBMC from long-term latently *M. tuberculosis*–infected individuals. In conclusion, these IVE-TB Ags are expressed during pulmonary infection in vivo, are immunogenic, induce strong T cell responses in long-term latently *M. tuberculosis*–infected individuals, and may therefore represent attractive Ags for new TB vaccines. The *Journal of Immunology*, 2013, 190: 000–000.

Tuberculosis (TB) remains a leading cause of death, particularly in low and middle income countries (1). One third of the world population is estimated to be latently infected with *Mycobacterium tuberculosis*, and 3–10% of these will develop active TB during their lifetime. In HIV-infected individuals this proportion increases to 7–10% per life year. The emergence of multidrug-resistant, extensively drug-resistant, and more recently, protease-resistant TB, strains is further aggravating the TB epidemic. Currently, *Mycobacterium bovis* bacillus Calmette-Guérin (BCG) is the only available vaccine against TB. Although BCG vaccination can prevent severe childhood TB (2), it induces highly variable and inconsistent protection against pulmonary TB, the contagious form of TB in adults (3). A more recently identified drawback of live BCG vaccination is the occurrence of disseminating BCG infections in HIV-infected children (4), similar to severe BCG infections in individuals with genetic defects in the IL-12/IL-23/IFN-γ axis (5, 6). Thus, new TB vaccines are needed that are more effective and safer than BCG.

Understanding the intracellular behavior of *M. tuberculosis* during in vivo infection is important not only for understanding its...
infection biology, but it is also essential for the identification of possible novel TB vaccine candidate Ags. Infection stage and site-related differences in in vivo *M. tuberculosis* gene expression patterns can clearly affect the repertoire of potential *M. tuberculosis* Ags that is available for immune recognition in the primary infected organ, the lung. Ags expressed in the lungs of *M. tuberculosis*–infected, susceptible individuals could represent interesting new candidate Ags for TB vaccination, because they would induce responses capable of recognizing in situ *M. tuberculosis*–infected cells.

*M. tuberculosis* has a remarkable ability to adapt to environmental changes by altering its metabolic state. A major environmental stress factor that *M. tuberculosis* is thought to encounter during host infection is the deprivation of oxygen and nutrients. In vitro hypoxia induces the expression of the *M. tuberculosis* dormancy regulon (7), which is controlled by the master regulator DosR (*Rv3133c*). The expression of the *M. tuberculosis* DosR regulon is also induced by low-dose NO, carbon monoxide exposure, and during infection in IFN-γ–activated macrophages (7, 8). Previously, we have reported broad human T cell responses to *M. tuberculosis* DosR regulon–encoded Ags and showed that responses to these Ags were prominent and associated with latent TB infection (LTBI) in ethnically and geographically distinct populations (9–13). Other work has shown that nutrient limitation can induce the expression of specific *M. tuberculosis* genes such as *Rv2660c* (14). This gene was found to encode a “starvation” Ag with promising long-term vaccine efficacy in preclinical TB infection models, both in mice (15) and in nonhuman primates (16). The more recently described enduring hypoxic response (EHR) genes represent an alternative hypoxia-induced response model, which includes most of the DosR regulon–encoded genes complemented with an additional number (>200) of *M. tuberculosis* stress response genes (17). This model has also helped to identify new *M. tuberculosis* Ags (18).

A limitation of the models discussed above is that the identification of differentially regulated *M. tuberculosis* genes relies on in vitro models supposed to recapitulate relevant environmental stress conditions that *M. tuberculosis* encounters upon host infection. First, however, many of these environmental stress factors may not be known as yet, limiting the value of such hypothesis-driven studies. Second, there may be additive or synergistic effects between multiple stress factors in vivo that may easily be missed when studied in isolation in vitro. Third, and perhaps more importantly, certain key features of host response–induced stress cannot readily be recapitulated in vitro, including granuloma formation and TB necrosis, both being cardinal features of TB. To overcome these limitations several laboratories have started to analyze the gene expression profiles of intracellular *M. tuberculosis*, either in infected human or murine macrophages (8, 19), in the infected tissue of different mouse strains (BALB/c, SCID) (20), or in artificial granuloma mouse models (21). However, none of these mouse models developed granulomatous necrotic TB lesions (22). We therefore have studied *M. tuberculosis* genomewide gene expression patterns in mice strains carrying different genotypes of the supersusceptibility to TB 1 (*sst1*) locus. This genetic locus is located on chromosome 1 and controls the progression of *M. tuberculosis* infection to severe and necrotic lesions in a lung-specific manner: C3HeB/FeJ (C3H) mice carrying the susceptible *sst1* allele develop TB pneumonia with strong inflammatory responses with exudation throughout the lung and early onset of massive necrosis, whereas C57BL/6J (B6) mice carrying the resistant *sst1* allele develop smaller, interstitial granulomas without necrotic lesions that control bacterial multiplication. C3H.B6-sst1 congenic mice carrying the (B6-derived) resistant *sst1* locus on the C3H background showed increased survival after *M. tuberculosis* infection compared with the susceptible C3H mice, but less prominently than did the resistant B6 mice. Finally, *M. tuberculosis*–infected B6.C3H-sst1 mice, carrying the susceptible C3H-sst1 locus on the B6 background, develop robust granulomas that are fenced from the healthy tissue where lesions contain foamy macrophages and develop late-onset necrosis, resembling pulmonary TB in human adults. In contrast, the B6 strain does not display this phenotype, confirming the specific role for *sst1* in the control of cell death (23).

The *sst1* locus carries 22 genes, 1 of which was highly expressed in *M. tuberculosis*–infected lungs of C3H.B6-sst1 but not of hypersusceptible C3H mice. Interestingly, the expression of this gene, termed intracellular pathogen resistance 1 (*lpr1*), decreased *M. tuberculosis* multiplication in susceptible macrophages and induced a switch from necrotic to apoptotic cell death (24). The lack of *lpr1* expression in C3H-susceptible *sst1* locus is therefore responsible for the development of lung-specific necrosis upon *M. tuberculosis* infection (25). The closest human homolog of *lpr1* is SP110b. The expression of both *lpr1* and SP110b is regulated by IFNs, indicating a role in immunity (26–28). Genetic association studies performed in West Africa identified three polymorphisms in the SP110b gene that were associated with genetic susceptibility to TB (29). However, a number of other studies performed in Ghana, Russia, South Africa, and Indonesia did not replicate this finding (30–33). A SP110b homolog was also identified in cattle, which correlated to susceptibility to *Mycobacterium avium* ssp. *paratuberculosis* (34).

