Published in final edited form as:
Tuberculosis (Edinb). 2007 July ; 87(4): 347–359.

Differential gene expression between *Mycobacterium bovis* and *Mycobacterium tuberculosis*

Germán Rehren[a], **Shaun Walters**[b,1], **Patricia Fontan**[b], **Issar Smith**[b], and **Ana M. Zárraga**[a,*]

a *Instituto de Bioquímica, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile.*

b *TB Center, The Public Health Research Institute of the New Jersey Medical School-UMDNJ, 225 Warren Street, Newark, NJ 07103, USA.*

Summary

The high sequence identity among the *Mycobacterium bovis* and *Mycobacterium tuberculosis* genomes contrasts with the physiological differences reported between these pathogens, suggesting that variations in gene expression may be involved. In this study, microarray hybridization was used to compare the total transcriptome of *M. bovis* and *M. tuberculosis*, during the exponential phase of growth. Differential expression was detected in 258 genes, representing a 6% of the total genome. Variable genes were grouped according to functional categories. The main variations were found in genes encoding proteins involved in intermediary metabolism and respiration, cell wall processes, and hypothetical proteins. It is noteworthy that, compared to *M. tuberculosis*, the expression of a higher number of transcriptional regulators were detected in *M. bovis*. Likewise, in *M. tuberculosis* we found a higher expression of the PE/PPE genes, some of which code for cell wall related proteins. Also, in both pathogens we detected the expression of a number of genes not annotated in the *M. tuberculosis* H37Rv or *M. bovis* 2122 genomes, but annotated in the *M. tuberculosis* CDC1551 genome.

Our results provide new evidence concerning differences in gene expression between both pathogens, and confirm previous hypotheses inferred from genome comparisons and proteome analysis. This study may shed some new light on our understanding of the mechanisms relating to differences in gene expression and pathogenicity in mycobacteria.

Keywords

Mycobacterium bovis; Mycobacterium tuberculosis; Transcriptome; Gene expression; Microarray

Introduction

Mycobacterium bovis and *Mycobacterium tuberculosis* are closely related pathogens, responsible for bovine and human tuberculosis, respectively. Bovine tuberculosis is enzootic in most developing countries, causing great economic losses. Lately, this zoonosis has raised in importance within the HIV infected population.[1,2] The *M. bovis* and *M. tuberculosis* genomes show 99.95% identity at the nucleotide level.[3] However, distinct phenotypes,

*Corresponding author. Instituto de Bioquímica, Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile. Phone: 56-63-221907; Fax: 56-63-221107. E-mail: anamariazarraga@uach.cl
[1]Present address: Department of Microbiology, Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10021, USA.

virulence and host tropism differentiate both pathogens.[4] Indeed, *M. bovi*s has shown increased virulence upon infection in mice, rabbits, calves and guinea pigs.[5,6] Moreover, it causes disease in bovines, humans and a broad range of mammals. In comparison, man is the main natural host of *M. tuberculosis.*[1,4,7,8] The high genome sequence homology and the lack of species-specific genes for *M. bovis*, suggest that distinctive mechanisms of gene expression might be involved in determining the differences among these bacilli.[3] Genome features such as deletions, SNPs and hypervariable regions are important sources for polymorphism, but their contribution needs to be correlated with gene expression studies in order to explain the differences observed. Transcriptional analysis using microarrays has become a useful tool to study whole-genome expression and to identify changes in gene expression in cells exposed to different environmental conditions.[9] Using this approach, diversity in gene expression has recently been reported among clinical isolates of *M. tuberculosis.*[10]

In this work, we used microarray hybridization to compare the total *in vitro* transcriptome of *M. bovis* and *M. tuberculosis* H37Rv. Our aim was to get a global view on the differences of gene expression among these pathogens as a mean to identify the genetic basis for their distinct phenotypes.

Materials and methods

Bacterial strains and culture conditions

The *M. tuberculosis* H37Rv (ATCC 27294) and *M. bovis* Ravenel (TMC 401) strains were used throughout and all cultures were grown under the same conditions. Bacteria were grown in 7H9 broth supplemented with 10% ADS (0.5% bovine serum albumin, fraction V, 0.2% glucose and 0.085% NaCl) and 0.05% Tween 80. For *M. tuberculosis*, 0.2% glycerol was also added to the media, while for *M. bovis* the media was supplemented with 0.4% sodium pyruvate. Liquid cultures were maintained at 37°C in plastic bottles in a roller apparatus and optical density (OD_{540}) was registered daily to measure growth.

Isolation of bacterial RNA

To minimize variations in gene expression analysis, all cultures were collected at the same growth phase. To prepare bacterial RNA, procedures previously published were followed.[11, 12] Bacterial cell pellets were recovered from 30 ml of exponential phase culture (OD_{540} ~0.25) by 3 min centrifugation and then quickly frozen on dry ice. Frozen pellets were then resuspended in 1 ml of TRI reagent (Molecular Research Center) and immediately transferred to a 2 ml screw cap microcentrifuge tube containing zirconia beads (0.1mm diameter). Samples were disrupted in a Mini-BeadBeater (BioSpec Products), and the RNA was extracted according to the TRI manufacturer's instructions. To remove residual DNA, samples were treated with Turbo™ DNAse (Ambion) and cleaned up using RNeasy columns (Qiagen). The integrity of all RNA samples was checked by non-denaturing agarose gel electrophoresis, with RNA concentration quantified by spectrophotometry. RNA samples were kept at −80°C until further use.

DNA microarray analysis

The *M. tuberculosis* microarray chips used in these assays were printed at the Center for Applied Genomics at the Public Health Research Institute. This array consists of 4,295 70-mer oligonucleotides representing 3,924 predicted open reading frames of the *M. tuberculosis* H37Rv strain, 371 non-redundant probes designed to the *M. tuberculosis* CDC-1551 strain, and 25 controls. Microarray analyses were performed according to previously described protocols [13] with modifications.[14] cDNA was synthesized and fluorescently labeled by a direct procedure. Two micrograms of total RNA extracted from either *M. tuberculosis* or *M. bovis* was reverse transcribed in a reaction mix with a final concentration of 0.17 ug/ul random

hexamers, 0.96x first strand buffer, 9.6 mM DTT, 0.44 mM dATP, dCTP and dGTP, 0.02 mM dTTP, 0.06mM Cyanine 3 or Cyanine 5 dUTP and 9.4 units Superscript II. The combination of reaction mix and total RNA was incubated for 10 minutes at 25°C followed by 90 minutes at 42°C. The labeled cDNA probes were then purified and concentrated using the MinElute Cleanup kit (Qiagen). The total purified cDNA probe was added to the arrays in a hybridization solution containing a final concentration of 0.5 ug/ul tRNA, 2.0x SSC, 0.25% formamide and 0.1% SDS. For each array, cDNA prepared from the *M. tuberculosis* RNA was mixed with cDNA from *M. bovis*. The slides were covered by a flat 22 × 22 mm coverslip and hybridized in sealed hybridization chambers for sixteen hours at 50°C in a water bath.

