Microbiology (2007), **153**, 3323–3336 DOI 10.1099/mic.0.2007/009894-0

Comparative transcriptomics reveals key gene expression differences between the human and bovine pathogens of the *Mycobacterium tuberculosis* complex

Paul Golby,[1] Kim A. Hatch,[2] Joanna Bacon,[2] Rory Cooney,[1]† Paul Riley,[2] Jon Allnutt,[2] Jason Hinds,[3] Javier Nunez,[1] Philip D. Marsh,[2] R. Glyn Hewinson[1] and Stephen V. Gordon[1]

Correspondence
Stephen Gordon
s.v.gordon@vla.defra.gsi.gov.uk

[1]Veterinary Laboratories Agency (Weybridge), New Haw, Addlestone, Surrey KT15 3NB, UK

[2]Health Protection Agency, Centre for Emergency Preparedness and Response, Porton Down, Salisbury, Wiltshire SP4 0JG, UK

[3]Bacterial Microarray Group, Department of Cellular and Molecular Medicine, St George's Hospital Medical School, Cranmer Terrace, London SW17 0RE, UK

Members of the *Mycobacterium tuberculosis* complex show distinct host preferences, yet the molecular basis for this tropism is unknown. Comparison of the *M. tuberculosis* and *Mycobacterium bovis* genome sequences revealed no unique genes in the bovine pathogen per se, indicating that differences in gene expression may play a significant role in host predilection. To define the key gene expression differences between *M. tuberculosis* and *M. bovis* we have performed transcriptome analyses of cultures grown under steady-state conditions in a chemostat. This revealed that the human and bovine pathogens show differential expression of genes encoding a range of functions, including cell wall and secreted proteins, transcriptional regulators, PE/PPE proteins, lipid metabolism and toxin–antitoxin pairs. Furthermore, we probed the gene expression response of *M. tuberculosis* and *M. bovis* to an acid-shock perturbation which triggered a notably different expression response in the two strains. Through these approaches we have defined a core gene set that shows differential expression between the human and bovine tubercle bacilli, and the biological implications are discussed.

Received 18 May 2007
Revised 7 June 2007
Accepted 8 June 2007

INTRODUCTION

The *Mycobacterium tuberculosis* complex can be viewed as a set of ecotypes with distinct host preferences (Smith *et al.*, 2006). Hence, while *M. tuberculosis* can only be sustained in human populations, *Mycobacterium bovis* can infect and be maintained in populations of wild and domesticated animals such as cattle, deer and badgers. The molecular origin for these phenotypes is encoded in the genomes of the pathogens. The genome sequences of *M. tuberculosis* H37Rv and *M. bovis* 2122 are >99.95 % identical at the nucleotide level, with *M. bovis* containing no unique genes per se compared to *M. tuberculosis* (Cole *et al.*, 1998; Garnier *et al.*, 2003). Comparative analyses of the *M. tuberculosis* and *M. bovis* genomes have revealed the basis for distinguishing phenotypes such as the pyruvate requirement of *M. bovis* in glycerol-based media, or the reason for eugonic/dysgonic colony morphology (Keating *et al.*, 2005). However, comparative genomics in itself does not reveal the basis for the complexity of phenotype between *M. tuberculosis* and *M. bovis*. Extra information needs to be layered onto the genome data, such as gene expression profiling, metabolic network analyses, signalling pathways, etc., to fully explore the biology of these pathogens.

To explore the hypothesis that transcriptome differences play a role in the differing ecotypes of the *M. tuberculosis* complex, we set out to define the key gene expression differences between the human and bovine tubercle bacilli. To achieve this, the culture conditions for both organisms must be identical. Clearly, the nature of batch culture makes the system highly dynamic, complicating interpretation of expression profiles. Chemostat systems on the other

†Present address: Veterinary Exotic Diseases, Research and Official Controls Division (VEROD) Defra Area 101, 1a Page Street, London SW1P 4PQ, UK.

Abbreviations: DAT, diacyltrehalose; PAT, polyacyltrehalose; PDIM, phthiocerol dimycoserosate; PGL, phenolic glycolipid; PIN, PilT amino terminus; qRT-PCR, quantitative real-time PCR; TA, toxin–antitoxin.

Supplementary tables of primer sequences used for qRT-PCR reactions, and *M. bovis* and *M. tuberculosis* genes regulated in response to acid-shock are available with the online version of this paper.

hand allow tight control of bacterial growth in a defined, reproducible system (Hoskisson & Hobbs, 2005). We therefore chose to define the transcriptomes of the genome sequenced strains *M. tuberculosis* H37Rv and *M. bovis* 2122 under physiological steady-state conditions achieved through growth in carbon-limited chemostat cultures.

During the pathogenic process, *M. tuberculosis* and *M. bovis* are thought to be exposed to a number of different stress conditions, including acidic pH, reactive nitrogen and oxygen species, and nutrient starvation (Chan *et al.*, 2005). Many of the proteins induced in response to stress are thought to be involved in the survival of the pathogen inside the host. We therefore also chose to study the effect of acid-shock on the steady-state transcriptomes of *M. tuberculosis* and *M. bovis* as a means to probe the response of both strains to a perturbation.

Hence, the aim of this study was to identify key gene expression differences between *M. tuberculosis* H37Rv and *M. bovis* 2122 under both steady-state and perturbation conditions. We found that the expression of genes encoding a range of functional activities varied between the strains, and we discuss the possible biological impact of this variation on strain phenotype.

METHODS

Bacterial strains and growth conditions. For the continuous culture experiments, *M. tuberculosis* H37Rv and *M. bovis* AF2122 were grown at 37 °C in defined minimal medium, CAMR Mycobacteria Media (CMM) Mod6 (pH 7.0), at a dissolved oxygen tension of 10 % (50 % air saturation), using the approach described previously for *M. tuberculosis* H37Rv (Bacon *et al.*, 2004). CMM Mod6 was a modified version of CMM (James *et al.*, 2000) in which all amino acids were removed except for L-asparagine. Glycerol and glucose were omitted and Tween 80 (0.2 %, v/v) was included as sole carbon source. As is standard practice, BSA (0.5 % w/v) was included to ameliorate the toxic side effects of Tween 80 degradation. The cultures were performed in 1 l glass fermentation vessels, which were operated as chemostats by controlling medium addition and effluent removal to maintain a working volume of 500 ml. The chemostats were operated at a constant dilution rate of 0.03 h^{-1}, which corresponds to a mean generation time of 23 h. The cultures were maintained at a constant pH 7.0 ± 0.1 and a temperature of 37 °C ± 0.1. Cultures were monitored for viability by measuring optical density at 540 nm and total viable count as described by Bacon *et al.* (2004).

Acid-shock experiments. Acid-shock studies were performed in a chemostat. Three independent aerobic steady-state cultures of *M. tuberculosis* and *M. bovis* were established at pH 7.0, then when the pH had dropped to 5.5 the samples were rapidly removed for transcriptome analysis. At the start of each pH perturbation experiment, maximum and minimum pH levels were set on the controller unit in the range of pH 5.48–5.52. HCl (1 M) was automatically added dropwise to the culture until the desired pH of the culture was reached, i.e. pH 5.5. The addition of the first drop of acid was taken as time point zero and the time taken for the pH to fall from 7.0 to 5.5 was in the range of 2.5 to 5 min. The culture was then maintained at pH 5.5 via the automatic addition of 1 M NaOH or 1 M HCl. Samples (20 ml) were rapidly withdrawn from the culture at 5, 15, 30, 60 min and 24 h after time point zero and RNA was extracted as described previously (Bacon *et al.*, 2004). Residual DNA was removed with the use of a DNA*free* kit (Ambion) and the integrity of the RNA was checked by non-denaturing gel electrophoresis.