These four (congenic) mouse models we have used in this study show a spectrum of TB susceptibility that ranges from highly susceptible (C3H) to resistant (B6) mice, with the development of necrotic lesions depending on the *sst1* locus and the modifying genetic background in which the locus is expressed. This mouse model replicates key features of human *M. tuberculosis* infection. In this study, we have taken advantage of this disease spectrum and 1) analyzed quantitative real-time expression patterns of all *M. tuberculosis* genes predicted to be the first gene in each operon, in the lungs of *M. tuberculosis*–infected mice, aiming to identify the *M. tuberculosis* genes that are highly or differentially expressed in the lung during in vivo infection (in vivo–expressed *Mycobacterium tuberculosis* [IVE-TB] genes); 2) compared these *M. tuberculosis* gene expression patterns between susceptible (B6.C3H-sst1 and C3H) and resistant (C3H.B6-sst1 and B6) mouse strains in an attempt to correlate expression patterns to infection phenotype; and 3) selected a set of the most consistently expressed *M. tuberculosis* genes, produced these as recombinant proteins, and analyzed their immunogenicity in tuberculin skin test (TST) healthy, TB-affected individuals as well as long-term LTBI (H1LTBI) as a first step toward their validation as new TB vaccine candidate Ags.

### Materials and Methods

#### Mouse strains

C3H and B6 mice were purchased from The Jackson Laboratory (Bar Harbor, ME). Congenic C3H.B6-sst1 and B6.C3H-sst1 mouse strains carrying the resistant and susceptible alleles of the *sst1* locus, respectively, were generated as previously described (24, 35). Briefly, an ∼20-cM segment of chromosome 1, containing the *sst1* locus, was introgressed in the opposite background strain via ≥10 backcrosses. Mice were bred and housed under specific pathogen-free conditions at the Harvard Medical School of Public Health.

#### Bacterial strains

*M. tuberculosis* suspensions were used as previously described (36). In short, *M. tuberculosis* (Erdman strain; Trudeau Institute, Saranac Lake,
NY) cultures were grown to midlog phase in Middlebrook 7H9 medium (BD Biosciences, Franklin Lakes, NJ) (10% oleic acid/albumin/dextrose/catalase [OADC; Difco]; 0.05% Tween 80 [Sigma-Aldrich], and 0.5% glycerol [Sigma-Aldrich]). Bacteria were washed and stored at −80°C. Prior to infection, bacteria were thawed, sonicated, and diluted in PBS to 10^5 CFU/ml.

**Infection of mice**

Mice were infected by aerosol with 25–50 CFU *M. tuberculosis* using a Madison chamber (College of Engineering shops at the University of Wisconsin–Madison, WI) with n = 2 per time point (24). B6 and C3H mice were sacrificed both 6 and 9 wk postinfection, whereas B6.C3H-sstl and C3H.B6-sstl mice were sacrificed 9 and 6 wk postinfection, respectively. For the reactivation model, B6 and B6.C3H-sstl mice were infected i.v. via the tail vein with 5 × 10^3 CFU *M. tuberculosis* per mouse as previously described (23). Twelve weeks after challenge the mice were given isoniazid (INH) supplied via the drinking water (10 mg/100 ml) for 90 d. Mice were sacrificed 8 wk after INH treatment withdrawal.

**Genome-wide *M. tuberculosis* transcription profiling via a two-step multiplex real-time RT-PCR**

Quantification of *M. tuberculosis* mRNA gene expression was performed as previously described (15, 37, 38). The protocol is based on first-strand cDNA synthesis and controlled multiplex amplification of cDNAs, which is followed by individual real-time PCR (TaqMan) quantification of amplified cDNAs in a 384-well format using a LightCycler 480.

Total *M. tuberculosis* RNA was isolated from the infected mouse lung tissue by homogenization in TRizol and bacillary disruption by bead beating (MP Biomedicals, Solon, OH). Total RNA was isolated using RNeasy columns (Qiagen, Valencia, CA). RNA was precipitated, cleaned with two consecutive off-column RQ1 DNase digestions (Promega, Madison, WI), and resuspended in 50 μl RNase-free water (Applied Biosystems/ Ambion, Austin, TX).

cDNA synthesis was performed with 50 ng total RNA, which was separated in reverse transcriptase (RT)+ and RT− reactions to control for DNA contamination. Exo-resistant random primer (0.5 μl), 1 μl 10 mM dNTPs, and nuclease-free water was added and incubated for 3 min at 70°C in a thermal cycler. Subsequently, 4 μl 5× Maxima RT buffer, 0.5 μl RiboLock RNase inhibitor, 0.5 μl Maxima RT enzyme (replaced by water for RT− control samples) (all Fermentas, Glen Burnie, MD), and nuclease-free water were added and incubated at 50°C for 1 h, 95°C for 2 min to inactivate, and then kept at 4°C.

The generated cDNAs were further amplified via controlled multiplex preamplification with a mix of 2179 *M. tuberculosis* gene-specific primers (23.8 μl primer mix: ∼50 nM per amplification reaction), 2 μl cDNA, 3 μl 10× Advantage 2 buffer (Clontech, Mountain View, CA), 0.6 μl 1× Advantage 2 polymerase mix (Clontech, CA), and 0.6 μl 10 mM dNTPs (Fermentas) to a volume of 30 μl (ftp://sam-dip.sam ford.edu/tbdb/rtpcr/taqman oligos fa). (15). Sequences and design of PCR primer/probe sets are available at: http://genes.stanford.edu/technology.php and http://www.tbdb.org/rtpcrData.shtml. A comparative control of 100 pg (2 are available at: http://genes.stanford.edu/technology.php and http://www.tbdb.org/rtpcrData.shtml) (15). Sequences and design of PCR primer/probe sets (Biosearch Technologies). Quantitative real-time PCR mix contained 0.07 μl preamplified cDNA, 2 μl TaqMan primer/probe mix, 5 μl 2× LightCycler 480 Probes Master Mix, and 2.95 μl Probes Master PCR-grade water (Roche) to a final volume of 10 μl. Reactions were heated at 95°C for 5 min, followed by 40 cycles at 95°C for 30 s and 60°C for 20 s. A cool down step of 40°C for 30 s was run for one cycle. Cycle threshold values were converted to relative gene copy numbers (RGCN) based on logarithmic transformation/linear regression equations devised from calibration curves. The data set is available at: http://www.tbdb.org/pubdata/tbdb/publications/ Raw-Data-Harvard-Mice.xls.

Correction for some biological heterogeneity between the different mice and mouse strains such as differences in bacterial load was not possible because these were inherent to the extensive differences in genetic TB susceptibility.

**IVE-TB gene selection procedure**

First, genes were selected that were expressed in one data set (>1000 RGCN) but not in the other data set (+/−), thus selecting *M. tuberculosis* genes that are differentially expressed owing to genetic host susceptibility and/or infection phenotype variations.

The second approach was to select genes that were highly expressed in both data sets from two different mouse strains in the chosen comparison (+/+), selecting for *M. tuberculosis* genes that are expressed independent of the genetic makeup of the host. For this selection, the RGCN data were ranked from the highest to the lowest value and overlapping genes were selected from the top 100 highest expressed genes of both data sets. This number of 100 genes was arbitrarily chosen to limit the number of candidates to be analyzed further.

The third and last approach included genes that were expressed in data set 2 (>1000 RGCN) but not in data set 1 (−/+), following the same rationale as approach 1. Hence, approaches 1 and 3 included differentially expressed genes. There was no number restriction limit (because fewer genes were identified compared with the second approach [+/+]) (Figs. 1, 2).