Microarray data analysis

Microarray slides were scanned using a GenePix 4000A scanner (Axon Instruments). Images were processed and the fluorescent intensity of each spot was quantified using the GenePix Pro 4.0 software. Four independent biological replicates were analyzed for each strain, and one swap-dye experiment was included. Median intensity values were corrected by background subtraction and negative corrected intensities were set to +1.[15] Further analysis was performed using GeneSpring 7.2 software (Silicon Genetics). Data was normalized using the locally weighted linear regression (Lowess) method, to remove the fluorescence-intensity dependent, dye-specific effect of low fluorescence intensity spots.[16] 20% of the data was used to calculate the Lowess fit at each point. Cy5/Cy3 intensity ratios were determined using normalized values and then log transformed. For each gene, the geometric mean was calculated from the intensity ratios of the four replicates and the resulting value was used to determine differences in mRNA abundance between both strains. Genes were classified as differentially expressed if they fulfilled both of the following criteria: a minimum 3-fold regulation difference and a p-value <0.05. Statistical significance of the chosen genes was verified by a t-test with the Benjamini Hochberg false discovery rate correction method[17] implemented in GeneSpring.

The data discussed in this publication have been deposited in NCBIs Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE6889.

Reverse transcription and real time PCR

RT and PCR primers were designed with the Primer3 web software[18] and purchased from IDT (Coralville, IA). Primer sequences used in this study are provided in supplementary Table S1. Reverse transcription was carried out at 60°C, using 100 ng of RNA, 0.3 μM antisense gene-specific primers (including *sigA* control primer) and the RNA Transcriptor First Strand cDNA Synthesis Kit (Roche), in a final volume of 20 μl, following the manufacturer's instructions. Control reactions, lacking reverse transcriptase, were performed for every RNA sample. Real time PCR reactions were accomplished using a LightCycler instrument (Roche), in a total volume of 10 μl, using 2 μl of diluted cDNA, 0.3 μM gene-specific primers and the LightCycler FastStart DNA Master SYBR Green I kit (Roche). After 10 min at 90°C, the PCR program consisted in 40-45 cycles at 95°C for 6 s, 62°C for 8 s and 72°C for 8 s. Fluorescence was measured at the end of the extension step. Reactions were run in triplicate for each gene and the specificity of the PCR products was verified by gel electrophoresis and melting curve analysis. For every PCR run, a standard curve was performed using serial dilutions of chromosomal DNA and the corresponding gene-specific primer set. These standard curves were used by the LightCycler software to calculate the initial amount of mRNA for each gene tested in the different samples. In order to compare gene expression between *M. bovis* and H37Rv strain, calculated values for each gene were normalized to the corresponding sample's *sigA* mRNA value,[19] and normalized values were used to calculate the *M. bovis*/*M. tuberculosis* gene ratio.[11] Ratios from 3 biological replicates were then averaged to give a final fold difference value.

Results and Discussion

Several lines of evidence have supported the fact that although slight differences are found in the genome sequence of *M. tuberculosis* and *M. bovis*, the physiology and host range spectrum is different. Variations in transcript abundance, even minor changes, may cause distinctive adaptive responses to changes in environmental conditions. To identify differences in the pattern of gene expression among *M. tuberculosis* and *M. bovis*, we used here a chip with all annotated *M. tuberculosis* genes. The comparative transcriptome profile was determined for both strains, grown to mid-exponential phase.

The results show that out of 4,295 genes that were represented on the chip, 6 % were differentially expressed. Of these genes, 95 exhibited higher expression in *M. bovis* compared to *M. tuberculosis*, while 163 genes were more highly expressed in *M. tuberculosis* compared to *M. bovis*. The list of differentially regulated genes in *M. bovis* and *M. tuberculosis* are shown in Table 1 and Table 2, respectively. The difference in transcript abundance for some of the selected genes was further confirmed by RT-qPCR (Table 3). Variations in transcript abundance were higher in RT-qPCR compared to microarray results, but followed the same pattern. Similar results have been observed in previous reports.[20,21]

Differentially regulated genes were grouped according to functional categories as described in TubercuList and BoviList (Table 4). The highest difference in gene expression was found in genes related to general metabolism, insertion sequences, hypothetical proteins and cell wall proteins. A significant proportion of genes coding for proteins associated with metabolic processes were identified for both pathogens. Using the Kyoto Encyclopedia of Genes and Genomes (KEGG) [22] to search for metabolic pathways assigned to these genes, we found that some of the differences observed within this category corresponded to genes encoding proteins related to amino acid, steroid and sugar metabolism. With the aim of support successful growth, the culture media was supplemented with different carbon sources. Keeping in consideration that this may have an effect on the expression of some genes, we particularly looked for variations on genes that are likely to be affected, such as the key glycolytic enzymes *glpK*, *pykA* and *pdhA*. Our results showed no significant differences using the 3-fold cut-off. On the other hand, Tween 80 was also added to both culture media. This oleic acid ester can be used as a carbon source *in vitro* by mycobacteria, and it has also been shown that fatty acids are the primary carbon source *in vivo*. [23] Therefore, taking together this data, we consider that although different carbon sources may influence gene expression to some extent, the results presented in this paper are representative of the *in vitro* differences on gene expression between *M. bovis* and *M. tuberculosis*.

Interestingly, 20 % of the differentially expressed genes code for hypothetical proteins in *M. bovis* compared to 13.5% in *M. tuberculosis*. These results point to the need of assigning a function to these proteins, as they may be involved in determining the physiological differences described for both bacilli.

Within the category of cell wall proteins, 18 genes of the PE/PPE family were found to be higher expressed in *M. tuberculosis* compared to three genes identified in *M. bovis*. These genes code for surface exposed proteins involved in host-pathogen interactions.[24-26] In addition, six ESAT6-like genes showed higher expression in *M. tuberculosis* whereas, in *M. bovis*, no increased expression of ESAT-6 genes was detected. Similar results were obtained in a previous report on proteome analysis that showed a differential pattern of expression for Rv2346c (esxO) and Rv3620c (esxW) between *M. bovis* (BCG) and *M. tuberculosis*.[27] The ESAT6-like proteins have been a focus of attention as they are highly immunogenic, secreted proteins capable of inducing a strong T cell response in the host.[28,29] The observed polymorphism in

the expression pattern of genes encoding cell wall and secreted proteins correlates with the variation in their sequences and could be an important source of antigenic diversity.

The major difference in the secretome of both bacilli is the elevated expression in *M. bovis* of the two serodominant antigens MPB70 and MPB83.[30] Behr and colleagues have recently reported that in both bacilli the *mpb70*/*mpb83* genes are under the positive control of *sigK*. Lately, a mutation in the gene encoding anti-SigK has been shown to be responsible for the high level of expression of MPB70/MPB83 in *M. bovis*.[31,32] Accordingly, in our array analysis *mpb70* showed the highest fold difference value. In addition, a higher expression was also found in *M. bovis* for the gene *mpb83* and the neighboring genes *dipZ* and Mb2901. The orthologous genes in *M. tuberculosis* have been described as part of a putative operon.[33] Recent data showed that these genes are members of the SigK-RskA regulon.[32] Interestingly, although the expression of *mpb70* and *mpb83* is low in *M. tuberculosis*, it greatly increases upon macrophage infection,[32,34] suggesting an important *in vivo* function.

A distinctive pattern of expression was also observed for the genes encoding for the phosphate-specific transport (Pst) system. This system comprises a periplasmic phosphate-binding protein (PstS), two transmembrane channel-forming proteins (PstA and PstC) and a cytoplasmic ATP binding protein (PstB) that probably interacts with PstA-PstC. The cluster of genes encoding for these proteins is formed by three putative operons: *pstS3*/*pstC2*/*pstA1*, *pstS2*/*pknD* and *pstB*/*pstS1*/*pstC1*/*pstA2*.[35,36] Interestingly, we observed a higher expression of the *pstS1*/*pstC1*/*pstA2* operon in *M. tuberculosis*, whereas the *pstS3*/*pstC2*/*pstA1* genes were increased in *M. bovis.* Another gene highly expressed in *M. tuberculosis*, *pknD*, is a pseudogene in *M. bovis*, thought to be involved in phosphate transport regulation.[37] Thus, these results suggest the use of different mechanisms controlling the transport of phosphate in each bacteria.