***M. tuberculosis–M. bovis* microarrays.** The *M. tuberculosis–M. bovis* composite microarrays used in these experiments were developed by the Bacterial Microarray Group (BμG@S, St Georges, University of London) in collaboration with the Veterinary Laboratories Agency (Weybridge). The array consists of 4410 PCR products (size range 60–1000 bp) that represent all the genes in the genomes of *M. tuberculosis* strains H37Rv and CDC1551 and *M. bovis* strain AF2122. Primer pairs for each ORF were designed with Primer 3 software and selected by BLAST analysis to have minimal cross-homology with all other ORFs. PCR products were generated and purified with the aid of a RoboAmp 4200 robot (MWG Biotech) and arrayed in duplicate onto UltraGap aminosilane-coated glass slides (Corning) by a Biorobotics Microgrid II microarrayer (Genomic Solutions). The quality of spots on the microarrays was evaluated using the Spot Check kit (Genetix Ltd). The array design is available in BμG@Sbase, accession no. A-BUGS-31 (http://bugs.sgul.ac.uk/A-BUGS-31) and also ArrayExpress, accession no. A-BUGS-31.

cDNA labelling and microarray hybridizations. The method used to synthesize labelled cDNA probes from RNA was similar to that described previously with some modifications (Stewart *et al.*, 2002). Complementary DNA synthesized from RNA was fluorescently labelled in a reverse transcription reaction in the presence of Cy5-CTP (Amersham Biosciences). In a total volume of 16 μl, RNA (3–5 μg) was mixed with 5 μg random hexamers, heated to 70 °C for 10 min and then snap-cooled on ice. To the chilled mixture, 6 μl 5 × first strand buffer, 3 μl 100 mM DTT, 15 nmol each of dATP, dGTP and dTTP, 6 nmol dCTP and 2.5 nmol Cy5-dCTP (Amersham Biosciences) were added. The labelling reaction was initiated by the addition of 400 units of Superscript II (Life Technologies), and the reactions were incubated for 10 min at 25 °C and then for 90 min at 42 °C.

cDNA synthesized using DNA as template was fluorescently labelled using Cy3-dCTP (Amersham Bioscience) and Klenow DNA polymerase. DNA (3 μg) was mixed with 3 μg random primers in a total volume of 41.5 μl, heated at 95 °C for 5 min and then cooled on ice. To the cooled mixture, 5 μl 10 × Klenow buffer, 5 nmol each of dATP, dGTP and dTTP, 2 nmol dCTP and 1.5 nmol Cy3-dCTP were added. The labelling reaction was initiated by the addition of 5 units Klenow fragment of DNA polymerase (New England Biolabs) and the reaction was incubated at 37 °C for 1.5 h. The Cy5- and Cy3-labelled cDNAs from the labelling reactions were co-purified in water using a mini elute purification kit (Qiagen). The volume of the purified labelled cDNAs was adjusted to 60 μl in 4 × SSC and 0.3 % SDS. Prior to use, the microarrays were prehybridized at 65 °C for 20 min in 50 ml of a solution containing 3.5 × SSC, 0.1 % SDS and 10 mg BSA ml^{-1}. The slides were washed at room temperature with H_2O for 1 min, propan-2-ol for 1 min and then dried by centrifugation. The labelled cDNAs were heated for 3 min at 95 °C, before being applied to the array under two 22 × 22 mm coverslips (LifterSlip; Erie Scientific).

Scanning and image analysis. Hybridized arrays were scanned using an Affymetrix 428 scanner with the photomultiplier tube gain set in the range 40–50, so that spots with the highest signal intensities were just below the level of saturation; this generated two images per microarray (one per channel) in TIF format. Fluorescent spots on each scanned image were quantified using Imagene 5.1 (Biodiscovery) using a segmentation signal percentage set to a high of 95 % and a low of 5 %. The signal for each spot was taken as the median of all pixel intensities within the defined area of the spot. Genomic DNA from *M. tuberculosis* was hybridized in the Cy3 channel, and was used as a common reference for all arrays.

Data normalization and statistical analysis. Data processing employed Mathematica Version 5.2 (Wolfram Research). For each channel background, values were first subtracted from signal values for each spot and values less than 0.01 were set to 0.01. For each microarray the log ratio of the Cy5 to Cy3 signal was calculated for every spot; each log ratio was then divided by the median of the log ratios of the microarray so as to centre the gene expression distributions of the whole set of microarrays to the same point (zero). After this step the microarrays for the *M. bovis* acid-shock experiment presented a similar scale; however, some of the arrays for the *M. tuberculosis* acid-shock data showed a wider distribution. To set all arrays to a common scale, a median absolute deviation scale transformation was therefore applied. As a final normalization step, duplicate spots for each gene on every microarray were averaged, and then the means of every gene across all technical replicate microarrays were calculated.

Means of the three biological replicates were used to compare gene expression between time points. Genes whose mean normalized expression value was larger than threefold different in one or more samples collected after the addition of acid compared to the zero time point sample were selected. A *t* test with a significance of 0.05 and with the Benjamini Hochberg correction was applied to verify statistical differential expression; those genes verifying this latter requirement were selected for further study. Files containing normalized data were imported into GeneSpring 7.1 (Silicon Genetics) for a graphical representation of the microarray data. Fully annotated microarray data has been deposited in BμG@Sbase (http://bugs.sgul.ac.uk/; accession no. E-BUGS-37) and ArrayExpress (www.ebi.ac.uk/arrayexpress/; accession no. E-BUGS-37).

Quantitative real-time PCR (qRT-PCR). In a total volume of 10 μl, RNA (0.5–1.0 μg) was mixed with 3 μg random hexamers, heated to 70 °C for 10 min and then snap-cooled on ice. To the chilled mixture, 4 μl 5× first strand buffer, 2 μl 100 mM DTT, and 20 nmol each of dATP, dCTP, dGTP and dTTP were added. The labelling reaction was initiated by the addition of 400 units Superscript II, and the reactions were incubated for 10 min at 25 °C and then for 90 min at 42 °C. The synthesized cDNA was purified with a DNA Clean and Concentrator kit (Zymo Research) and serial dilutions of the purified cDNA were made. qRT-PCR reactions were performed using the QuantiTect SYBR Green PCR kit (Qiagen) in a total volume of 25 μl containing 5 μl diluted cDNA, 12.5 μl 2× QuantiTect Master Mix and 0.5 μl gene-specific primers. PCR amplification was performed using a RotorGene 3000 (Corbett Research) and a programme consisting of 1 cycle at 95 °C for 15 min, and then 40–50 cycles at 95 °C for 15 s (denaturation), 55 °C for 30 s (extension step) and 80 °C for 30 s (signal acquisition). Melting curve analysis and gel electrophoresis were performed to confirm the specificity of the qRT-PCR products. The sequences of the gene-specific primer pairs were designed to show minimal cross-hybridization to all other genes within the genome. Sequences of the primer pairs are listed in supplementary Table S1 (available with the online version of the journal). The relative expression of each of the genes was normalized to *sigA*.

RESULTS AND DISCUSSION

Establishing continuous cultures of *M. tuberculosis* and *M. bovis*

A key issue in setting up the culture system was the choice of carbon source. Previously in batch culture, glycerol has been used as the carbon source for *M. tuberculosis*, and pyruvate has been used for *M. bovis*; clearly, comparative transcriptome profiling required both strains to be grown on the same carbon source, so neither glycerol nor pyruvate were used in this study. Instead, we exploited the presence of the oleic acid ester, Tween 80, in mycobacterial liquid media where it acts as a detergent to ensure dispersed growth; however, it can also be used as a carbon source by mycobacteria (Dubos & Davis, 1946). Tween 80 (0.2 %, v/v, in water) was therefore used as sole carbon source. The composition of the medium was manipulated to ensure that Tween 80 was the growth-limiting nutrient, and this was confirmed by adding a pulse of Tween 80 (0.4 %, v/v) to the steady-state cultures which resulted in an increase in OD_{540} over 48 h from 1.05 to 1.50 for the *M. bovis* culture, and from 1.325 to 1.55 for the *M. tuberculosis* culture. Hence, as well as its action as a detergent to prevent clumping, Tween 80 can be used as a carbon source by both *M. tuberculosis* and *M. bovis* under continuous culture conditions.

Cultures of *M. tuberculosis* and *M. bovis* were maintained in steady-state at a dissolved oxygen tension of 10 % (equivalent to 50 % air saturation) at a dilution rate of 0.03 h^{-1}, giving a constant mean generation time of 23 h. The mean OD_{540} and viable counts for three *M. bovis* cultures (Mb1–3) under steady-state conditions were, respectively, as follows: Mb1, 0.95 and 3.4×10^8 c.f.u. ml^{-1}; Mb2, 0.8 and 3.5×10^8 c.f.u. ml^{-1}; Mb3, 1.16 and 1.8×10^8 c.f.u. ml^{-1}. Mean OD_{540} and viable counts for three *M. tuberculosis* cultures (Mtb1–3) under steady-state conditions were, respectively, as follows: Mtb1, 1.29 and 3.7×10^8 c.f.u. ml^{-1}; Mtb2, 1.23 and 3.9×10^8 c.f.u. ml^{-1}; Mtb3, 1.32 and 3.1×10^8 c.f.u. ml^{-1}.

Differential gene expression in *M. tuberculosis* H37Rv and *M. bovis* AF2122 during steady-state growth

Ninety-two genes were identified that showed a minimum of threefold differential expression between *M. bovis* and *M. tuberculosis* during steady-state growth in continuous culture. Forty-three *M. bovis* genes showed higher expression than their *M. tuberculosis* orthologues, while 49 genes showed higher expression in *M. tuberculosis* (Tables 1 and 2, respectively). To focus our discussion we highlight differences in the key categories of cell wall and secreted proteins, lipid metabolism, gene regulators, PE/PPE family, and toxin–antitoxin (TA) gene pairs, and examine the potential role of these variably expressed genes in phenotypic differences between the human and bovine bacilli.