**Recombinant proteins**

Recombinant proteins were produced from the selected *M. tuberculosis* genes as described previously (39). Briefly, *M. tuberculosis* genes were amplified by PCR from genomic H37Rv DNA and cloned by Gateway technology (Invitrogen, Carlsbad, CA) in a bacterial expression vector containing a histidine tag at the N terminus. Vectors were overexpressed in *Escherichia coli* BL21 (DE3) and purified. Size and purity of recombinant proteins were analyzed by gel electrophoresis and Western blotting with an anti-His Ab (Invitrogen) and an anti-*E. coli* polyclonal Ab (gift of Statens Serum Institute, Copenhagen, Denmark). RV2380c, RV2435c, and RV2377c proteins were prepared as two or three recombinant protein fragments owing to their large sizes (C, middle [M], and N termini). Endotoxin contents were <50 IU/ml as tested using a Limulus amebocyte lysate assay (Cambrex, East Rutherford, NJ). All recombinant proteins were tested in lymphocyte stimulation assays to exclude Ag nonspecific T cell stimulation and cellular toxicity using PBMC of in vitro–purified protein derivative of *M. tuberculosis*—negative healthy Dutch donors (Sanquin Blood Bank, Leiden, The Netherlands) (12, 40–42). Purified protein derivative of *M. tuberculosis* was purchased from Statens Serum Institute.

**FIGURE 1.** Overview of IVE-TB gene selection procedure. (A) The *M. tuberculosis* RGCN profiles of each mouse model were independently compared with each other. (B) Three gene selection procedures were used to select genes for each comparison: genes that were expressed in data set 1 (>1000 RGCN) but not in data set 2 (+/−); genes highly expressed in both data sets (top 100 highest expressed genes in both models and select overlapping genes) (+/+); and genes not expressed in data set 1 but expressed in data set 2 (>1000 RGCN) (−/+).
was harvested and stored at 2°C for future use in IFN-γ ELISAs.

IFN-γ ELISAs

IFN-γ concentrations in supernatants were measured with a standard ELISA technique (U-CyTech, Utrecht, The Netherlands). ELISA samples were tested in duplicate and the assay was performed according to the manufacturer’s guidelines. Detection limit of the assay was set arbitrary at 20 pg/ml for whole blood assay (WBA) and 100 pg/ml for a lymphocyte stimulation test.

Flow cytometric analysis

PBMC (1–2 × 10^6/tube) were thawed and rested overnight and subsequently stimulated for 16 h with protein (10 μg/ml) in the presence of costimulatory Abs anti-CD28 and anti-CD49d (Sanquin and BD Biosciences, respectively). Brefeldin A (3 μg/ml; Sigma-Aldrich) was added after the first 4–6 h. Cells were stained with Live/Dead fixable violet dead cell stain (VIVID; Invitrogen) to discriminate between live and dead cells according to manufacturer’s instructions. Cells were stained for 1 h at 4°C with the following surface markers: anti-CD3 PE-Cy5 (BD Biosciences), anti-CD4 Texas Red (Caltag), and anti-CD8 V500 (BD Biosciences). Additionally, anti-CD14 Pacific Blue and anti-CD19 Pacific Blue (both Invitrogen) were included to select for CD14+ and CD19+ live cells. Intracellular staining was performed with anti–IFN-γ Alexa 700 (BD Pharmingen), anti–TNF-α PE-Cy7 (BD Biosciences), anti–IL-2 PE (BD Pharmingen), and CD154 alkaline phosphocyanin-Alexa 780 (eBioscience) using the ADG Fix&Perm kit (An Der Grub Bio Research, Vienna, Austria). Data were acquired on a BD LSRII (BD Biosciences) and analyzed using FlowJo version 7.6.5 (Tree Star, Ashland, OR). Single CD14+/CD19−/CD3+ live cells were gated to analyze CD4+ and CD8+ cytokine responses. Final Ag-specific CD4+ and CD8+ T cell subset populations all contained at least 200 events. For comparative purposes, medium background values were subtracted for each response in each donor.

Statistical analysis

Statistical analysis was performed using GraphPad Prism (version 5.1). A Mann–Whitney U test was used to analyze 1) the difference between cumulative IFN-γ responses for the ESAT6/CFP10 hybrid (E/C)+, E/C−, and HC individuals and 2) the difference between PHA-induced IFN-γ responses measured in E/C+ and E/C− donors. A p value ≤0.05 was considered significant.

Results

Identification of IVE-TB genes in the lungs of genetically resistant and/or susceptible mice

To start identifying novel candidate *M. tuberculosis* Ags in an unbiased and *M. tuberculosis* genome-wide fashion, we analyzed the gene expression patterns of 2170 *M. tuberculosis* genes, most of which represent the first gene of each predicted *M. tuberculosis* operon, in the lungs of four different *M. tuberculosis*–infected mouse strains (B6, C3H, C3H.B6-sst1, and B6.C3H-sst1) that show a spectrum of TB susceptibility (Table I). The RGCN were determined using quantitative PCR (15). This allows for absolute quantification of the level of transcripts per sample, because the data are normalized against a standard reference gene number copy (as described in Materials and Methods).

In this IVE-TB gene selection screen we used the following criteria to select candidate genes for further analysis. First, we used the most strongly upregulated *M. tuberculosis* genes in all four analyzed mouse models. This group of genes includes genes that are expressed independently of the host susceptibility background. Second, we used *M. tuberculosis* genes specifically upregulated in either B6, C3H, B6.C3H-sst1, or C3H.B6-sst1. The expression of these genes is influenced by the host genetic background and may therefore include genes whose expression is associated to particular TB disease characteristics such as granuloma formation and

![FIGURE 2. Selection of IVE-TB genes. Numbers of *M. tuberculosis* genes obtained after each comparison of two different mouse strains for every approach described in Fig. 1 are shown in (A)–(G).](http://www.jimmunol.org/)
necrosis. The obtained RGCN of all four mouse models were independently compared with each other as visualized in Fig. 1A. All four mouse models were infected with a low-dose aerosol M. tuberculosis inoculum. Additionally, a subset of B6.C3H-sstl was infected using a previously described TB relapse model to study features of gene expression during M. tuberculosis reactivation (23).