The family of genes involved in lipid metabolism and cell wall composition are of particular relevance as their proteins have been related to *M. tuberculosis* pathogenesis. In *M. bovis*, a higher expression was found for *fadD26* and the *ppsA-D* genes, which codes for fatty-acid-coA ligase and polyketide synthases, respectively. These enzymes are involved in phtiocerol dimycocerosate (PDIM) and phenolphtiocerol glycolipids (PGLs) biosynthesis.[38,39] PGLs are not produced by most strains of *M. tuberculosis*, includin H37Rv and CDC1551 strains, due to a frameshift mutation in the *pks1* gene.[40] Noteworthy, some strains of *M. tuberculosis* that produce PGLs display a hypervirulent phenotype,[41,42] thus, supporting the role of these lipids in mycobacteria pathogenesis. Moreover, *fadD26* and *ppsA-E* genes have been associated with virulence by signature-tagged transposon mutagenesis and bioinformatics analyses.[43,44] We also detected an increased expression in *M. tuberculosis* of two clusters of genes with similar organization: *pks4*/*papA3*/*mmpL10* and *mmpL8*/*papA1*/*pks2*. The *pks2* and *mmpL8* genes codes for a polyketide synthase and a lipid transporter, respectively. Both genes are required for the biosynthesis and transport of SL-1, a sulfolipid found exclusively in *M. tuberculosis* and thought to be implicated in host-pathogen interactions.[45,46]

Special attention was given to transcriptional regulators as they play a crucial role in the survival of the mycobacteria. Within this group, we observed higher expression for the *M. tuberculosis* Rv2160c and Rv0232 genes, which code for putative transcriptional regulators of the TetR/AcrR-family. In contrast, seven transcriptional regulatory genes were differentially expressed in *M. bovis*. Among these, *whiB6* exhibited the highest degree of difference (30.5-fold). This gene is markedly up-regulated in *M. tuberculosis* cultured under SDS, ethanol, heat-shock or oxidative stress conditions.[47] Due to the high expression observed for whiB6 and keeping in mind that minor changes in the expression of transcriptional regulators may cause significant changes to the phenotype, we searched for the other members of the *whiB* family. The data showed that in *M. bovis*, *whiB3* and *whiB4* expression were close to the cut-off level (2.7 and 2.8 fold difference, respectively). Whereas, in M. *tuberculosis*, a similar result was

observed for *whiB1* (2.6 fold difference). These results support the view that differential gene regulation might contribute to the physiological differences observed between these pathogens.

The presence of several deletions in one genome relative to the others have been proposed as key features for evolution in members of the *M. tuberculosis* complex.[4,48] We detected the expression of 31 genes annotated in the RD 4, 5, 6, 7, 8, 9, 10, 11 and 13 regions, present in *M. tuberculosis* but deleted in the *M. bovis* genome. These groups of genes include potential candidates for virulence factors and phenotype variation such as phospholipase C, prophages, ESAT-6 and PE family of proteins. In comparison, the expression of five genes in the RvD regions 1, 2 and 5, were detected in *M. bovis*. These genes code for a trans-membrane transporter protein, two hypothetical proteins, a molybdenum cofactor biosynthesis protein and a putative transcriptional regulator. Further research is needed to assess the role of the deleted regions in the physiology of the pathogen. So far, only two deletions have been attributed function across the *M. tuberculosis* complex; RD1 plays a key role in the attenuation of BCG, [49] while Wilkinson and colleagues recently described loss of RD750 as being involved in host interaction of the EAI clade of *M. tuberculosis*.[50]

An unexpected result was the finding of expression for some genes annotated in the *M. tuberculosis* CDC1551 genome whose annotation was missing from the *M. bovis* or *M. tuberculosis* H37Rv genomes. These genes hybridized to a set of probes designed for CDC1551 open reading frames. BLAST analysis of the probes sequences showed no significant identity with the annotated coding sequences of *M. bovis* and *M. tuberculosis* H37Rv genomes. Similar results were previously found at the proteome level by analysis of H37Rv supernatant proteins. [27] One of the identified genes with higher expression in *M. bovis* is MT2941. This gene is located between *mpb83* and *dipZ*, which suggest it might be co-transcribed as part of the *mpb83*/Mb2901 region. A previous report has shown that when comparing CDC1551 and H37Rv strain, important differences where found in gene prediction due to sequence polymorphism.[51] In this respect, we detected a high expression in *M. bovis* of MT1812, a gene absent in *M. tuberculosis* H37Rv and non-predicted in *M. bovis* 2122. Although the genes detected in our study were not included in the re-annotation of the *M. tuberculosis* H37Rv genome,[52] the data presented here suggests that gene prediction approaches used in the annotation of the *M. tuberculosis* H37Rv and *M. bovis* 2122 genomes have overlooked some coding sequences and hence, their annotation needs to be updated.

A 15 % variation in gene expression among *M. tuberculosis* clinical isolates[10] has been previously reported. Due to differences in the experimental approach, it is not feasible to directly compare that report with our data. However, a similar tendency in the distribution pattern of differentially expressed genes can be depicted across functional categories.

The transcriptional analysis presented in this work showed a good correlation with previously reported differences at the proteome level and with differences predicted by comparative genomics, which supports the validity of our study. Mutational studies will help to understand the particular contribution of variable genes in the mechanisms underlying the differences between both pathogens.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Alejandro Araya (C.N.R.S., Bordeaux Cedex, France) for his helpful comments and review of the manuscript. We also thank Eugenie Dubnau (TB Center, P.H.R.I) for valuable discussions and Saleena Ghanny (CAG, P.H.R.I), for her advice with microarray processing. This work was funded by FONDEF grant D02I1111

(awarded to A.M.Z.) and NIH grant HL068513 (awarded to I.S.). G.R. was supported by MECESUP AUS0006 fellowship and DID-UACH grant D2004-08.