Cell wall and secreted proteins

Two of the most highly expressed genes in *M. bovis* compared to their orthologues in *M. tuberculosis* were *mpb83* and *mpb70*, encoding the serodominant antigens MPB83 and MPB70 which have been shown to be the major protein differences between *M. tuberculosis* and *M.*

Table 1. Genes expressed at higher levels in *M. bovis* than *M. tuberculosis* under steady-state growth

Coding sequence		Gene name	-Fold change	Product
M. bovis 2122	*M. tuberculosis* H37Rv			
Mb0066	*Rv0065*	*vapC-1*	6.6	PIN domain protein
Mb0222	*Rv0216*		17.7	Acyl-CoA hydrolase
Mb0223c	*Rv0217c*	*lipW*	5.8	Esterase
Mb0454c	*Rv0446c*		4.1	Conserved hypothetical protein
Mb0455c	*Rv0447c*	*ufaA1*	3.6	Fatty acid synthase
Mb0456c	*Rv0448c*		5.2	Conserved hypothetical protein
Mb0457c	*Rv0449c*		13.7	Putative amine oxidase
Mb0474c	*Rv0465c*		3.8	Putative transcriptional regulator
Mb0563c	*Rv0549c*	*vapC-3′*	3.5	Truncated PIN domain protein
Mb0607	*Rv0592*	*mce2D*	4.8	Mce family protein
Mb0635	*Rv0619*	*galTb*	3.6	Galactose uridyltransferase
Mb0792	*Rv0769*		3.4	Probable short-chain dehydrogenase
Mb0804	*Rv0782*	*ptrBb*	7.4	Probable protease II
Mb1405	*Rv1371*		5.2	Conserved transmembrane protein
Mb1529	*Rv1492*	*mutA*	3.3	Probable methylmalonyl-CoA mutase
Mb1649c	*Rv1623c*	*cydA*	5.2	Probable cytochrome oxidase
Mb2003c	*Rv1981c*	*nrdF1*	7.0	Ribonucleotide reductase small subunit
Mb2050c	*Rv2025c*		4.4	Transmembrane protein
Mb2095	*Rv2069*	*sigC*	3.2	RNA polymerase sigma factor
Mb2651	*Rv2618*		4.1	Possible transcriptional regulator
Mb2652c	*Rv2619c*		8.6	Conserved hypothetical protein
Mb2653c	*Rv2620c*		15.2	Conserved transmembrane protein
Mb2654c	*Rv2621c*		13.6	Possible transcriptional regulator
Mb2655	*Rv2622*		12.5	Possible methyltransferase
Mb2898	*Rv2873*	*mpb83*	60.3	Cell surface lipoprotein
Mb2899	*Rv2874*	*dipZ*	34.3	CcdA-like membrane protein
Mb2900	*Rv2875*	*mpb70*	63.4	Secreted immunogenic lipoprotein
Mb2901	*Rv2876*		7.8	Transmembrane protein
Mb2902c	*Rv2877c*		21.4	Membrane protein
Mb2970c	*Rv2945c*	*lppX*	3.8	PDIM export protein
Mb2971c	*Rv2946c*	*pks1*	3.5	PDIM synthesis
Mb2982c	*Rv2958c*		3.9	PGL and *p*-hydroxybenzoic acid biosynthesis
Mb3012c	*Rv2988c*	*leuC*	3.5	Leucine synthase
Mb3109c	*Rv3082c*	*virS*	10.2	Transcriptional regulator
Mb3379c	*Rv3347c*	*PPE55*	4.0	PPE family protein
Mb3427c	*Rv3395c*		5.4	Conserved hypothetical protein
Mb3477c	*Rv3447c*		8.6	FtsK/SpoIIIE family
Mb3706c	*Rv3681c*	*whiB4*	10.7	Possible transcriptional regulator
Mb3829c	*Rv3862c*	*whiB6*	4.4	Possible transcriptional regulator
Mb3894	*Rv3864*		4.7	Conserved hypothetical protein
Mb3902	*Rv3872*	*PE35*	3.9	PE family protein
Mb3903	*Rv3873*	*PPE68*	3.8	PPE family protein
Mb3904	*Rv3874*	*esxB*	4.5	ESAT-6 antigen

bovis (Hewinson *et al.*, 1996). Using a combination of expression profiling and comparative genomics, Behr and colleagues have shown that the inactivation of a negative regulator is the basis for high expression in *M. bovis* (Charlet *et al.*, 2005; Said-Salim *et al.*, 2006). Hence, expression of *mpb83* and *mpb70* is controlled by an anti-sigma factor/sigma factor pair, RskA/SigK; a mutation in the RskA (Mb0452c) anti-sigma factor in *M. bovis* prevents repression of SigK activity, leading to constitutive *mpb70* and *mpb83*. Another animal-adapted member of the *M. tuberculosis* complex, the 'oryx' bacillus (Mostowy *et al.*, 2005; van Soolingen *et al.*, 1994), is also a high producer of MPB70 and MPB83, and has an independent mutation in RskA that abolishes its activity (Said-Salim *et al.*, 2006).

These independent mutations would suggest positive selection for overexpression of MPB70 and MPB83 in *M. bovis* and the oryx bacillus, the phenotypic advantage of which remains to be determined. The genes *Mb0455c*, *Mb0456c*, *Mb0457c*, *dipZ*, *Mb2901* and *Mb2902c* are also part of the SigK regulon, and show higher expression in *M. bovis* than in *M. tuberculosis*; however, as with MPB83 and MPB70, the precise functions of their encoded proteins are unknown.

Rv0987 and *Rv0988* encode part of an ABC transporter and a putative secreted hydrolase, respectively. Both genes showed higher expression in *M. tuberculosis* H37Rv (Table 2), and in *M. bovis* 2122 the *Rv0987* orthologue is frameshifted (*Mb1013/Mb1014*). Inactivation of *Rv0987* in *M. tuberculosis* reduces the ability of the mutant to bind to THP-1 macrophages, but does not reduce virulence in the mouse model (Rosas-Magallanes *et al.*, 2007). An *Rv0987* mutant was also found to be incapable of invasion across the blood–brain barrier (Jain *et al.*, 2006). Whether mutation and decreased expression of the *M. bovis* *Rv0987-0988* orthologues reflect inactivation of factors no longer needed for full virulence of the bovine bacillus remains to be verified.

Lipid metabolism

Sulfolipid is a trehalose-containing sulfatide which is absent from *M. bovis*, but present in *M. tuberculosis*. It is therefore not surprising that the *mmpL8-papA1-pks2* locus responsible for sulfolipid synthesis (Bhatt *et al.*, 2007; Converse *et al.*, 2003; Sirakova *et al.*, 2001) showed approximately fourfold higher expression in *M. tuberculosis* (Table 2). The cluster *pks4-Rv1184c-Rv1185c* involved in the synthesis of another group of non-mycolated trehalose esters, the diacyltrehalose (DAT) and polyacyltrehaloses (PATs) (Dubey *et al.*, 2002), also showed higher expression in *M. tuberculosis* (Table 2). Both of these gene clusters have been shown to be under control of the PhoPR regulon (Gonzalo Asensio *et al.*, 2006; Walters *et al.*, 2006). It is therefore interesting that *phoR* contains two non-synonymous single nucleotide polymorphisms in *M. bovis*, with glycine at position 71 of the *M. tuberculosis* PhoR replaced by isoleucine in *M. bovis* and proline-172 replaced by leucine. Hence, mutation of PhoR may lead to the divergent expression of sulfolipid, DAT and PAT biosynthetic genes in *M. bovis*.

M. bovis showed higher expression of the *lppX-pks1* pair, which are involved in the transport (Sulzenbacher *et al.*, 2006) and synthesis (Constant *et al.*, 2002) of phthiocerol dimycoserosate (PDIM), respectively (Table 1). TLC analyses did not reveal significant differences in PDIM production between *M. bovis* 2122 and *M. tuberculosis* H37Rv (data not shown), suggesting the expression difference may relate to the synthesis of the PDIM-derived mycoside B by *M. bovis*, a glycolipid which is lacking in *M. tuberculosis* H37Rv. *Rv2958c*, encoding a glycosyltransferase involved in the synthesis of the trisaccharide phenolic glycolipid (PGL) that is derived from PDIM, was also found to be expressed at higher levels in *M. bovis* (Table 1). PGL is not produced by *M. tuberculosis* H37Rv, so the higher expression of this gene in *M. bovis* may simply be down to low expression of the synthetic pathways in *M. tuberculosis*. However, it should also be noted that Rv2958c is non-functional in *M. bovis*, so the higher expression of its gene may simply be due to loss of negative feedback inhibition. As the regulation of PDIM and PGL synthesis is currently poorly understood, comparisons between *M. bovis* and *M. tuberculosis* offer an ideal starting point to tease the system apart.