Every individual M. tuberculosis gene’s expression data (e.g., B6 versus C3H, B6 versus B6.C3H-sstl) was subjected to three different selection approaches as indicated in Fig. 1B and described in detail in Materials and Methods. The gene selection results for each comparison are visualized in Fig. 2 and Supplemental Table I. The results shown in Fig. 2 were then used to select IVE-TB genes whose expression was associated with particular disease characteristics as indicated in Fig. 3. These included 1) M. tuberculosis genes highly expressed independently of host genetic background (expressed in all four mouse models; these M. tuberculosis genes included esxA encoding ESAT6 and other esx genes); 2) M. tuberculosis genes expressed in association with necrosis (expressed in both C3H and B6.C3H-sstl, but not in B6 or C3H.B6-sstl); 3) M. tuberculosis genes expressed in association with severe necrotic infection or susceptibility (expressed in C3H, but not in B6, C3H.B6-sstl, or B6.C3H-sstl); 4) M. tuberculosis genes expressed in association with dense granuloma development (only expressed in B6.C3H-sstl but not in C3H, B6, or C3H.B6-sstl); 5) M. tuberculosis genes expressed in association with diffuse granuloma development (expressed in C3H but not in B6, C3H-sstl, B6, or C3H.B6-sstl); 6) M. tuberculosis genes expressed in association with resistance (expressed in every more resistant mouse strain per comparison); 7) M. tuberculosis genes expressed in association with low inflammation (expressed in B6, but not in C3H.B6-sstl); 8) inflammation (expressed in C3H.B6-sstl, but not in B6); and 9) relapse (expressed in i.v. M. tuberculosis–infected, INH-treated B6.C3H-sstl, but not low-dose aerosol–infected B6.C3H-sstl). An overview of the resulting IVE-TB genes is presented in Supplemental Table II.

Immunogenicity of newly identified IVE-TB Ags

Further down selection of IVE-TB genes. The goal of the above selection of IVE-TB genes was to identify potentially interesting new vaccine candidate Ags. Thus, we next determined their immunogenicity. To this end, a number of M. tuberculosis genes were further selected that were either present in more than one group or were among the top number of genes in the IVE-TB selections performed (Table II). Some M. tuberculosis genes identified using our unbiased genome-wide approach were also identified in previous studies as environmental stress induced proteins (7, 14, 17).

A subsequent literature search revealed that almost all further selected IVE-TB proteins (14 of the 16) have been identified previously in M. tuberculosis proteomic studies, confirming the protein expression of the M. tuberculosis genes identified in our study (44–57) (Table III). Indeed, we observed that M. tuberculosis–

Table I. Mouse strains, genetic background, and TB infection phenotypes

<table>
<thead>
<tr>
<th>Mouse Strain</th>
<th>Genetic Background</th>
<th>Sstl Allele</th>
<th>Inflammation</th>
<th>Lung Necrosis</th>
<th>Clinical TB Correlate</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3H</td>
<td>C3H</td>
<td>Susceptible</td>
<td>+++</td>
<td>+ (Early)</td>
<td>Caseous pneumonia</td>
</tr>
<tr>
<td>C3H.B6-sstl</td>
<td>C3H</td>
<td>Resistant</td>
<td>+++</td>
<td>–</td>
<td>Progressive interstitial granuloma without necrosis</td>
</tr>
<tr>
<td>B6.C3H-sstl</td>
<td>B6</td>
<td>Susceptible</td>
<td>++</td>
<td>+ (Late)</td>
<td>Causative granuloma</td>
</tr>
<tr>
<td>B6.C3H-sstl</td>
<td>B6</td>
<td>Resistant</td>
<td>+</td>
<td>–</td>
<td>Chronic persistent interstitial granuloma without necrosis</td>
</tr>
</tbody>
</table>

++, +++, and +++ indicate intensity of inflammation.

*Pichugin et al. (23) and Pan et al. (24).

FIGURE 3. Selection of IVE-TB genes associated with particular TB disease characteristics. Flowchart of analysis to identify IVE-TB genes related to TB disease phenotypes is shown. Letters in graphs refer to the specific selections indicated in Fig. 2. M. tuberculosis genes highly expressed independent of host genetic background are presented in Fig. 2A (b), 2B (e), 2C (h), 2D (k), 2E (n), and 2F (q). M. tuberculosis genes highly expressed in association with necrosis are presented in Fig. 2A (c), 2B (f), 2D (j), and/or 2F (p). M. tuberculosis genes highly expressed in association with diffuse granuloma are presented in Fig. 2C (g), 2B (f), and/or 2F (p). M. tuberculosis genes highly expressed in association with severe necrotic infection or susceptibility are presented in Fig. 2A (c), 2D (f), 2F (p), and/or 2G (q). M. tuberculosis genes highly expressed in association with inflammation, resistance, or relapse are presented in Fig. 2A (a), 2B (d), 2D (l), 2F (m), and/or 2G (s). M. tuberculosis genes highly expressed in association with dense granuloma are presented in Fig. 2C (i), but also Fig. 2A (c) and 2D (j). M. tuberculosis genes highly expressed in association with inflam-


dation are presented in one or more of the following selections: Fig. 2A (a), 2B (d), 2D (l), 2E (m) (although not related to the susceptible sst1 locus), 2F (r) and/or 2G (s). M. tuberculosis genes highly expressed in association with low inflammation are presented in Fig. 2E (m). M. tuberculosis genes highly expressed in association with inflammation are presented in Fig. 2E (o). M. tuberculosis genes highly expressed in association with relapse-associated genes presented in Fig. 2G (u).
infected C3H mice recognized most of the selected IVE-TB Ags as measured by Ag-specific IFN-γ production by splenocytes (Supplemental Fig. 1). Furthermore, we analyzed the conservation of these IVE-TB proteins using protein BLAST searches on different *M. tuberculosis* strains as well as other mycobacterial species. This showed that the IVE-TB protein sequences are strongly conserved among all tested *M. tuberculosis*, *M. bovis*, and *Mycobacterium africanum* strains. Strong conservation was observed also for other mycobacterial strains (Table IV).