References

1. Grange JM. Mycobacterium bovis infection in human beings. Tuberculosis (Edinb) 2001;81:71–77. [PubMed: 11463226]
2. Grange JM, Daborn C, Cosivi O. HIV-related tuberculosis due to Mycobacterium bovis. Eur Respir J 1994;7:1564–1566. [PubMed: 7995382]
3. Garnier T, Eiglmeier K, Camus JC, Medina N, Mansoor H, Pryor M, Duthoy S, Grondin S, Lacroix C, Monsempe C, Simon S, Harris B, Atkin R, Doggett J, Mayes R, Keating L, Wheeler PR, Parkhill J, Barrell BG, Cole ST, Gordon SV, Hewinson RG. The complete genome sequence of Mycobacterium bovis. Proc Natl Acad Sci U S A 2003;100:7877–7882. [PubMed: 12788972]
4. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, van Soolingen D, Cole ST. A new evolutionary scenario for the Mycobacterium tuberculosis complex. Proc Natl Acad Sci U S A 2002;99:3684–3689. [PubMed: 11891304]
5. Dunn PL, North RJ. Virulence ranking of some Mycobacterium tuberculosis and Mycobacterium bovis strains according to their ability to multiply in the lungs, induce lung pathology, and cause mortality in mice. Infect Immun 1995;63:3428–3437. [PubMed: 7642273]
6. Medina E, Ryan L, LaCourse R, North RJ. Superior virulence of Mycobacterium bovis over Mycobacterium tuberculosis (Mtb) for Mtb-resistant and Mtb-susceptible mice is manifest as an ability to cause extrapulmonary disease. Tuberculosis (Edinb) 2006;86:20–27. [PubMed: 16253563]
7. Collins DM. Virulence factors of Mycobacterium bovis. Tuberculosis (Edinb) 2001;81:97–102. [PubMed: 11463229]
8. Cosma CL, Sherman DR, Ramakrishnan L. The secret lives of the pathogenic mycobacteria. Annu Rev Microbiol 2003;57:641–676. [PubMed: 14527294]
9. Boshoff HI, Myers TG, Copp BR, McNeil MR, Wilson MA, Barry CE 3rd. The transcriptional responses of Mycobacterium tuberculosis to inhibitors of metabolism: novel insights into drug mechanisms of action. J Biol Chem 2004;279:40174–40184. [PubMed: 15247240]
10. Gao Q, Kripke KE, Saldanha AJ, Yan W, Holmes S, Small PM. Gene expression diversity among Mycobacterium tuberculosis clinical isolates. Microbiology 2005;151:5–14. [PubMed: 15632420]
11. Walters SB, Dubnau E, Kolesnikova I, Laval F, Daffe M, Smith I. The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. Mol Microbiol 2006;60:312–330. [PubMed: 16573683]
12. Manganelli R, Voskuil MI, Schoolnik GK, Smith I. The Mycobacterium tuberculosis ECF sigma factor sigmaE: role in global gene expression and survival in macrophages. Mol Microbiol 2001;41:423–437. [PubMed: 11489128]
13. Wilson M, DeRisi J, Kristensen HH, Imboden P, Rane S, Brown PO, Schoolnik GK. Exploring drug-induced alterations in gene expression in Mycobacterium tuberculosis by microarray hybridization. Proc Natl Acad Sci U S A 1999;96:12833–12838. [PubMed: 10536008]
14. Schoolnik GK, Voskuil MI, Schnappinger D, Yildiz FH, Meibom K, Dolganov NA, Wilson MA, Chong KH. Whole genome DNA microarray expression analysis of biofilm development by Vibrio cholerae O1 E1 Tor. Methods Enzymol 2001;336:3–18. [PubMed: 11398407]
15. Mostowy S, Cleto C, Sherman DR, Behr MA. The Mycobacterium tuberculosis complex transcriptome of attenuation. Tuberculosis (Edinb) 2004;84:197–204. [PubMed: 15207489]
16. Quackenbush J. Microarray data normalization and transformation. Nat Genet 2002;32(Suppl):496–501. [PubMed: 12454644]
17. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc Ser B 1995;57:289–300.
18. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 2000;132:365–386. [PubMed: 10547847]

19. Dubnau E, Fontan P, Manganelli R, Soares-Appel S, Smith I. Mycobacterium tuberculosis genes induced during infection of human macrophages. Infect Immun 2002;70:2787–2795. [PubMed: 12010964]
20. Talaat AM, Howard ST, Hale Wt, Lyons R, Garner H, Johnston SA. Genomic DNA standards for gene expression profiling in Mycobacterium tuberculosis. Nucleic Acids Res 2002;30:e104. [PubMed: 12384606]
21. Butcher PD. Microarrays for Mycobacterium tuberculosis. Tuberculosis (Edinb) 2004;84:131–137. [PubMed: 15207482]
22. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27–30. [PubMed: 10592173]
23. Keating LA, Wheeler PR, Mansoor H, Inwald JK, Dale J, Hewinson RG, Gordon SV. The pyruvate requirement of some members of the Mycobacterium tuberculosis complex is due to an inactive pyruvate kinase: implications for in vivo growth. Mol Microbiol 2005;56:163–174. [PubMed: 15773987]
24. Banu S, Honore N, Saint-Joanis B, Philpott D, Prevost MC, Cole ST. Are the PEPGRS proteins of Mycobacterium tuberculosis variable surface antigens? Mol Microbiol 2002;44:9–19. [PubMed: 11967065]
25. Brennan MJ, Delogu G. The PE multigene family: a 'molecular mantra' for mycobacteria. Trends Microbiol 2002;10:246–249. [PubMed: 11973159]
26. Brennan MJ, Delogu G, Chen Y, Bardarov S, Kriakov J, Alavi M, Jacobs WR Jr. Evidence that mycobacterial PE_PGRS proteins are cell surface constituents that influence interactions with other cells. Infect Immun 2001;69:7326–7333. [PubMed: 11705904]
27. Mattow J, Schaible UE, Schmidt F, Hagens K, Siejak F, Brestrich G, Haeselbarth G, Muller EC, Jungblut PR, Kaufmann SH. Comparative proteome analysis of culture supernatant proteins from virulent Mycobacterium tuberculosis H37Rv and attenuated M. bovis BCG Copenhagen. Electrophoresis 2003;24:3405–3420. [PubMed: 14595687]
28. Brodin P, Rosenkrands I, Andersen P, Cole ST, Brosch R. ESAT-6 proteins: protective antigens and virulence factors? Trends Microbiol 2004;12:500–508. [PubMed: 15488391]
29. Sorensen AL, Nagai S, Houen G, Andersen P, Andersen AB. Purification and characterization of a low-molecular-mass T-cell antigen secreted by Mycobacterium tuberculosis. Infect Immun 1995;63:1710–1717. [PubMed: 7729876]
30. Hewinson RG, Michell SL, Russell WP, McAdam RA, Jacobs WR Jr. Molecular characterization of MPT83: a seroreactive antigen of Mycobacterium tuberculosis with homology to MPT70. Scand J Immunol 1996;43:490–499. [PubMed: 8633206]
31. Charlet D, Mostowy S, Alexander D, Sit L, Wiker HG, Behr MA. Reduced expression of antigenic proteins MPB70 and MPB83 in Mycobacterium bovis BCG strains due to a start codon mutation in sigK. Mol Microbiol 2005;56:1302–1313. [PubMed: 15882422]
32. Said-Salim B, Mostowy S, Kristof AS, Behr MA. Mutations in Mycobacterium tuberculosis Rv0444c, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in Mycobacterium bovis. Mol Microbiol. 2006
33. Juarez MD, Torres A, Espitia C. Characterization of the Mycobacterium tuberculosis region containing the mpt83 and mpt70 genes. FEMS Microbiol Lett 2001;203:95–102. [PubMed: 11557146]
34. Schnappinger D, Ehrt S, Voskuil MI, Liu Y, Mangan JA, Monahan IM, Dolganov G, Efron B, Butcher PD, Nathan C, Schoolnik GK. Transcriptional Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the Phagosomal Environment. J Exp Med 2003;198:693–704. [PubMed: 12953091]
35. Torres A, Juarez MD, Cervantes R, Espitia C. Molecular analysis of Mycobacterium tuberculosis phosphate specific transport system in Mycobacterium smegmatis. Characterization of recombinant 38 kDa (PstS-1). Microb Pathog 2001;30:289–297. [PubMed: 11373123]
36. Lefevre P, Braibant M, de Wit L, Kalai M, Roeper D, Grotzinger J, Delville JP, Peirs P, Ooms J, Huygen K, Content J. Three different putative phosphate transport receptors are encoded by the Mycobacterium tuberculosis genome and are present at the surface of Mycobacterium bovis BCG. J Bacteriol 1997;179:2900–2906. [PubMed: 9139906]