Mb0222 showed 17-fold higher expression in *M. bovis* compared with the *M. tuberculosis* orthologue *Rv0216*, with the adjacent *lipW* also up-regulated (Table 1). The similarity of Mb0222 to acyl dehydratases, and its location close to genes encoding a fatty acyl-CoA ligase (*Mb0220c*), an acyl dehydrogenase (*Mb0221c*) and an esterase (*Mb0223c*), indicates a role in fatty acid catabolism. *Rv0216* has been shown to be required for virulence in *M. tuberculosis* (Sassetti & Rubin, 2003), while *Mycobacterium leprae* has retained a functional orthologue, *ML2625* (Cole *et al.*, 2001); the *M. bovis* orthologue may also, therefore, play a role in virulence. Up-regulation of *Mb0222* may indicate that the encoded product is primed for immediate activity on *M. bovis* entering the host.

Gene regulators

Transcriptional regulators that showed variable expression were *Mb2651*, *Mb2654c*, *Mb3109c*/*virS*, *Mb3477c* (expressed at higher levels in *M. bovis;* Table 1), and *Rv0196*, *Rv0275c* and *Rv2160A-Rv2160c* (higher in *M. tuberculosis*; Table 2). The latter gene pair, *Rv2160A-Rv2160c*, are frameshifted in *M. tuberculosis* H37Rv, with the intact orthologue in *M. bovis* (*Mb2184c*) showing similarity to TetR-family transcriptional repressors. As the flanking genes *Rv2159c* and *Rv2161c* also show higher expression in *M. tuberculosis* H37Rv, it suggests that loss of the repressor from *M. tuberculosis* leads to up-regulation of this locus.

Mb3706c/*whiB4* and *Mb3829c*/*whiB6* were expressed at higher levels in *M. bovis* 2122 (Table 1). The WhiB family were originally designated as transcription factors based on the presence of putative DNA-binding motifs (Davis & Chater, 1992), but direct proof for this activity is lacking. Recent work has shown that this family bind Fe–S clusters (Jakimowicz *et al.*, 2005), and indeed WhiB4 binds Fe–S and functions as a disulfide reductase (Alam *et al.*, 2007). Furthermore, comparing expression of all seven *whiB* genes in *M. tuberculosis* to a range of stress conditions, Bishai and colleagues found that *whiB6* expression was the most stress-responsive (Geiman *et al.*, 2006). The WhiB proteins of tubercle bacilli may therefore function in redox signalling during oxidative and nitrosative stress, conditions which prevail in the *in vivo* milieu. Up-regulation of *whiB4* and *whiB6* in *M. bovis* under *in vitro* conditions may therefore indicate an increased need to sense and respond quickly to oxidative stress.

Table 2. Genes expressed at higher levels in *M. tuberculosis* than *M. bovis* under steady-state growth

Coding sequence		Gene name	-Fold change	Product
M. tuberculosis H37Rv	*M. bovis* 2122			
Rv0120c	*Mb0124c*	*fusA2*	3.1	Translation elongation factor
Rv0196	*Mb0202*		3.5	Putative transcriptional regulator
Rv0197	*Mb0203*		5.5	Possible oxidoreductase
Rv0275c	*Mb0281c*		4.8	Putative transcriptional regulator
Rv0276	*Mb0282*		12.0	Conserved hypothetical protein
Rv0557	*Mb0572*	*pimB*	3.8	Mannosyltransferase
Rv0933	*Mb0958*	*pstB*	3.1	ABC-type phosphate transport system
Rv0934	*Mb0959*	*pstS1*	5.0	Phosphate binding lipoprotein
Rv0987	*Mb1013*; *Mb1014*		5.1	ABC-type transport system
Rv0988	*Mb1015*		3.8	Secreted hydrolase
Rv1161	*Mb1193*	*narG*	4.0	Nitrate reductase subunit
Rv1162	*Mb1194*	*narH*	10.8	Nitrate reductase subunit
Rv1181	*Mb1213*	*pks4*	4.4	DAT and PAT synthesis
Rv1184c	*Mb1216c*		7.8	DAT and PAT synthesis
Rv1185c	*Mb1217c*	*fadD21*	3.7	DAT and PAT synthesis
Rv1186c	*Mb1218c*		5.6	Conserved hypothetical protein
Rv1374c	*Mb1409c*		3.3	Conserved hypothetical protein
Rv1397c	*Mb1432c*		4.2	Conserved hypothetical protein
Rv1398c	*Mb1433c*		7.8	Conserved hypothetical protein
Rv1535	*Mb1562*		4.5	Conserved hypothetical protein
Rv1595	*Mb1621*	*nadB*	3.1	NAD synthase
Rv1596	*Mb1622*	*nadC*	4.7	NAD synthase
Rv1809	*Mb1838*; *Mb1839*	*PPE33*	3.3	PPE family protein
Rv1810	*Mb1840*		3.6	Conserved hypothetical protein
Rv1899c	*Mb1934c*	*lppD*	5.2	Lipoprotein
Rv1925	*Mb1960*	*fadD31*	4.1	Fatty acid synthase
Rv2077c	*Mb2102c*		10.2	Transmembrane protein
Rv2159c	*Mb2183c*		21.8	Alkylhydroperoxidase-like protein
Rv2160A	*Mb2184c*		8.4	Conserved hypothetical protein
Rv2160c	*Mb2184c*		12.7	Conserved hypothetical protein
Rv2161c	*Mb2185c*		20.2	Flavin-dependent oxidoreductase
Rv2331	*Mb2358*		6.9	Conserved hypothetical protein
Rv3094c	*Mb3121c*		3.4	Conserved hypothetical protein
Rv3136	*Mb3160*	*PPE51*	9.6	PPE family protein
Rv3137	*Mb3161*		3.6	Mg^{2+}-dependent phosphatase
Rv3312A	*Mb3341c*		5.9	Secreted antigen
Rv3407	*Mb3441*	*vapB-44*	17.1	PhD antitoxin
Rv3408	*Mb3442*	*vapC-44*	4.8	PIN domain
Rv3409c	*Mb3443c*	*choD*	3.8	Putative cholesterol oxidase precursor
Rv3420c	*Mb3454c*	*rimI*	3.1	Ribosomal protein acetyltransferase
Rv3479	*Mb3506*		3.5	Transmembrane protein
Rv3581c	*Mb3612c*	*ispF*	3.7	Erythritol cyclodiphosphate synthase
Rv3679	*Mb3704*		12.4	Putative anion transporter
Rv3686c	*Mb3711c*		3.4	Conserved hypothetical protein
Rv3750c	*Mb3376c*		6.6	Putative excisionase
Rv3823c	*Mb3853c*	*mmpL8*	4.1	Sulfolipid biosynthesis
Rv3824c	*Mb3854c*	*papA1*	4.8	Sulfolipid biosynthesis
Rv3825c	*Mb3855c*	*pks2*	4.1	Sulfolipid biosynthesis
Rv3848	*Mb3878*		4.5	Transmembrane protein

The level of expression of *virS* was 10-fold higher in *M. bovis* compared with *M. tuberculosis* (Table 1). In *M. tuberculosis*, VirS has been shown to regulate expression of the genes *Rv3083-Rv3089*, the products of which could be involved in the modification of fatty acids in the cell envelope (Singh *et al.*, 2003). However, no differential

expression of these latter genes between *M. tuberculosis* and *M. bovis* was observed.

Clearly, a full description of the regulatory networks controlled by the transcription factors named above is required if their role in the phenotypic discrimination of human and bovine bacilli is to be defined. The application of genome-wide location studies, such as chromatin-immunoprecipitation assays, will aid in the identification of these gene networks (Rodrigue *et al.*, 2007).

PE and PPE family

The PE and PPE proteins are two large protein families with >160 members in *M. tuberculosis* and *M. bovis*. Approximately 60 % of the PE and PPE proteins differ across the two strains, a feature clearly at odds with the high sequence identity across the majority of encoded proteins in human and bovine bacilli (Cole *et al.*, 1998; Garnier *et al.*, 2003). While the precise function of the PE and PPE proteins remains to be defined, members have been implicated in virulence and antigenic variation. As well as variation at the nucleotide level, it is therefore interesting that the genes encoding PPE33, PE35, PPE51 and PPE55 showed variation in expression between *M. tuberculosis* H37Rv and *M. bovis* AF2122 (Tables 1 and 2). The higher expression of the gene encoding PPE33, *Rv1809*, in *M. tuberculosis* may be due to the *M. bovis* orthologue being frameshifted (*Mb1838/Mb1839*), thus disrupting its correct regulation. *Rv3136*, which encodes PPE51, has been shown to be one of 22 genes that are consistently up-regulated in *M. tuberculosis* H37Rv compared to the attenuated derivative H37Ra (Mostowy *et al.*, 2004), a result that parallels its decreased expression in *M. bovis*. The differential expression of PPE genes between human and bovine bacilli points to subtle variation in the control of these genes among strains, the biological implications of which remain to be defined.