### Table II. Predicted function and classification of selected IVE-TB genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Name</th>
<th>IVE-TB Selection</th>
<th>Selections</th>
<th>Function</th>
<th>Category</th>
<th>Classification</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rv1284</td>
<td>Rv1284</td>
<td>High expression</td>
<td>b + c + h + k + n + q + t</td>
<td>Conserved hypothetical protein (carbonic anhydrase)</td>
<td>7</td>
<td>EHR/starvation</td>
<td>14, 17, 43</td>
</tr>
<tr>
<td>Rv2380c</td>
<td>mbtE</td>
<td>High expression</td>
<td>b + c + h + k + n + q + t</td>
<td>Peptide synthetase mbtE</td>
<td>1</td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>Rv3515c</td>
<td>fadD19</td>
<td>High expressed stt1</td>
<td>h + q + t</td>
<td>Probable fatty acid-CoA ligase fadD19 (involved in lipid degradation)</td>
<td>1</td>
<td>EHR</td>
<td>17, 21</td>
</tr>
<tr>
<td>Rv0079</td>
<td>Rv0079</td>
<td>Necrosis and sni and relapse</td>
<td>c + f + j + p</td>
<td>Hypothetical protein</td>
<td>10</td>
<td>DosR</td>
<td>7, 9, 12, 20, 21</td>
</tr>
<tr>
<td>Rv2324</td>
<td>Rv2324</td>
<td>Necrosis and sni and relapse</td>
<td>c + j + u</td>
<td>Probable transcriptional regulator, asnc family</td>
<td>9</td>
<td>EHR</td>
<td>17</td>
</tr>
<tr>
<td>Rv2737c</td>
<td>recA</td>
<td>Necrosis and sni</td>
<td>c + f + j + p</td>
<td>Recombination protein recombinase A (recA; <em>M. tuberculosis</em> recA intein)</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2838c</td>
<td>rfbA</td>
<td>sni, diffuse granuloma, and relapse</td>
<td>c + i + j + u</td>
<td>Probable ribosome-binding factor A (P15B protein)</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv3420c</td>
<td>rimI</td>
<td>sni, diffuse granuloma, and relapse</td>
<td>c + i + j + u</td>
<td>Ribosomal-protein-alanine acetyltransferase rimI</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2034</td>
<td>Rv2034</td>
<td>Resistance, diffuse granuloma, low inflammation, and relapse</td>
<td>o</td>
<td>Transcriptional regulator (possible antitoxin; TA operon with Rv1955)</td>
<td>9</td>
<td>EHR/starvation</td>
<td>14, 17, 20</td>
</tr>
<tr>
<td>Rv1956</td>
<td>Rv1956</td>
<td>Resistance, diffuse granuloma, low inflammation, and relapse</td>
<td>d + i + j + m + u</td>
<td></td>
<td>0</td>
<td>EHR/starvation</td>
<td>14, 17, 21</td>
</tr>
<tr>
<td>Rv2225</td>
<td>panB</td>
<td>Dense granuloma</td>
<td>f + g + p</td>
<td>Probable 3-methyl-2-oxobutanoate hydroxymethyltransferase (panB)</td>
<td>7</td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>Rv2465c</td>
<td>rpiB</td>
<td>Dense granuloma</td>
<td>f + g + p</td>
<td>Isomerase (ribose 5-phosphate isomerase)</td>
<td>7</td>
<td>EHR</td>
<td>17, 20</td>
</tr>
<tr>
<td>Rv2982c</td>
<td>gpsA</td>
<td>Resistance, inflammation, and relapse</td>
<td>l + o + r + u</td>
<td>Probable glycerol-3-phosphate dehydrogenase (gpdA2)</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv3353c</td>
<td>Rv3353c</td>
<td>Relapse</td>
<td>u</td>
<td>Conserved hypothetical protein</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv1363c</td>
<td>Rv1363c</td>
<td>Resistance</td>
<td>s</td>
<td>Possible membrane protein</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2435c</td>
<td>Rv2435c</td>
<td>Resistance and inflammation</td>
<td>l + o</td>
<td>Probable cyclase (adenylate or guanylate cyclase)</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*In reference to Figs. 2 and 3.


*TubercuList functional classification codes are available at: http://genolist.pasteur.fr/tuberculist.

sni, Severe necrotic infection.

### Table III. Identification of IVE-TB proteins in *M. tuberculosis*

<table>
<thead>
<tr>
<th>Gene</th>
<th>Protein Name</th>
<th>Protein Identification</th>
<th>Protein Location</th>
<th>Membrane/Lipid Associated</th>
<th>Cytosol</th>
<th>CF</th>
<th>WCL</th>
<th>Essential In Vitro</th>
<th>Essential In Vivo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rv0079</td>
<td>Rv0079</td>
<td>49, 51, 52</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No (53)</td>
<td>Yes (57)</td>
</tr>
<tr>
<td>Rv1284</td>
<td>Rv1284</td>
<td>45–48, 51</td>
<td>X</td>
<td>X</td>
<td></td>
<td>Yes (53)</td>
<td>Yes (54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv1363c</td>
<td>Rv1363c</td>
<td>45, 46</td>
<td>X</td>
<td></td>
<td>No (53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv1956</td>
<td>Rv1956</td>
<td>46</td>
<td>X</td>
<td></td>
<td>No (53, 55)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2034</td>
<td>Rv2034</td>
<td>45, 46</td>
<td>X</td>
<td></td>
<td>No (53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2225</td>
<td>panB</td>
<td>46, 52</td>
<td>X</td>
<td>X</td>
<td>Yes (53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2324</td>
<td>Rv2324</td>
<td>45, 46</td>
<td>X</td>
<td>X</td>
<td>No (53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2380c</td>
<td>mbtE</td>
<td>51, 52</td>
<td>X</td>
<td>X</td>
<td>Probable yes (55)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2435c</td>
<td>Rv2435c</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No (53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2465c</td>
<td>rpiB</td>
<td>44–48</td>
<td>X</td>
<td>X</td>
<td></td>
<td>No (53)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2737c</td>
<td>recA</td>
<td>52</td>
<td>X</td>
<td></td>
<td>Slow grow mutant (53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2838c</td>
<td>rfbA</td>
<td>45, 46</td>
<td>X</td>
<td></td>
<td>NT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv2982c</td>
<td>gpsA</td>
<td>45, 46, 50, 51</td>
<td>X</td>
<td></td>
<td>No (53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv3353c</td>
<td>Rv3353c</td>
<td></td>
<td></td>
<td></td>
<td>No (53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv3420c</td>
<td>rimI</td>
<td>52</td>
<td>X</td>
<td></td>
<td>No (53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rv3515c</td>
<td>fadD19</td>
<td>46, 49, 50</td>
<td>X</td>
<td>X</td>
<td>No (53, 55); yes (56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note that the locations of the proteins indicated may not be exclusive given limitations and difficulties in annotating exact protein localization (46). CF, Culture filtrate; NT, not tested; WCL, whole cell lysate.
Table IV. IVE-TB protein sequence identity among mycobacterial strains

<table>
<thead>
<tr>
<th>Mycobacterial Species</th>
<th>Strain</th>
<th>Rv0079</th>
<th>Rv1284</th>
<th>Rv1363c</th>
<th>Rv1956</th>
<th>Rv2034</th>
<th>Rv2225</th>
<th>Rv2324</th>
<th>Rv2380c</th>
<th>Rv2435c</th>
<th>Rv2465c</th>
<th>Rv2737c</th>
<th>Rv2838c</th>
<th>Rv2982c</th>
<th>Rv3353c</th>
<th>Rv3420c</th>
<th>Rv3515c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erdmann</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>tuberculosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>Haarlem</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>ND</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>tuberculosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F11</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>tuberculosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDC1551</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>tuberculosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>africanum</td>
<td>GM041182</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>bovis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paratuberculosis</td>
<td>K-10</td>
<td>80</td>
<td>74</td>
<td>ND</td>
<td>41</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>ulcerans</td>
<td>104</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>87</td>
<td>89</td>
<td>88</td>
<td>88</td>
<td>89</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
<td>94</td>
</tr>
<tr>
<td>M. leprae</td>
<td>M-155</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>M. marinum</td>
<td>TN</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>


calosis–specific E/C protein (32%, 61–9980 pg/ml). ESAT6 and CFP10 M. tuberculosis–specific proteins are frequently used as immunodiagnostic Ags to identify M. tuberculosis exposure (12). One third of our population also responded to these diagnostic Ags (ESAT6, CFP10 complemented with TB7.7 [p4] peptides) using the commercial QFT-GIT. The responses in the QFT-GIT correlated very well with responses to our E/C protein measured in WBA (84%). Division of the group into WBA E/C and WBA E/C− donors showed that 74% of the E/C+ donors responded to M. tuberculosis lysate (87–3112 pg/ml) and 40% to Ag85B Ag (97–735 pg/ml) (Fig. 4A). Importantly, high levels of recognition of the newly identified M. tuberculosis proteins were seen within this population, with responses ranging from 10 to 60% of the donors.