37. Peirs P, Parmentier B, De Wit L, Content J. The Mycobacterium bovis homologous protein of the Mycobacterium tuberculosis serine/threonine protein kinase Mbk (PknD) is truncated. FEMS Microbiol Lett 2000;188:135–139. [PubMed: 10913696]
38. Azad AK, Sirakova TD, Fernandes ND, Kolattukudy PE. Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria. J Biol Chem 1997;272:16741–16745. [PubMed: 9201977]
39. Camacho LR, Constant P, Raynaud C, Laneelle MA, Triccas JA, Gicquel B, Daffe M, Guilhot C. Analysis of the phthiocerol dimycocerosate locus of Mycobacterium tuberculosis. Evidence that this lipid is involved in the cell wall permeability barrier. J Biol Chem 2001;276:19845–19854. [PubMed: 11279114]
40. Constant P, Perez E, Malaga W, Laneelle MA, Saurel O, Daffe M, Guilhot C. Role of the pks15/1 gene in the biosynthesis of phenolglycolipids in the Mycobacterium tuberculosis complex. Evidence that all strains synthesize glycosylated phydroxybenzoic methly esters and that strains devoid of phenolglycolipids harbor a frameshift mutation in the pks15/1 gene. J Biol Chem 2002;277:38148–38158. [PubMed: 12138124]
41. Reed MB, Domenech P, Manca C, Su H, Barczak AK, Kreiswirth BN, Kaplan G, Barry CE 3rd. A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. Nature 2004;431:84–87. [PubMed: 15343336]
42. Tsenova L, Ellison E, Harbacheuski R, Moreira AL, Kurepina N, Reed MB, Mathema B, Barry CE 3rd, Kaplan G. Virulence of selected Mycobacterium tuberculosis clinical isolates in the rabbit model of meningitis is dependent on phenolic glycolipid produced by the bacilli. J Infect Dis 2005;192:98–106. [PubMed: 15942899]
43. Camacho LR, Ensergueix D, Perez E, Gicquel B, Guilhot C. Identification of a virulence gene cluster of Mycobacterium tuberculosis by signature-tagged transposon mutagenesis. Mol Microbiol 1999;34:257–267. [PubMed: 10564470]
44. Zubrzycki IZ. Analysis of the products of genes encompassed by the theoretically predicted pathogenicity islands of Mycobacterium tuberculosis and Mycobacterium bovis. Proteins 2004;54:563–568. [PubMed: 14748003]
45. Converse SE, Mougous JD, Leavell MD, Leary JA, Bertozzi CR, Cox JS. MmpL8 is required for sulfolipid-1 biosynthesis and Mycobacterium tuberculosis virulence. Proc Natl Acad Sci U S A 2003;100:6121–6126. [PubMed: 12724526]
46. Sirakova TD, Thirumala AK, Dubey VS, Sprecher H, Kolattukudy PE. The Mycobacterium tuberculosis pks2 gene encodes the synthase for the hepta- and octamethyl-branched fatty acids required for sulfolipid synthesis. J Biol Chem 2001;276:16833–16839. [PubMed: 11278910]
47. Geiman DE, Raghunand TR, Agarwal N, Bishai WR. Differential Gene Expression in Response to Exposure to Antimycobacterial Agents and Other Stress Conditions among Seven Mycobacterium tuberculosis whiB-Like Genes. Antimicrob Agents Chemother 2006;50:2836–2841. [PubMed: 16870781]
48. Gordon SV, Brosch R, Billault A, Garnier T, Eiglmeier K, Cole ST. Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. Mol Microbiol 1999;32:643–655. [PubMed: 10320585]
49. Pym AS, Brodin P, Brosch R, Huerre M, Cole ST. Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti. Mol Microbiol 2002;46:709–717. [PubMed: 12410828]
50. Newton SM, Smith RJ, Wilkinson KA, Nicol MP, Garton NJ, Staples KJ, Stewart GR, Wain JR, Martineau AR, Fandrich S, Smallie T, Foxwell B, Al-Obaidi A, Shafi J, Rajakumar K, Kampmann B, Andrew PW, Ziegler-Heitbrock L, Barer MR, Wilkinson RJ. A deletion defining a common Asian lineage of Mycobacterium tuberculosis associates with immune subversion. Proc Natl Acad Sci U S A 2006;103:15594–15598. [PubMed: 17028173]
51. Fleischmann RD, Alland D, Eisen JA, Carpenter L, White O, Peterson J, DeBoy R, Dodson R, Gwinn M, Haft D, Hickey E, Kolonay JF, Nelson WC, Umayam LA, Ermolaeva M, Salzberg SL, Delcher A, Utterback T, Weidman J, Khouri H, Gill J, Mikula A, Bishai W, Jacobs Jr WR Jr. Venter JC, Fraser CM. Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains. J Bacteriol 2002;184:5479–5490. [PubMed: 12218036]

52. Camus JC, Pryor MJ, Medigue C, Cole ST. Re-annotation of the genome sequence of Mycobacterium tuberculosis H37Rv. Microbiology 2002;148:2967–2973. [PubMed: 12368430]

Table 1

Genes with higher expression level in *M. bovis* compared with *M. tuberculosis* H37Rv