TAs

TA modules were originally described as plasmid maintenance systems that trigger post-segregational killing of plasmid-free cells (Gerdes *et al.*, 2005). The TA pair is transcriptionally coupled, with the antitoxin typically acting both to repress transcription of the locus and inhibiting the action of the toxin. In response to various stresses, toxin release can lead to growth arrest and antibiotic tolerance (Correia *et al.*, 2006). At least 60 TA genes have been discovered in the genome of *M. tuberculosis*, a very high number compared to most other sequenced bacteria (Arcus *et al.*, 2005; Pandey & Gerdes, 2005).

TA pairs showed variation in expression between *M. tuberculosis* and *M. bovis*. *Rv3407-Rv3408* were up-regulated in *M. tuberculosis* H37Rv (Table 2) and encode a TA pair (belonging to the *vapBC* family) designated *vapBC-44* by Pandey & Gerdes (2005); *Rv3407* encodes the antitoxin, and *Rv3408* the PilT amino terminus (PIN) domain toxin. *M. tuberculosis* also showed up-regulation of *Rv1397c-Rv1398c*, another TA pair designated *vapBC-20* (D. P. Pandey & K. Gerdes, unpublished). Comparative proteome analysis of BCG and *M. tuberculosis* showed that Rv3407 is present at lower levels in BCG than *M. tuberculosis* (Jungblut *et al.*, 1999). Inactivation of four of the five *M. tuberculosis rpf* genes, encoding glycolytic hydrolases implicated in resuscitation of dormant bacilli, led to up-regulation of *Rv3407/3408*; however, inactivation of the one *rpf* gene which is different between *M. bovis* and *M. tuberculosis*, *rpfA*, had no effect on *Rv3407/3408* expression (Downing *et al.*, 2004). Interestingly, an RD1 knockout of *M. tuberculosis* showed a 10-fold up-regulation of *Rv3407* and *Rv3408* (Mostowy *et al.*, 2004). However, BCG is deleted for RD1, yet shows low expression of its *Rv3407/Rv3408* orthologues (Jungblut *et al.*, 1999), again pointing to defective regulation of the *vapBC-44* locus in *M. bovis* strains.

In *M. bovis*, *Mb0066*, encoding a PIN domain protein, was expressed at higher levels than in *M. tuberculosis*; however, no upstream antitoxin had been annotated (Garnier *et al.*, 2003). Using exhaustive searches for TA loci, Pandey & Gerdes (2005) identified a putative CDS upstream of *Mb0066/Rv0065* that they designated *vapB-1* and that encodes the antitoxin for *Mb0066/Rv0065* (*vapC-1*). In keeping with *M. bovis* genome nomenclature, we designated this upstream CDS *Mb0066A*. Using qRT-PCR analyses we confirmed both the transcription of the *Mb0066A*/*vapB-1* CDS in *M. bovis* and *M. tuberculosis* and the increased expression of *Mb0066A* in *M. bovis* compared to *M. tuberculosis* (Fig. 1).

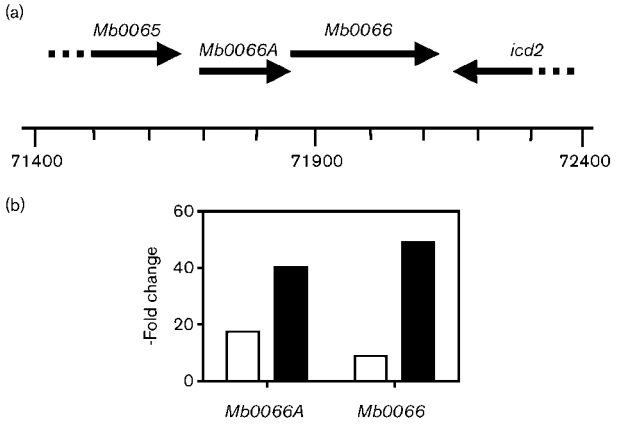

Fig. 1. Up-regulation of a TA pair in *M. bovis* compared with *M. tuberculosis*. (a) Map of the chromosomal location of *Mb0066A*/*Mb0066* in *M. bovis*. (b) -Fold increase in expression of *Mb0066A*/*Mb0066* in *M. bovis* (black bars) compared with *M. tuberculosis* (white bars) during steady-state growth as measured by qRT-PCR. Data are the means of two independent experiments.

The biological importance of TA pair differential expression is unclear. However, evidence from *E. coli* has shown that some TAs play roles in the generation of drug-resistant 'persister' populations (Correia *et al.*, 2006; Keren *et al.*, 2004). Given the increased emergence of drug-resistant populations of *M. tuberculosis*, exploration of the function of TA proteins is a high priority, and further studies of the pairs that show differential expression between *M. tuberculosis* and *M. bovis* may offer an ideal starting point to elucidate function.

Verification by qRT-PCR

To confirm gene expression differences predicted by array analysis, real-time qRT-PCR was performed. All of the 92 genes showing differential regulation between *M. tuberculosis* and *M. bovis* could not be analysed due to cost implications, so a subset of nine was selected that included genes involved in transcriptional regulation (*Mb2654c*, *virS*, *whiB4*), lipid metabolism (*Rv1184c*, *Mb0222*), and a TA pair (*Rv3407*). Fig. 2 compares the -fold changes in expression measured by microarray and qRT-PCR for these nine genes. The results showed that both methods display broad agreement in their measurements of expression for each gene, confirming that experimental procedures and data handling had not introduced any significant bias to the dataset.

Acid-shock response of *M. tuberculosis* and *M. bovis*

Our chemostat experiments defined a core gene set that showed differential expression between *M. tuberculosis* and *M. bovis* during steady-state growth. However, we also wished to determine if the response to a perturbation of steady-state was significantly different between the strains. We chose acid shock as this perturbation, a biologically relevant stress experienced by mycobacteria during infection. Cultures were grown to steady-state, then the pH dropped to 5.5 and gene expression was monitored over a 24 h period at 5, 15, 30 and 60 min and 24 h after acid shock. We identified 150 *M. tuberculosis* genes and 66 *M. bovis* genes that showed a minimum of a threefold change in expression post-acid-shock compared with the steady-state, zero time point (see supplementary Tables S2 and S3, available with the online version of this paper, for a full listing of all *M. tuberculosis* and *M. bovis* genes that were responsive to acid shock).

While *M. tuberculosis* showed repression of 44 genes in response to acid shock across all time points (Table S2), *M. bovis* repressed only six genes (Table S3), with just *Rv3269/Mb3298*, encoding a protein of unknown function, common to both species. In terms of gene induction, 106 *M. tuberculosis* genes showed increased expression, compared to 60 *M. bovis* genes, with 15 genes in common (Tables S2 and S3). Temporal changes in gene expression were evident in *M. tuberculosis* and *M. bovis*, with some

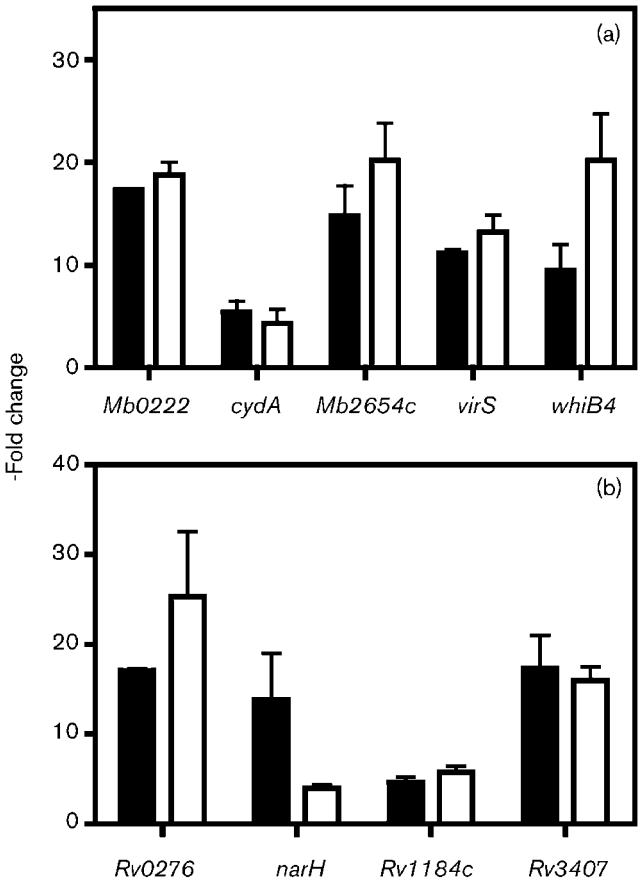

Fig. 2. Confirmation of microarray results using qRT-PCR. Selected genes that show higher levels of expression in *M. bovis* compared with *M. tuberculosis* (a), or higher levels of expression in *M. tuberculosis* compared with *M. bovis* (b) in steady-state samples, as measured by microarray (black bars) and qRT-PCR (white bars). Data are the means of two independent experiments.

genes showing 'early' induction (5–60 min), while others were only induced at the 'late' (24 h) time point. At the global level it was apparent that *M. tuberculosis* and *M. bovis* showed significant differences in their response to acid stress. Below we highlight the commonality and differences in this temporal response.