To identify the Ags that are well recognized, we first examined which ones were recognized by ≥25% of the E/C+ donors. From all 20 proteins and protein fragments (derived from the 16 selected Ags) tested, 13 were recognized by ≥25% (n ≥ 3 of 10 and ≥11 of 43) of the donors. Six of 13 Ags induced only intermediate levels of IFN-γ (60–600 pg/ml) after stimulation, whereas high levels of IFN-γ were produced after stimulation with Rv1363c (102–1053 pg/ml), Rv1956 (100–1861 pg/ml), Rv2034 (77–1256 pg/ml), Rv2324 (63–2130 pg/ml), Rv2380c-C (80–1159 pg/ml), Rv3353c (70–899 pg/ml), and Rv3420c (88–660 pg/ml), representing the seven best recognized proteins.

Of the TST+ E/C+ donors still 47% responded to M. tuberculosis lysate and 20% to Ag85B (Fig. 4B). Additionally, HC were analyzed for possible IVE-TB responses (Fig. 4C). Most importantly, limited responses to no responses against the IVE-TB proteins were detected in the TST+ E/C+ and HC groups, demonstrating clear specificity of recognition, presumably linked to M. tuberculosis exposure.

To further verify that the IVE-TB Ags are specifically recognized by E/C+ donors, the cumulative IFN-γ response to the 20 tested Ags per individual was calculated, as described before (9). A highly significant difference between the E/C+ and E/C− population was observed (p < 0.0001). As expected, a significant difference was also observed between E/C+ and HC donors (p = 0.049) (Fig. 5), confirming the association between Ag recognition and E/C test positivity. Interestingly, of the seven best recognized proteins, Rv3420c was the most discriminatory between the E/C+ (≥25%) and E/C− group (≤1%) (p < 0.0001) and was not recognized by HC donors (p = 0.016), suggesting a possible role as M. tuberculosis–specific biomarker Ag in addition to ESAT6, CFP10, and TB7.7.

Recognition of IVE-TB proteins by PBMC from TB patients. We next investigated whether TB patients could also recognize these Ags. PBMC from WBA E/C+ TST+ donors (Fig. 6A) and TB patients (Fig. 6B) were therefore stimulated with the 20 proteins and protein fragments and IFN-γ production was measured. High IFN-γ responses (123–3391 pg/ml) to the IVE-TB Ags were observed in the PBMC cultures from the WBA E/C+ TST+ population, confirming the results obtained in the WBA assay above. Nine of the tested Ags were recognized by ≥50% of the donors and eight Ags by 38% of the donors. Only one protein fragment was not recognized in this assay (Rv2380c-M). In contrast, seven of the tested IVE-TB Ags did not induce detectable IFN-γ production in PBMC from TB patients, whereas most of the IVE-TB Ags induced only low levels of IFN-γ compared with the WBA E/C+ TST+ individuals (107–1825 pg/ml). Only two Ags induced high levels of IFN-γ in the TB patients. Additionally, only one Ag was recognized by 50% of the TB patients, whereas the remainder of the Ags were recognized by relatively fewer TB patients (14–43%). Thus, IVE-TB Ags seem to be less immunogenic in TB patients than in TST+ individuals.
T cell responses toward IVE-TB Ags in long-term latent M. tuberculosis--infected individuals. Because the IVE-TB Ags were strongly recognized by TST+ individuals, we subsequently analyzed the immune responses toward the seven best recognized Ags in more detail using PBMC from donors that had been exposed to M. tuberculosis decades ago, but had never developed TB despite the lack of any preventive treatment, designated ltLTBI (41, 42). Of additional importance, the availability of several vials of PBMC also allowed more detailed cell subset analysis.

Interestingly, high frequencies of TNF-α- and IL-2–producing CD4+ T cells were observed after stimulation with the IVE-TB Ags, whereas only intermediate frequencies of IFN-γ–producing CD4+ T cells were detected (Fig. 7A). In contrast, high frequencies of IFN-γ–producing as well as TNF-α+ CD8+ T cells were present in these donors, whereas fewer IL-2+ T cells were detected compared with CD4+ T cells. Besides IFN-γ, TNF-α, and IL-2, also the Ag-induced CD4+ T cell activation marker CD154 (58) was expressed.

More detailed analysis of the multifunctional Th1 responses among CD4+ and CD8+ T cells showed that CD154+CD4+ T cells were mostly TNF-α+/IL-2– and TNF-α– (Fig. 7B). Furthermore, intermediate frequencies of IFN-γ+/TNF-α+/IL-2+ CD154+CD4+ T cells were detected. Finally, a CD154+ population was detected, producing none of the IFN-γ, TNF-α, and IL-2 cytokines. Intriguingly, the same pattern was observed for every IVE-TB Ag or E/C control Ag. Furthermore, interindividual variation of Ag recognition was observed. Remarkably, few TNF-α+/IL-2– CD8+ T cells were detected compared with TNF-α+/IL-2+ CD4+ T cells. IFN-γ+/TNF-α+CD8+ T cells were the most prominent population present, followed by TNF-α+ CD8+ T cells. Also, intermediate IFN-γ+/TNF-α+/IL-2+ and IFN-γ−+CD8+ T cells were observed. Again, as mentioned for CD4+ T cells, the same patterns were observed for every Ag within the CD8+ T cell population as well as interindividual variation of Ag recognition.

The integrated median fluorescence intensity (iMFI) was calculated to determine the quantitative contribution of cytokines produced by the different multiple and single cytokine producing CD154+/CD4+ and CD8+ T cells (Fig. 8). IFN-γ+/TNF-α+/IL-2+ CD154+/CD4+ T cells had the highest iMFI, which gradually declined for double producing and single IFN-γ+CD154+/CD4+ T cells. IFN-γ−/TNF-α−CD8+ T cells contributed the most to IFN-γ production, directly followed by the IFN-γ−/TNF-α+/IL-2+ CD8+ T cells. IFN-γ+/TNF-α+CD8+ T cells were also the main contributors for TNF-α, whereas IFN-γ−/TNF-α−/IL-2− CD8+ T cells showed a higher IL-2 iMFI. TNF-α and IL-2 iMFI were also the highest for TNF-α−/IL-2− CD154+/CD4+ T cells, followed by the IFN-γ−/TNF-α−/IL-2− CD154+/CD4+ T cells. Thus the TNF-α−/IL-2− CD4 and IFN-γ−/TNF-α− CD8 T cells contribute strongly to the production of Th1 cytokines, followed by the triple-positive T cells. Single cytokine–producing cells only showed a relatively minor contribution.