Mtb CDS	Mb CDS	Gene	Fold difference[a]	*p*-value	RvD region	Predicted function
Rv0065	Mb0066		14.8	0.00384		Conserved hypothetical protein
Rv0077c	Mb0079c		4.6	0.0415		Probable oxidoreductase
Rv0215c	Mb0221c	*fadE3*	4.1	0.00384		Probable acyl-coa dehydrogenase
Rv0216	Mb0222		10.8	0.0086		Conserved hypothetical protein
Rv0260c	Mb0266c		4.3	0.0133		Posible transcriptional regulatory protein
Rv0448c	Mb0456c		11.1	0.0111		Conserved hypothetical protein
Rv0449c	Mb0457c		27.5	0.00568		Conserved hypothetical protein
Rv0465c	Mb0474c		6	0.0141		Posible transcriptional regulatory protein
Rv0549c	Mb0563c		4.3	0.0179		Conserved hypothetical protein
Rv0550c	Mb0565c		3.8	0.027		Hypothetical protein
Rv0591	Mb0606	*mce2C*	9.2	0.0219		Thought to be involved in host cell invasion
Rv0592	Mb0607	*mce2Da*	8	0.0045		Thought to be involved in host cell invasion
Rv0619	Mb0635	*galT*	17.3	0.00073		Probable galactose 1-phosphate uridyl transferase
Rv0769	Mb0792		9.7	0.00384		Probable dehydrogenase/reductase
Rv0770	Mb0793		3.4	0.0111		Probable dehydrogenase/reductase
Rv0771	Mb0794		3.7	0.00802		Possible 4-carboxymuconolactone decarboxylase
Rv0782	Mb0804	*ptrBb*	25.5	0.00384		Probable protease II (oligopeptidase b)
Rv0892	Mb0916		5.5	0.00389		Probable monooxygenase
Rv0928	Mb0951	*pstS3*	4.5	0.00562		Periplasmic phosphate-binding lipoprotein
Rv0929	Mb0952	*pstC2*	3	0.0129		Phosphate-transport integral membrane ABC transporter
Rv0930	Mb0953	*pstA1*	3	0.0049		Phosphate-transport integral membrane ABC transporter
Rv1135c	Mb1166c	*PPE16*	7	0.00664		PPE family protein
Rv1421	Mb1456		3.6	0.00523		Conserved hypothetical protein
Rv1493	Mb1530	*mutB*	3.1	0.0224		Probable methylmalonyl-coa mutase
Rv1588c	Mb1614c		3.3	0.0164		Partial REP13E12 repeat protein
Rv1652	Mb1680	*argC*	4.1	0.0366		Probable n-acetyl-gamma-glutamyl-phoshate reductase
Rv1653	Mb1681	*argJ*	3.1	0.0147		Probable glutamate n-acetyltransferase
Rv1656	Mb1684	*argF*	3	0.0152		Probable ornithine carbamoyltransferase
MT1802	Mb1787	*mmpL14*	27.7	0.00562	2	Probable conserved transmembrane transport protein
Rv1758	Mb1788	*cut1*	28	0.0289		Probable cutinase
Rv1947	Mb1982		3.3	0.0426		Hypothetical protein
Rv2004c	Mb2027c		3.7	0.0266		Conserved hypothetical protein
Rv2005c	Mb2028c		3.7	0.0105		Conserved hypothetical protein
MT2081	Mb2048c		20.6	0.0166	1	Conserved hypothetical protein
Rv2025c	Mb2050c		5.4	0.00802		Possible conserved membrane protein
Rv2277c	Mb2300c		21.3	0.00384		Possible glycerolphosphodiesterase
Rv2386c	Mb2407c	*mbtI*	7	0.00476		Putative isochorismate synthase
Rv2421c	Mb2444c	*nadD*	3.1	0.0119		Probable nicotinate-nucleotide adenylyltransferase
Rv2431c	Mb2457c	*PE25*	3	0.0336		PE family protein
Rv2435c	Mb2461c		3.2	0.0134		Probable cyclase
Rv2562	Mb2591		5.8	0.0166		Conserved hypothetical protein
Rv2619c	Mb2652c		3.8	0.0255		Conserved hypothetical protein
Rv2620c	Mb2653c		10	0.00384		Probable conserved transmembrane protein
Rv2621c	Mb2654c		13.1	0.00362		Possible transcriptional regulatory protein
Rv2622	Mb2655		13.4	0.00137		Possible methyltransferase
Rv2623	Mb2656	*TB31.7*	4.1	0.0105		Conserved hypothetical protein
Rv2629	Mb2662		3.1	0.00384		Conserved hypothetical protein
Rv2779c	Mb2801c		3.2	0.0181		Possible transcriptional regulatory protein
Rv2789c	Mb2812c	*fadE21*	3.5	0.0206		Probable acyl-coa dehydrogenase
Rv2873	Mb2898	*mpb83*	30.9	0.00361		Cell surface lipoprotein
Rv2874	Mb2899	*dipz*	11.5	0.00362		Possible membrane cytochrome biogenesis protein
Rv2875	Mb2900	*mpb70*	118	0.00384		Major secreted immunogenic protein

Mtb CDS	Mb CDS	Gene	Fold difference[a]	*p*-value	RvD region	Predicted function
Rv2876	Mb2901		5.7	0.0049		Possible conserved transmembrane protein
Rv2917	Mb2941		7.9	0.0086		Conserved hypothetical alanine and arginine rich protein
Rv2930	Mb2955	*fadD26*	5.8	0.0086		Fatty-acid-coa ligase
Rv2931	Mb2956	*ppsA*	10.7	0.00384		Phenolpthiocerol synthesis type-I polyketide synthase
Rv2932	Mb2957	*ppsB*	6.9	0.0166		Phenolpthiocerol synthesis type-I polyketide synthase
Rv2933	Mb2958	*ppsC*	7.9	0.0049		Phenolpthiocerol synthesis type-I polyketide synthase
Rv2934	Mb2959	*ppsD*	3.9	0.0166		Phenolpthiocerol synthesis type-I polyketide synthase
Rv2945c	Mb2970c	*lppX*	3.3	0.0226		Probable conserved lipoprotein
Rv2955c	Mb2979c		4.5	0.00568		Conserved hypothetical protein
Rv2972c	Mb2997c		3.8	0.0172		Possible conserved membrane or exported protein
Rv2973c	Mb2998c	*recG*	7.3	0.0166		Probable atp-dependent dna helicase
Rv2974c	Mb2999c		5.3	0.0119		Conserved hypothetical alanine rich protein
Rv2987c	Mb3011c	*leuD*	4.3	0.015		Probable 3-isopropylmalate dehydratase (small subunit)
Rv2988c	Mb3012c	*leuC*	3.4	0.0336		Probable 3-isopropylmalate dehydratase (large subunit)
Rv2989	Mb3013		6	0.00568		Probable transcriptional regulatory protein
Rv2998A	Mb3023c		3.8	0.00384		Conserved hypothetical protein
Rv3054c	Mb3080c		5.4	0.0331		Conserved hypothetical protein
Rv3082c	Mb3109c	*virS*	4.2	0.019		Virulence-regulating transcriptional regulator
MT3248	Mb3184c	*PPE70*	7.5	0.0166		PPE family protein
Rv3324c	Mb3353c	*moaC3*	13	0.00612		Probable molybdenum cofactor biosynthesis protein C3
MT3427	Mb3355c	*moaA3*	10.3	0.0133	5	Probable molybdenum cofactor biosynthesis protein A
MT3427.1	Mb3356		22.6	0.0119	5	Hypothetical protein
MT3428	Mb3358	*embR2*	8	0.027	5	Possible transcriptional regulatory protein
Rv3331	Mb3364	*sugI*	15.5	0.00401		Probable sugar-transport integral membrane protein
Rv3332	Mb3365	*nagA*	5.2	0.0413		Probable n-acetylglucosamine-6-phosphate deacetylase
Rv3340	Mb3372	*metC*	3.1	0.015		Probable o-acetylhomoserine sulfhydrylase
Rv3397c	Mb3430c	*phyA*	3.4	0.0219		Probable phytoene synthase
Rv3398c	Mb3431c	*idsA*	4.4	0.0086		Probable geranylgeranyl pyrophosphate synthetase
Rv3453	Mb3483		6.2	0.0184		Possible conserved transmembrane protein
Rv3454	Mb3483		3.2	0.0086		Probable conserved integral membrane protein
Rv3466	Mb3495		3.5	0.0086		Conserved hypothetical protein
Rv3530c	Mb3560c		3.9	0.0224		Possible oxidoreductase
Rv3651	Mb3675		3.2	0.0131		Conserved hypothetical protein
Rv3697c	Mb3723c		3.1	0.00641		Possible conserved membrane protein
Rv3862c	Mb3892c	*whiB6*	30.4	0.00663		WhiB-like possible transcriptional regulatory protein
Rv3897c	Mb3927c		5.7	0.00523		Conserved hypothetical protein
MT0573.1			3.5	0.0105		Hypothetical protein
MT0915.1			6.3	0.00568		Hypothetical protein
MT1812			55.5	0.00384		Hypothetical protein
MT1813			62.5	0.0266		Hypothetical protein
MT2083			5.1	0.00568		Hypothetical protein
MT2941			42.6	0.00861		Hypothetical protein
MT3718.1			6.6	0.0179		Hypothetical protein

[a]Fold differences are the average of normalized intensity ratios from 4 microarray experiments using 4 independent biological replicates.