Common acid-shock response

Two genes encoding putative methyltransferases, *Rv1403c/Mb1438c* and *Rv1405c/Mb1440c*, were induced early after acid shock in both *M. tuberculosis* and *M. bovis* (Table 3). *Rv1404/Mb1439* lies between the genes and encodes a MarR family transcriptional repressor that regulates expression of *Rv1405c* and *Rv1403c* (P. Golby, unpublished). Previous work has shown that *Rv1405c* was strongly up-regulated in macrophages (Schnappinger *et al.*, 2003), while transposon site hybridization studies identified *Rv1405c* as one of 194 genes predicted to be essential for survival of *M. tuberculosis in vivo* (Sassetti & Rubin, 2003). The target(s)

Table 3. *M. tuberculosis* and *M. bovis* genes up-regulated in response to acid-shock

Coding sequence	Gene name	-Fold change at:					Putative function
		5 min	15 min	30 min	60 min	1440 min	
Rv0263c		0.57	0.81	1.56	2.75	3.76	Putative allophanate hydrolase subunit 1
Mb0269c		0.9	0.75	2.01	4.51	1.64	
Rv1057		5.96	4.38	2.56	2.64	1.46	Conserved hypothetical protein
Mb1086		2.87	4.98	1.69	1.69	2.59	
Rv1131	*gltA1*	0.58	2.34	7.11	17.13	22.14	Citrate synthase
Mb1162		0.82	1.23	3.68	8.37	2.33	
Rv1195	*PE13*	2.86	4.93	3.25	2.47	0.4	PE family protein
Mb1227		1.76	3.94	1.1	0.87	0.89	
Rv1403c		2.46	8.71	8.86	13.09	13.0	Putative methyltransferase
Mb1438c		0.91	6.22	4.99	5.45	5.44	
Rv1405c		5.06	25.95	30.79	30.37	8.56	Putative methyltransferase
Mb1440c		1.3	28.66	21.2	16.83	11.75	
Rv1463		1.07	3.30	1.93	0.71	0.87	Probable Fe–S cluster assembly protein
Mb1498		1.01	3.39	1.16	0.87	0.99	
Rv1464	*csd*	1.02	4.41	2.11	0.94	1.2	Probable cysteine desulfurase
Mb1499		1.12	3.21	1.1	0.82	1.17	
Rv1553	*frdB*	0.82	0.95	0.86	1.45	3.85	Fumarate reductase subunit B
Mb1579		1.22	1.17	1.09	1.95	3.09	
Rv2390c		2.9	6.7	8.0	10.7	7.5	Conserved hypothetical protein
Mb2411c		1.0	2.2	2.6	3.0	1.8	
Rv2428	*ahpC*	10.17	6.53	4.35	2.88	0.63	Alkyl hydroperoxide reductase
Mb2454		5.49	2.58	1.3	1.48	0.73	
Rv2557		1.13	0.87	1.15	1.68	12.62	Conserved hypothetical protein
Mb2587		0.99	0.85	0.92	1.56	5.54	
Rv2930	*fadD26*	4.94	5.45	2.95	1.48	0.67	Fatty acyl-CoA ligase
Mb2955		2.86	3.72	1.7	1.46	1.07	
Rv2931	*ppsA*	1.07	5.04	3.08	1.35	0.74	PDIM synthase
Mb2956		0.98	3.69	1.6	0.93	0.94	
Rv2959c		3.63	7.66	4.0	2.36	0.8	PGL and *p*-hydroxybenzoic acid biosynthesis
Mb2983c		1.7	3.29	1.78	1.63	0.84	
Rv3601c	*panD*	2.1	5.6	4.92	3.07	1.01	Aspartate decarboxylase
Mb3631c		1.37	4.35	4.71	4.49	1.56	
Rv3603c		3.23	4.08	3.84	2.15	1.29	Hypothetical alanine and leucine rich protein
Mb3633c		2.09	3.09	2.94	2.18	1.52	

of Rv1405c and Rv1403c are unknown. Methyltransferase-mediated modification of unsaturated fatty acids to produce cyclopropane fatty acids is a known acid resistance mechanism in *E. coli* (Grogan & Cronan, 1997). The gene encoding the cyclopropane fatty-acyl synthase, *cfa*, is also up-regulated in response to low pH in *E. coli* (Maurer *et al.*, 2005). However, Rv1403c and Rv1405c show low similarity to cyclopropane fatty-acyl synthases, suggesting this is not their function. Identification of the methylated target will help to shed light on the function of Rv1403c/Mb1438c and Rv1405c/Mb1440c both in virulence and in response to stress.

Genes involved in PDIM synthesis were up-regulated in both *M. tuberculosis* and *M. bovis*, with *fadD26*, *ppsA* and *Rv2959c* showing maximal induction after 5–10 min, with expression then decreasing over the 24 h period (Table 3). *ppsB* and *ppsE* also showed up-regulation over the same time frame in *M. tuberculosis* (Table 4). It has previously been shown that mutants lacking PDIM have increased permeability (Camacho *et al.*, 2001), suggesting that the early up-regulation of PDIM biosynthesis may be a general stress response that attempts to remodel the cell wall, decreasing permeability to damaging agents. This is similar to the suggested role for increased levels of cyclopropane fatty acids in membrane lipids of acid-tolerant *E. coli*, with conversion of unsaturated to cyclopropane fatty acids leading to a decrease in proton permeability (Grogan & Cronan, 1997).

Rv0263c/Mb0269c was up-regulated 30–60 min after acid-shock (Table 3) and its predicted protein sequence shows similarity to allophanate hydrolase subunit 2 which hydrolyses urea carboxylate to carbon dioxide and ammonia. Up-regulation would lead to increased ammonia production, hence acting to buffer the cytoplasmic pH

Table 4. *M. tuberculosis* and *M. bovis* genes that showed differential expression between the strains after acid shock