In conclusion, seven of the identified IVE-TB Ags are strongly immunogenic, triggering specific and high cellular immune responses in E/C+ TST+ individuals and long-term ltLTBI individuals, but not in E/C− TST+ individuals, healthy mycobacterial naive individuals, and TB patients. The strong IVE-TB responses that were measured in the ltLTBI group were identified as IFN-γ−/ TNF-α− CD8+ T cells and TNF-α−/IL-2− CD4+ T cells, which were the most prominent contributors to the produced cytokines, followed by triple-positive T cells.

**Discussion**

Using quantitative genome-wide M. tuberculosis transcriptional profiling, we have identified a series of M. tuberculosis genes that

---

**FIGURE 4.** IFN-γ responses to IVE-TB Ags in E/C+ (A), E/C− (B), and HC donors (C). A total of 43 E/C+ donors (A), 90 E/C− donors (B), and 11 HC donors (C) were analyzed for their IFN-γ WBA responses to Ags and controls; the Ags were *M. tuberculosis* RV1284, RV1956, RV2034, RV2324, RV3353c, and RV3420c. Ten E/C+ donors (A), 36 E/C− donors (B), and 9 HC donors (C) were also analyzed for responses to *M. tuberculosis* Ags RV0079, RV1363c, RV2225, RV2380c, RV2435c, RV2465c, RV2737c, RV2828c, RV2982c, and RV3515c IFN-γ. The proportion of responders for each Ag is indicated at the top of the graph. For comparative purposes, medium background values were subtracted for each response in each donor. Horizontal bars represent the median IFN-γ responses. The dotted line indicates the cut-off value for positivity, arbitrarily calculated as 3 × medium value.
are expressed during in vivo *M. tuberculosis* infection in the lungs of resistant and susceptible mice, which we term IVE-TB. Most of the genes identified have previously been found to be expressed in the *M. tuberculosis* proteome, and thus encode bona fide *M. tuberculosis* proteins. This is further supported by their immunogenicity profiles, as many of these proteins triggered IFN-γ production in human WBA and lymphocyte stimulating assays in *M. tuberculosis* ESAT6/CFP10-responsive patients, but not in ESAT6/CFP10−TST+ individuals, HC donors, or TB patients. This is particularly relevant in the case of Rv2435c and Rv3353c, as their protein products have not been identified yet; however, IFN-γ responses were demonstrated in E/C + TST+ individuals, indirectly showing that these *M. tuberculosis* proteins are presented to the human immune system during mycobacterial infection. Many of the IVE-TB genes we identified have been described previously in relationship to the adaptative response of *M. tu-

**FIGURE 5.** Cumulative IFN-γ responses induced by IVE-TB Ags, calculated per individual in the E/C+, E/C−, and HC groups. Cumulative IFN-γ responses to all 20 IVE-TB protein and protein fragments in E/ C+ (n=43), E/C− (n=90), and HC donors (n=11). Squares indicate cumulative IFN-γ response of all 20 IVE-TB Ags, and circles indicate cumulative IFN-γ response of Rv1284, Rv1956, Rv2034, Rv2324, Rv3353c, and Rv3420c Ags. Horizontal bars represent the median cumulative IFN-γ responses.

**FIGURE 6.** PBMC IFN-γ responses toward IVE-TB Ags in TB patients and WBA E/C+ TST+ donors. PBMC of WBA E/C+ TST+ donors (n=8) (A) and TB patients (n=7) (B) were stimulated with IVE-TB Ags and control conditions for 6 d. Levels of IFN-γ were measured and medium background values were subtracted for each response in each donor for comparative purposes. The proportion of responders for each Ag is indicated at the top of the graph. Horizontal bars represent the median IFN-γ responses. The dotted line indicates the cut-off value for positivity, arbitrarily set at 100 pg/ml.
berculosis to environmental stress conditions, especially those that *M. tuberculosis* likely encounters during host infection. We identified one *M. tuberculosis* DosR regulon-encoded gene (*Rv0079*) (7) as well as six genes that are part of the EHR regulon (*Rv1284*, *Rv1956*, *Rv2034*, *Rv2324*, *Rv2465*, and *Rv3515*) (17). Three of these have also been described as starvation/nutritional stress genes (*Rv1284*, *Rv1956*, and *Rv2034*) (14). This function of the IVE-TB genes in responding to host-induced stress conditions during in vivo pulmonary infection enhances the biological plausibility of our findings and lends validity to our approach.

Of further interest, nine of the *M. tuberculosis* genes identified in this study have not been described previously in relationship to *M. tuberculosis* host infection, although some of their functions have been linked to possible adaptation to in vitro host-induced stress conditions (Supplemental Table III). Several of these genes have a role in metal transport, metalloregulatory transcriptional regulation, or represent metalloenzymes. Furthermore, genes were identified that play a role in lipid metabolism. This is in agreement with the documented shift toward using fatty acids as an alternative carbon source instead of carbohydrates under nutrient-limiting conditions. Altogether, many of the IVE-TB genes we have identified appear to be related to the adaptation of *M. tuberculosis* to environmental stress conditions encountered in the host.

Of additional importance, the identification of these genes in our in vivo model supports previous findings mostly obtained in in vitro models by showing that they are induced during pulmonary *M. tuberculosis* infection in vivo. On a cautionary note, however, our data do not allow us to discriminate whether the observed differential *M. tuberculosis* gene expression patterns are cause or
consequence of variations in host susceptibility (background and/or \textit{sst} locus).

As already mentioned, \textit{M. tuberculosis} gene expression profiling has been performed in the past, mostly focusing on in vitro–cultured bacteria grown under a variety of different conditions. Subsequent work assessed \textit{M. tuberculosis} gene expression profiles following infection of host cells (8, 19, 43), and some recent studies have analyzed \textit{M. tuberculosis} gene expression patterns also in vivo (20, 59). Ward et al. (59) showed that there was little overlap in the \textit{M. tuberculosis} genes reported to be expressed in different studies reporting on \textit{M. tuberculosis} intracellular infection, likely as a result of methodological differences. Nonetheless, the two studies Ward et al. described (8, 60) indicate that similar functional categories of \textit{M. tuberculosis} genes are expressed during intracellular infection, likely as a result of methodological differences. Nonetheless, the two studies Ward et al. described (8, 60) indicate that similar functional categories of \textit{M. tuberculosis} genes are expressed during intracellular infection. In line with this notion, when comparing our data to previous reports there are few overlapping individual \textit{M. tuberculosis} genes, but we nevertheless do identify genes with previous described functional categories. These differences are probably due to differences in selection criteria, in experimental settings such as infection route, and the specific mouse models we have used, which have not been analyzed previously.

Despite these differences, several of our selected IVE-TB genes do overlap with \textit{M. tuberculosis} genes identified in other studies as indicated in Table II. The \textit{M. tuberculosis} gene \textit{Rv}2225 whose expression was TB granuloma associated was also significantly expressed in the artificial granuloma model of Karakousis et al. (21). This strengthens their association with host granuloma formation. Our in vivo pulmonary TB granuloma-associated \textit{M. tuberculosis}–expressed genes did not overlap with the granuloma-associated genes or macrophage-associated genes described by Ramakrishnan and colleagues (61, 62) for \textit{M. marinum}, which might be due to differences between the mycobacterial species studied. Moreover, several of the IVE-TB genes we identified to be highly expressed have also been described previously, including \textit{Rv}0467 (\textit{icl}), which encodes an enzyme in the glyoxylate pathway, which is important for \textit{M. tuberculosis} persistence of \textit{M. tuberculosis} (63, 64), and \textit{Rv}0991c, which is part of the so-called in vivo–expressed genomic island (20).