Table 2

Genes with higher expression level in *M. tuberculosis* H37Rv compared with *M. bovis*

Mtb CDC	Gene	Fold difference[a]	*p*-value	RD region	Predicted function
Rv0032	*bioF2*	3.7	0.0383		Possible 8-amino-7-oxononanoate synthase
Rv0040c	*mtc28*	3.1	0.0365		Secreted proline rich protein
Rv0112	*gca*	3.1	0.0146		GDP-D-mannose dehydratase
Rv0166	*fadD5*	3.3	0.0157		Fatty-acid-coa ligase
Rv0167	*yrbe1a*	3.6	0.0233		Conserved hypothetical integral membrane protein
Rv0221		42.6	0.00818		Conserved hypothetical protein
Rv0222	*echA1*	10.4	0.0235	10	Probable enoyl-coa hydratase
Rv0223c		12.9	0.0227		Probable aldehyde dehydrogenase
Rv0232		4	0.0196		Probable transcriptional regulatory protein (tetr/acrr-family)
Rv0276		40.5	0.00818		Conserved hypothetical protein
Rv0520		7.4	0.0487		Possible methyltransferase/methylase
Rv0544c		6.1	0.00659		Possible Conserved transmembrane protein
Rv0794c		4	0.0487		Probable oxidoreductase
Rv0796		6.5	0.00268		Putative transposase for insertion sequence IS6110
Rv0888		4.7	0.0108		Probable exported protein
Rv0931c	*pknD*	10.6	0.00659		Transmembrane serine/threonine-protein kinase
Rv0934	*pstS1*	7.3	0.0108		Periplasmic phosphate-binding lipoprotein
Rv0935	*pstC1*	4.3	0.016		Phosphate-transport integral membrane ABC transporter
Rv0936	*pstA2*	4.2	0.032		Phosphate-transport integral membrane ABC transporter
Rv1038c	*esxJ*	3.2	0.00707		Esat-6 like protein
Rv1076	*lipU*	5.2	0.0235		Possible lipase
Rv1130		28.1	0.00659		Conserved hypothetical protein
Rv1131	*gltA1*	28.4	0.00899		Probable citrate synthase
Rv1172c	*PE12*	9.1	0.00659		PE family protein
Rv1181	*pks4*	3.9	0.00707		Probable polyketide beta-ketoacyl synthase
Rv1182	*papA3*	5.4	0.0146		Probable polyketide synthase associated protein
Rv1183	*mmpL10*	4.3	0.0121		Probable Conserved transmembrane transport protein
Rv1184c		7.6	0.00268		Possible exported protein
Rv1195	*PE13*	7.7	0.00818		PE family protein
Rv1196	*PPE18*	14.8	0.00268		PPE family protein
Rv1197	*esxK*	3.2	0.00268		Esat-6 like protein
Rv1200		3.1	0.0235		Probable Conserved integral membrane transport protein
Rv1220c		3	0.0105		Probable methyltransferase
Rv1257c		28.4	0.0183	13	Probable oxidoreductase
Rv1361c	*PPE19*	11.7	0.00659		PPE family protein
Rv1369c		4.2	0.0156		Probable transposase
Rv1370c		7.9	0.0227		Putative transposase for insertion sequence IS6110
Rv1386	*PE15*	6	0.0108		PE family protein
Rv1387	*PPE20*	8.5	0.00966		PPE family protein
Rv1396c	*PE PGRS25*	6.2	0.0356		PE PGRS family protein
Rv1397c		7.1	0.0369		Conserved hypothetical protein
Rv1398c		5.6	0.0248		Conserved hypothetical protein
Rv1497	*lipL*	3.2	0.0477		Probable esterase
Rv1506c		15	0.0149	4	hypothetical protein
Rv1507c		31.9	0.0217	4	Conserved hypothetical protein
Rv1508c		44.4	0.0252	4	Probable membrane protein
Rv1509		5.6	0.0486	4	hypothetical protein
Rv1510		4.3	0.0365	4	Probable Conserved membrane protein
Rv1511	*gmdA*	40.5	0.016	4	GDP-D-mannose dehydratase
Rv1512	*epiA*	22.5	0.0393	4	Probable nucleotide-sugar epimerase
Rv1513		8.7	0.00351	4	Conserved hypothetical protein
Rv1515c		69.4	0.00739	4	Conserved hypothetical protein

Mtb CDC	Gene	Fold difference[a]	*p*-value	RD region	Predicted function
Rv1516c		15.2	0.0235	4	Probable sugar transferase
Rv1535		3.4	0.0208		Hypothetical protein
Rv1563c	*treY*	6.7	0.0435		Maltooligosyltrehalose synthase
Rv1611	*trpC*	6	0.00659		Probable indole-3-glycerol phosphate synthase
Rv1612	*trpB*	6.1	0.00268		Probable tryptophan synthase, beta subunit
Rv1613	*trpA*	7	0.0183		Probable tryptophan synthase, alpha subunit
Rv1614	*lgt*	11.1	0.00651		Possible prolipoprotein diacylglyceryl transferases
Rv1639c		5.4	0.00268		Conserved hypothetical membrane protein
Rv1646	*PE17*	3.2	0.0291		PE family protein
Rv1651c	*PE PGRS30*	16.2	0.0421		PE PGRS family protein
Rv1730c		5	0.00268		Possible Penicillin-binding protein
Rv1757c		6.9	0.016		Putative transposase for insertion sequence IS6110
Rv1763		7.1	0.0162		Putative transposase for insertion sequence IS6110
Rv1764		5.4	0.0116		Putative transposase for insertion sequence IS6110
Rv1792	*esxM*	3.2	0.00707		Esat-6 like protein
Rv1809	*PPE33*	11.2	0.00659		PPE family protein
Rv1872c	*lldD2*	4.5	0.0296		Possible l-lactate dehydrogenase
Rv1884c	*rpfC*	5.5	0.0288		Probable resuscitation-promoting factor
Rv1925	*fadD31*	3.1	0.00984		Possible dehydrogenase
Rv1965	*yrbE3B*	14.1	0.0217	7	Conserved hypothetical integral membrane protein
Rv1976c		29.3	0.0176	7	Conserved hypothetical protein
Rv1977		6.9	0.0452	7	Conserved hypothetical protein
Rv2071c	*cobM*	3.7	0.00765		Precorrin-4 c11-methyltransferase
Rv2073c		4.8	0.032	9	Probable shortchain dehydrogenase
Rv2074		83.3	0.00268	9	Conserved hypothetical protein
Rv2077c		13	0.0108		Possible Conserved transmembrane protein
Rv2105		8	0.0157		Putative transposase for insertion sequence IS6110
Rv2106		5.2	0.016		Probable transposase for insertion sequence IS6110
Rv2137c		5.7	0.0157		Conserved hypothetical protein
Rv2159c		24.3	0.00965		Conserved hypothetical protein
Rv2160c		44.4	0.0121		Conserved hypothetical protein (putative tetr family)
Rv2161c		9.8	0.00268		Conserved hypothetical protein
Rv2162c	*PE PGRS38*	17.9	0.00659		PE PGRS family protein
Rv2167c		4.3	0.00268		Probable transposase
Rv2168c		9.2	0.0156		Putative transposase for insertion sequence IS6110
Rv2187	*fadD15*	5.6	0.016		Probable long-chain-fatty-acid-coa ligase
Rv2278		9.9	0.0103		Putative transposase for insertion sequence IS6110
Rv2279		6.6	0.0121		Probable transposase
Rv2346c	*esxO*	3.1	0.0121	5	Putative Esat-6 like protein
Rv2347c	*esxP*	3.4	0.00268	5	Putative Esat-6 like protein
Rv2348c		87.7	0.00268	5	hypothetical protein
Rv2349c	*plcC*	67.6	0.00659	5	Probable phospholipase C3
Rv2350c	*plcB*	64.5	0.00966	5	Probable membrane-associated phospholipase C2
Rv2351c	*plcA*	19	0.016	5	Probable membrane-associated phospholipase C
Rv2352c	*PPE38*	69.4	0.00565	5	PPE family protein
Rv2354		9.5	0.016		Probable transposase for insertion sequence IS6110
Rv2355		4.2	0.0217		Probable transposase
Rv2395		3.9	0.00588		Probable Conserved integral membrane protein
Rv2479c		4	0.0266		Probable transposase
Rv2480c		6.5	0.0208		Possible transposase for insertion sequence IS6110
Rv2490c	*PEPGRS43*	3.7	0.00707		PE PGRS family protein
Rv2600		3.8	0.0268		Probable Conserved integral membrane protein
Rv2648		9.6	0.0126	11	Probable transposase for insertion sequence IS6110
Rv2649		6.7	0.0162	11	Probable transposase for insertion sequence IS6110
Rv2651c		6	0.0108	11	Possible phirv2 prophage protease