Coding sequence	Gene name	-Fold change at:					Predicted function
		5 min	15 min	30 min	60 min	1440 min	
***M. tuberculosis* H37Rv**							
Rv1028c	*kdpD*	1.13	0.92	1.00	1.41	4.54	Sensor kinase
Rv1305	*atpE*	0.80	0.75	0.76	0.57	0.23	ATP synthase
Rv1306	*atpF*	0.93	0.87	0.79	0.60	0.20	ATP synthase
Rv1307	*atpH*	0.67	1.04	0.92	0.58	0.21	ATP synthase
Rv1308	*atoA*	1.04	0.79	0.81	0.66	0.21	ATP synthase
Rv1311	*atpC*	0.88	0.82	0.77	0.60	0.22	ATP synthase
Rv1611	*trpC*	0.97	0.94	0.91	0.75	0.18	Tryptophan synthesis
Rv1612	*trpB*	1.31	1.13	0.99	0.63	0.20	Tryptophan synthesis
Rv1613	*trpA*	0.78	0.83	0.73	0.62	0.24	Tryptophan synthesis
Rv2244	*acpM*	0.76	0.77	0.26	0.17	0.25	Mycolic acid synthesis
Rv2245	*kasA*	1.10	0.84	0.31	0.21	0.35	Mycolic acid synthesis
Rv2246	*kasB*	0.91	0.69	0.29	0.24	0.38	Mycolic acid synthesis
Rv2321c	*rocD2*	0.88	0.84	1.05	0.84	4.24	Ornithine aminotransferase
Rv2322c	*rocD1*	0.81	1.08	1.12	1.56	9.85	Ornithine aminotransferase
Rv2323c		1.06	1.21	1.22	1.69	16.26	Putative amidinotransferase
Rv2846c	*efpA*	1.08	0.52	0.33	0.24	0.58	Putative efflux protein
Rv2930	*fadD26*	4.94	5.03	2.95	1.49	0.67	PDIM biosynthesis
Rv2931	*ppsA*	1.07	5.05	3.08	1.35	0.74	PDIM biosynthesis
Rv2932	*ppsB*	1.35	4.42	3.86	1.97	1.02	PDIM biosynthesis
Rv2935	*ppsE*	1.63	1.94	4.45	2.34	0.83	PDIM biosynthesis
Rv2959c		3.63	7.66	4.00	2.36	0.80	PDIM biosynthesis
Rv3146	*nuoB*	0.38	0.53	0.55	0.72	0.28	Type I NADH dehydrogenase subunit
Rv3148	*nuoD*	0.49	0.48	0.54	0.72	0.30	Type I NADH dehydrogenase subunit
Rv3149	*nuoE*	0.60	0.35	0.43	0.49	0.18	Type I NADH dehydrogenase subunit
Rv3150	*nuoF*	0.97	0.43	0.54	0.60	0.22	Type I NADH dehydrogenase subunit
Rv3153	*nuoI*	1.08	0.39	0.50	0.57	0.24	Type I NADH dehydrogenase subunit
Rv3157	*nuoM*	1.04	0.75	0.59	0.64	0.25	Type I NADH dehydrogenase subunit
Rv3158	*nuoN*	0.81	0.54	0.44	0.42	0.24	Type I NADH dehydrogenase subunit
***M. bovis* 2122**							
Rv1029	*kdpA*	1.18	1.36	1.54	1.56	19.07	High affinity K^+ transporter
Rv1030	*kdpB*	1.04	0.94	1.08	0.92	7.19	High affinity K^+ transporter
Rv1031	*kdpC*	1.34	1.31	1.59	1.55	6.66	High affinity K^+ transporter
Rv1181	*pks4*	1.81	1.95	3.34	5.30	9.03	DAT and PAT synthesis
Rv1182	*papA3*	1.45	1.81	2.51	5.14	8.69	DAT and PAT synthesis
Rv1183	*mmpL10*	2.18	1.02	2.05	3.75	5.74	DAT and PAT synthesis
Rv2874	*dipZ*	1.09	0.91	1.03	1.49	3.07	DsbD homologue

during acid stress. Tubercle bacilli also contain the 'classic' bacterial nickel-containing UreABC enzyme, but the genes encoding this complex were not up-regulated. While pathogens such *Helicobacter pylori* are known to produce high levels of urease as a means of protection against host acidic environments (Weeks *et al.*, 2000), this is the first report of up-regulation of urease genes in tubercle bacilli in response to acid stress.

Differential acid-shock response

Among the 44 genes repressed in *M. tuberculosis* in response to acid-shock were two clusters of genes involved in energy transduction (Table 4). Five *atp* genes encoding components of the membrane-bound proton-translocating ATP synthase were down-regulated in *M. tuberculosis*. As this system serves to import protons, and hence acidify the cytoplasm, its repression under acid stress would aid in maintaining cytoplasmic pH near neutrality. However, seven *nuo* genes encoding subunits of the type I NADH dehydrogenase, which serves to export protons, were also down-regulated. As proton export would help to maintain cytoplasmic pH under acid-stress, the repression of the *nuo* operon is surprising. A previous analysis of the *M. tuberculosis* acid-shock response also noted down-regulation of genes encoding type I NADH dehydrogenase subunits *nuoA* and *nuoG* (Fisher *et al.*, 2002). The repression of both proton influx and efflux systems in *M. tuberculosis*, and their differential expression between *M. tuberculosis* and *M. bovis*, suggests a complex dynamic in the control of these systems in response to acid stress. This does not of course mean that constitutive pH homeostasis

systems that direct proton efflux, such as cytochrome respiratory chains (Booth, 1985), are not important.

The *kdpABC* operon encodes a high affinity K^+ uptake system that is regulated by turgor pressure in *E. coli* (Malli & Epstein, 1998). In *M. bovis* 2122, this operon was found to be induced at the 24 h time point, while in *M. tuberculosis* H37Rv only *kdpD*, encoding the sensor kinase that controls *kdpABC* expression, was up-regulated after 24 h (Table 4). Under conditions of cytoplasm acidification, protons need to be removed from the cytoplasm; however, to maintain electrochemical neutrality, positively charged ions must be introduced. Hence, induction of K^+ uptake systems allows the cell to maintain electrochemical gradients (Booth, 1985). As proton extrusion would need to occur quickly after acidification of the cytoplasm begins, it is therefore puzzling that *kdpABC* was induced only at the late time point in *M. bovis*. For example, for osmotic stress in *Salmonella*, it has been shown that up-regulation of *kdpABC* occurs less than 10 min after exposure to the stress (Balaji *et al.*, 2005). This suggests that the up-regulation of *kdpABC* in *M. bovis* seen here is either not a direct response to acid stress, or that the temporal response is markedly different in *M. bovis* compared to other bacteria. Monitoring the expression of *kdpABC* in *M. bovis* and *M. tuberculosis* in response to osmotic stresses will help to unravel the role of K^+ import for homeostasis in the tubercle bacilli.

Differential expression of amino acid biosynthesis genes in response to acid shock was seen. *M. tuberculosis* repressed the tryptophan biosynthetic genes *trpCBA* 30 min after the pH drop, but induced the expression of the *rocD2-rocD1-Rv2323c* locus over the same time period, reaching maximal expression by 24 h (Table 4). This latter frameshifted locus encodes an ornithine aminotransferase (RocD), and a putative amidinotransferase (Rv2323c), enzymes that are involved in the urea cycle and hence lead to ammonia production. As with up-regulation of ureases, increased production of ammonia will serve to increase cytoplasmic pH in the face of proton efflux.

While the *pks4-papA3-mmpL10* locus involved in DAT and PAT synthesis was expressed at lower levels in *M. bovis* than *M. tuberculosis* under steady-state conditions (Table 1; see above), this region was up-regulated in *M. bovis* after acid shock, reaching maximum expression after 24 h (Table 4). In *M. tuberculosis*, *acpM-kasA-kasB*, involved in mycolic acid biosynthesis, were down-regulated across all time points, as was *efpA*; repression of these genes has been shown to be a signature for agents that effect fatty acid synthesis (Boshoff *et al.*, 2004). As with up-regulation of PDIM synthetic loci that was seen in both strains in response to acid shock, altered expression of genes involved in cell wall synthesis suggests remodelling of cell wall architecture in response to stress. Indeed, the hydrophobic cell wall acts as a major protective barrier for the bacillus, so maintaining its integrity is clearly of major importance for the cell.

The *dipZ* gene, one of the SigK regulon genes whose expression is higher in *M. bovis* than *M. tuberculosis*, was up-regulated threefold after acid shock only in *M. bovis* at the 24 h time point. As this gene is already highly expressed in *M. bovis* (Table 1), its up-regulation is intriguing. None of the other SigK regulon genes was expressed at higher levels after acid shock, indicating SigK-independent regulation. DipZ is a homologue of DsbD in *E. coli*, which functions to transfer reducing power from the cytoplasm to the periplasm and plays key roles in the cytochrome *c* assembly pathway and correct disulfide-bond formation in proteins (Kadokura *et al.*, 2003). The crystal structure of the C-terminal domain of DipZ revealed a dimeric structure with an expected thioredoxin-like fold, but also an unexpected fold with similarity to carbohydrate-binding modules, suggesting a role in carbohydrate processing (Goldstone *et al.*, 2005). As cytochrome respiratory chains are key mediators of pH homeostasis through proton efflux (Booth, 1985), up-regulation of *dipZ* may reflect an increased requirement for cytochrome biogenesis. However, the unique structural features of DipZ suggest it could play a novel role that remains to be elucidated.

Conclusions

The *M. tuberculosis* complex can be viewed as a set of ecotypes with distinct host preferences (Smith *et al.*, 2006), the basis of which is encoded in the pathogen's genomes. Comparative analyses of the *M. tuberculosis* and *M. bovis* genomes has identified a set of genetic differences between these strains, some of which have been linked to known *in vitro* phenotypic differences (Keating *et al.*, 2005; Said-Salim *et al.*, 2006; Stermann *et al.*, 2004). However, comparative genomics cannot provide a full picture of the impact of genetic change on phenotype. To accelerate our identification of candidate genes that contribute to the phenotypes of *M. tuberculosis* and *M. bovis*, we therefore sought to identify key gene expression differences between the strains.

In defining core gene expression differences it is essential that growth conditions are identical between cultures. We chose continuous culture since it allows environmental conditions and bacterial growth rates to be tightly and reproducibly controlled, reducing variation between biological replicates. A seminal example of the need to control growth rate in transcriptome experiments has been provided by a global analysis of central carbon metabolism in *E. coli* using gene expression, protein and metabolome data (Ishii *et al.*, 2007). Using glucose-limited chemostat cultures of *E. coli*, it was found that the effect of variation in growth rate on the transcriptome and proteome was larger than the effects of most gene disruptions in carbon metabolism. We are therefore confident that our profile of steady-state gene expression in *M. tuberculosis* and *M. bovis* identifies key gene expression differences between the strains, rather than being simple growth rate or media effects.