The new \textit{M. tuberculosis} Ags we have identified in this study may represent interesting targets for vaccination, as they are expressed during \textit{M. tuberculosis} infection in the (genetically susceptible) lung, which we consider a critical parameter for appropriate Ag selection. Moreover, successful vaccine Ags should be conserved between multiple \textit{M. tuberculosis} strains. All protein sequences examined were conserved among the tested \textit{M. tuberculosis} strains. Additionally, for almost all IVE-TB genes multiple proteome studies have documented their expression as proteins in \textit{M. tuberculosis} (Table III). A subset of the analyzed IVE-TB proteins was shown to be strongly immunogenic as judged by Th1 responses in WBA, lymphocyte stimulation assays, and polychromatic flow cytometry. Indeed, the highest IFN-\gamma responses were identified within the E/C+ population of our TST+ cohort.
whereas no differences in mitogen-induced responses were seen. No responses were seen in *M. tuberculosis* nonresponder healthy individuals, suggesting that T cell recognition of IVE-TB Ags is indeed Ag specific and is correlated with *M. tuberculosis* exposure based on TST and QFT-GIT conversions. Interestingly, TB patients showed relatively low recognition of the IVE-TB Ags, suggesting that they did not develop strong Th1 immunity against these Ags. Importantly, IVE-TB Ag-specific responses could be detected in IL-17 production, immune regulation, or yet other functions, which need further study. Finally, we previously reported multiple differential T cell subsets may be unrelated to the nature of the specific protein Ag involved.

CD8+ T cells are activated upon recognition of epitopes presented via MHC class I molecules, indicating that the Ags are present and processed via the canonical cytotoxic pathway or via alternative (e.g., cross-priming) pathways (65). Both CD4+ and CD8+ T cells are important in *M. tuberculosis* control, and CD4+ and CD8+ T cell–deficient mice, for example, have increased susceptibility to *M. tuberculosis* (66). CD4+ T cells were recently shown to play an important (IFN-γ–independent) role in the indirect activation of IFN-γ–CD8+ T cells (67). In any case, our data obtained in the IL-17 individuals show that the *M. tuberculosis* Ag-specific CD4+ and CD8+ T cells recognizing IVE-TB Ags must be long lived.

The immunogenicity of some of the IVE-TB Ags has been analyzed previously. The immunogenicity of the DosR family protein was analyzed in TST* (endemic) individuals as well as (cured) TB patients (9, 12, 68). In these studies, Rv0079 protein was recognized by a minority of individuals, in agreement with our results in this study. The immunogenicity of EHR and starvation Ags Rv1284 and Rv1956 was previously analyzed in *M. tuberculosis*–exposed cattle (17, 69). Rv1284 was one of the five best recognized Ags, whereas Rv1956 was also highly recognized. In contrast to the responses observed in *M. bovis*–exposed cattle, Rv1284 was moderately recognized in our study, whereas Rv1956 was better recognized.

In conclusion, by combining *M. tuberculosis* genome-wide transcriptional profiling in the lungs of infected mice with strikingly differing host susceptibility backgrounds, we have identified *M. tuberculosis* genes that are specifically expressed in resistant or susceptible animals during pulmonary infection. These genes reveal a signature of the *M. tuberculosis* stress response in vivo depending on the genetic host background and host susceptibility. From these genes we selected 16 proteins, of which proved to be highly immunogenic in E/C* TST* donors and IL-17 donors and therefore represent interesting TB vaccine candidates and possibly TB biomarker Ags (70).

### Acknowledgments

We thank Louis Wilson for the production of the *M. tuberculosis* lysates, as well as all blood donors who participated in this study.

### Disclosures

T.H.M.O. is coinventor of an *M. tuberculosis* latency Ag patent, which is owned by Leiden University Medical Center. The other authors have no financial conflicts of interest.

### References

2014.
Dominant role of the sstl locus in pathogenesis of necrotizing lung granulomas
during chronic tuberculosis infection and reactivation in genetically resistant
Acad. Sci. USA 103: 10364–10368.
morphisms are not associated with pulmonary tuberculosis in a South African
31. Yan, B. S., A. V. Pichugin, O. Jobe, L. Helming, E. B. Ersuslanov, J. A. Gutiérrez-
Pabello, M. Rojas, Y. V. Sherezhkov, L. Kobzik, and I. Kramnik. 2007. Pro-
gression of pulmonary tuberculosis and efficiency of bacillus Calmette-Guérin
vaccination are genetically controlled via a common sstl-mediated mechanism of
32. DosR antigens and peptides in long-
morphisms are not associated with pulmonary tuberculosis in a South African
36. Yan, B. S., A. V. Pichugin, O. Jobe, L. Helming, E. B. Ersuslanov, J. A. Gutiérrez-
Pabello, M. Rojas, Y. V. Sherezhkov, L. Kobzik, and I. Kramnik. 2007. Pro-
gression of pulmonary tuberculosis and efficiency of bacillus Calmette-Guérin
vaccination are genetically controlled via a common sstl-mediated mechanism of
37. DosR antigens and peptides in long-
morphisms are not associated with pulmonary tuberculosis in a South African
41. Yan, B. S., A. V. Pichugin, O. Jobe, L. Helming, E. B. Ersuslanov, J. A. Gutiérrez-
Pabello, M. Rojas, Y. V. Sherezhkov, L. Kobzik, and I. Kramnik. 2007. Pro-
gression of pulmonary tuberculosis and efficiency of bacillus Calmette-Guérin
vaccination are genetically controlled via a common sstl-mediated mechanism of
42. DosR antigens and peptides in long-
morphisms are not associated with pulmonary tuberculosis in a South African
46. Yan, B. S., A. V. Pichugin, O. Jobe, L. Helming, E. B. Ersuslanov, J. A. Gutiérrez-
Pabello, M. Rojas, Y. V. Sherezhkov, L. Kobzik, and I. Kramnik. 2007. Pro-
gression of pulmonary tuberculosis and efficiency of bacillus Calmette-Guérin
vaccination are genetically controlled via a common sstl-mediated mechanism of
47. DosR antigens and peptides in long-
morphisms are not associated with pulmonary tuberculosis in a South African
51. Yan, B. S., A. V. Pichugin, O. Jobe, L. Helming, E. B. Ersuslanov, J. A. Gutiérrez-
Pabello, M. Rojas, Y. V. Sherezhkov, L. Kobzik, and I. Kramnik. 2007. Pro-
gression of pulmonary tuberculosis and efficiency of bacillus Calmette-Guérin
vaccination are genetically controlled via a common sstl-mediated mechanism of
52. DosR antigens and peptides in long-
morphisms are not associated with pulmonary tuberculosis in a South African