Mtb CDC	Gene	Fold difference[a]	*p*-value	RD region	Predicted function
Rv2657c		29.1	0.0188	11	Probable phirv2 prophage protein
Rv2814c		4.8	0.0176		Probable transposase
Rv2815c		8.6	0.0108		Probable transposase for insertion sequence IS6110
Rv2822c		3.7	0.0216		Hypothetical protein
Rv2823c		13.6	0.0472		Conserved hypothetical protein
Rv2877c	*merT*	54.1	0.00659		Probable Conserved integral membrane protein
Rv3061c	*fadE22*	3.1	0.0236		Probable acyl-coa dehydrogenase
Rv3083		5.3	0.0121		Probable monooxygenase
Rv3084	*lipR*	3.6	0.00966		Probable acetyl-hydrolase/esterase
Rv3086	*adhD*	3.9	0.0299		Aldehyde reductase
Rv3088		3.7	0.0277		Conserved hypothetical protein
Rv3093c		6.6	0.0176		Hypothetical oxidoreductase
Rv3094c		10.9	0.0176		Conserved hypothetical protein
Rv3136	*PPE51*	47.2	0.00707		PPE family protein
Rv3184		10.4	0.00707		Probable transposase for insertion sequence IS6110
Rv3185		38.5	0.032		Probable transposase
Rv3186		10.4	0.00818		Probable transposase for insertion sequence IS6110
Rv3187		5.6	0.0108		Probable transposase
Rv3325		6.5	0.024		Probable transposase for insertion sequence IS6110
Rv3326		7.4	0.00872		Probable transposase
Rv3354		4.2	0.00707		Conserved hypothetical protein
Rv3380c		5	0.0162		Probable transposase
Rv3381c		8.5	0.0234		Probable transposase for insertion sequence IS6110
Rv3390	*lpqD*	3.1	0.0121		Probable Conserved lipoprotein
Rv3407		15.9	0.00818		Conserved hypothetical protein
Rv3408		13	0.00659		Conserved hypothetical protein
Rv3426	*PPE58*	30.5	0.0272	6	PPE family protein
Rv3429	*PPE*	45.7	0.00659		PPE family protein
Rv3474		11	0.016		Possible transposase for insertion sequence IS6110
Rv3475		4.4	0.0133		Possible transposase for insertion sequence IS6110
Rv3477	*PE31*	20.4	0.00659		PE family protein
Rv3478	*PPE60*	16.6	0.00268		PPE family protein
Rv3479		11.2	0.0235		Possible transmembrane protein
Rv3487c	*lipF*	12.9	0.00965		Probable esterase/lipase
Rv3618		65.8	0.00659	8	Possible monooxygenase
Rv3620c	*esxW*	3.9	0.0108	8	Putative Esat-6 like protein
Rv3623	*lpqG*	75.8	0.00659		Probable Conserved lipoprotein
Rv3633		3	0.0104		Conserved hypothetical protein
Rv3726		7.2	0.00818		Possible dehydrogenase
Rv3749c		5.2	0.0171		Conserved hypothetical protein
Rv3750c		9.3	0.00899		Possible excisionase
Rv3760		3.2	0.0255		Possible Conserved membrane protein
Rv3822		4.4	0.016		No description
Rv3823c	*mmpL8*	4.2	0.0291		Probable Conserved integral membrane transport protein
Rv3824c	*papA1*	10	0.0119		Probable polyketide synthase associated protein
Rv3825c	*pks2*	11.9	0.0116		Probable polyketide synthase
MT0291.3		3.7	0.0104		Hypothetical protein
MT1178		3.9	0.0176		Hypothetical protein
MT2420		17.7	0.0116		Conserved hypothetical protein
MT2421		6	0.0237		Conserved hypothetical protein
MT2466		6.8	0.00659		Hypothetical protein
MT2467		3.6	0.0149		Hypothetical protein
MT2880.1		3.8	0.0267		Hypothetical protein
MT3580.2		22.2	0.00659		Hypothetical protein
MT3846		25.2	0.0307		Hypothetical protein

Mtb CDC	Gene	Fold difference[a]	*p*-value	RD region	Predicted function
MT4026.1		8.9	0.00659		Hypothetical protein

[a]Fold differences are the average of normalized intensity ratios from 4 microarray experiments using 4 biological replicates.

Table 3

Validation of microarray results by RT-qPCR

CDS	Gene	Fold difference[a] *M. bovis/M. tuberculosis*
Mb0606	*mce2C*	16.2
Mb0607	*mce2D*	23
Mb2898	*mpb83*	38.8
Mb2900	*mpb70*	166.9
Mb2955	*fadD26*	3.4
Mb2956	*ppsA*	7.1
Mb2957	*ppsB*	3.6
Mb2958	*ppsC*	22
Mb2959	*ppsD*	28.6
Mb3358	*embR2*	173.3[b]
Mb3450	*WhiB3*	4.5
Mb3706c	*WhiB4*	3.1
Mb3892c	*WhiB6*	77.4

[a] Fold differences are averaged ratios from 3 biological replicates. Each ratio was calculated between the numbers of cDNA copies for each gene in both strains, normalized to *sigA*.

[b] The *embR2* gene is not present in the *M. tuberculosis* H37Rv genome, therefore, the value represent the absolute number of cDNA copies detected in *M. bovis*.

Table 4

Distribution of differentially expressed genes according to functional categories

Functional category[a]	*M. bovis*		*M. tuberculosis*	
	n	%	n	%
Cell wall and cell processes	17	17.9	31	19
Conserved hypotheticals	19	20	22	13.5
Conserved hypotheticals with an orthologue in *M. bovis/M. tuberculosis*	4	4.2	2	1.2
Information pathways	1	1.1	-	-
Insertion seqs and phages	2	2.1	33	20.2
Intermediary metabolism and respiration	22	23.2	30	18.4
Lipid metabolism	10	10.5	8	4.9
PE/PPE	3	3.2	18	11
Regulatory proteins	7	7.4	3	1.8
Unknown	-	-	3	1.8
Virulence, detoxification, adaptation	3	3.2	3	1.8
Hypothetical proteins CDC1551[b]	7	7.4	10	6.1
Total	95	100	163	100

[a] Genes were grouped according to functional classifications as annotated in TubercuList.

[b] ORFs annotated in the *M. tuberculosis* CDC1551 genome.