In a recent publication, Rehren *et al.* (2007) described the transcriptome of *M. tuberculosis* H37Rv and *M. bovis* Ravanel grown in batch culture, identifying 278 genes with differential expression. Comparing their results with our data, 45 genes appear in both datasets, with 325 gene differences. Whether these differences are due to variation between *M. bovis* Ravanel and *M. bovis* 2122, or are down to technical differences between the experiments, is difficult to assess. For example, while we grew both *M. tuberculosis* H37Rv and *M. bovis* 2122 on Tween 80 as sole carbon source, Rehren and colleagues grew H37Rv in Middlebrook 7H9 medium with glycerol, while the Ravanel strain was grown in 7H9 medium with pyruvate; hence, some of their gene expression differences may be due to different carbon source utilization. The two studies also used different microarrays and data handling procedures, a further source of variation. Hence, the 'minimal' set of differentially expressed genes between *M. tuberculosis* and *M. bovis* remains to be defined.

Prior to our work, Fisher *et al.* (2002) studied the effect of acid on the transcriptome of *M. tuberculosis* H37Rv, but in batch culture. Overlap between the genes differentially expressed in response to acid stress identified here and in their work is apparent, with *Rv1130*, *gltA1*, *PE13*, *acpM*, *kasAB*, *ahpCD* and *efpA* identified in both analyses. However, some key differences in the data were also apparent, such as the induction of a putative polyketide synthase operon (*Rv3083-Rv3089*) in their work which we did not see in *M. tuberculosis* or *M. bovis*. Conversely, the strong induction of the putative methyltransferase-encoding genes *Rv1405c* and *Rv1403c* seen by us in both *M. tuberculosis* and *M. bovis* was not reported by Fisher and colleagues, although they did see a slight induction of *Rv1404*, which encodes the putative regulator of *Rv1405c* and *Rv1403c*. It is difficult to determine whether the variation in expression profile between the two studies is due to choice of strain, growth condition, experimental procedures or data handling.

In summary, we have defined key gene expression differences between the sequenced *M. tuberculosis* H37Rv and *M. bovis* 2122 strains, and have identified candidate genes that may play pivotal roles in the pathogenic process. The focus now is to explore the role of this differential gene expression in the distinct host preference of the human and bovine tubercle bacilli.

ACKNOWLEDGEMENTS

This work was funded by the UK Department for Environment, Food and Rural Affairs (Defra), the Department of Health, and the Health Protection Agency. The authors would like to thank Paul Wheeler and Lisa Keating (VLA) for their help in the development of the chemostat growth medium for *M. bovis*, and for useful discussions on the microarray data. We would like to acknowledge Jason Sawyer (VLA) for developing a protocol for qRT-PCR, to thank Lucy Brooks and Adam Whitney (Bacterial Microarray Group, St Georges) for help with depositing data in BμG@Sbase and ArrayExpress, and acknowledge Colorado State University for the provision of *M. tuberculosis* H37Rv genomic DNA produced under NIH contract HHSN266200400091C/ADB NO1-A1-40091 'Tuberculosis Vaccine Testing and Research Materials Contract'. The views expressed in the publication are those of the authors and not necessarily those of the Department of Health and Health Protection Agency.

REFERENCES

Alam, M. S., Garg, S. K. & Agrawal, P. (2007). Molecular function of WhiB4/Rv3681c of *Mycobacterium tuberculosis* H37Rv: a [4Fe–4S] cluster co-ordinating protein disulphide reductase. *Mol Microbiol* **63**, 1414–1431.

Arcus, V. L., Rainey, P. B. & Turner, S. J. (2005). The PIN-domain toxin-antitoxin array in mycobacteria. *Trends Microbiol* **13**, 360–365.

Bacon, J., James, B. W., Wernisch, L., Williams, A., Morley, K. A., Hatch, G. J., Mangan, J. A., Hinds, J., Stoker, N. G. & other authors (2004). The influence of reduced oxygen availability on pathogenicity and gene expression in *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* **84**, 205–217.

Balaji, B., O'Connor, K., Lucas, J. R., Anderson, J. M. & Csonka, L. N. (2005). Timing of induction of osmotically controlled genes in *Salmonella enterica* serovar *Typhimurium*, determined with quantitative real-time reverse transcription-PCR. *Appl Environ Microbiol* **71**, 8273–8283.

Bhatt, K., Gurcha, S. S., Bhatt, A., Besra, G. S. & Jacobs, W. R., Jr (2007). Two polyketide-synthase-associated acyltransferases are required for sulfolipid biosynthesis in *Mycobacterium tuberculosis*. *Microbiology* **153**, 513–520.

Booth, I. R. (1985). Regulation of cytoplasmic pH in bacteria. *Microbiol Rev* **49**, 359–378.

Boshoff, H. I., Myers, T. G., Copp, B. R., McNeil, M. R., Wilson, M. A. & Barry, C. E., III (2004). The transcriptional responses of *Mycobacterium tuberculosis* to inhibitors of metabolism: novel insights into drug mechanisms of action. *J Biol Chem* **279**, 40174–40184.

Camacho, L. R., Constant, P., Raynaud, C., Laneelle, M. A., Triccas, J. A., Gicquel, B., Daffe, M. & Guilhot, C. (2001). Analysis of the phthiocerol dimycocerosate locus of *Mycobacterium tuberculosis*. Evidence that this lipid is involved in the cell wall permeability barrier. *J Biol Chem* **276**, 19845–19854.

Chan, J., Silver, R. F., Kampmann, B. & Wallis, R. S. (2005). Intracellular models of *Mycobacterium tuberculosis* infection. In *Tuberculosis and the Tubercle Bacillus*, pp. 437–461. Edited by S. T. Cole, K. D. Eisenach, D. N. McMurray & W. R. Jacobs Jr. Washington, DC: American Society for Microbiology.

Charlet, D., Mostowy, S., Alexander, D., Sit, L., Wiker, H. G. & Behr, M. A. (2005). Reduced expression of antigenic proteins MPB70 and MPB83 in *Mycobacterium bovis* BCG strains due to a start codon mutation in *sigK*. *Mol Microbiol* **56**, 1302–1313.

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S. & other authors (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544.

Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honoré, N., Garnier, T., Churcher, C. & other authors (2001). Massive gene decay in the leprosy bacillus. *Nature* **409**, 1007–1011.

Constant, P., Perez, E., Malaga, W., Laneelle, M. A., Saurel, O., Daffe, M. & Guilhot, C. (2002). Role of the *pks15/1* gene in the biosynthesis of phenolglycolipids in the *M. tuberculosis* complex: evidence that all strains synthesize glycosylated P^{-}-hydroxybenzoic methyl esters and that strains devoid of phenolglycolipids harbour a frameshift mutation in the *pks15/1* gene. *J Biol Chem* **277**, 38148–38158.

Converse, S. E., Mougous, J. D., Leavell, M. D., Leary, J. A., Bertozzi, C. R. & Cox, J. S. (2003). MmpL8 is required for sulfolipid-1 biosynthesis and *Mycobacterium tuberculosis* virulence. *Proc Natl Acad Sci U S A* **100**, 6121–6126.

Correia, F. F., D'Onofrio, A., Rejtar, T., Li, L., Karger, B. L., Makarova, K., Koonin, E. V. & Lewis, K. (2006). Kinase activity of overexpressed HipA is required for growth arrest and multidrug tolerance in *Escherichia coli*. *J Bacteriol* **188**, 8360–8367.

Davis, N. K. & Chater, K. F. (1992). The *Streptomyces coelicolor whiB* gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. *Mol Gen Genet* **232**, 351–358.

Downing, K. J., Betts, J. C., Young, D. I., McAdam, R. A., Kelly, F., Young, M. & Mizrahi, V. (2004). Global expression profiling of strains harbouring null mutations reveals that the five *rpf*-like genes of *Mycobacterium tuberculosis* show functional redundancy. *Tuberculosis (Edinb)* **84**, 167–179.

Dubey, V. S., Sirakova, T. D. & Kolattukudy, P. E. (2002). Disruption of *msl3* abolishes the synthesis of mycolipanoic and mycolipenic acids required for polyacyltrehalose synthesis in *Mycobacterium tuberculosis* H37Rv and causes cell aggregation. *Mol Microbiol* **45**, 1451–1459.

Dubos, R. J. & Davis, B. D. (1946). Factors affecting the growth of tubercle bacilli in liquid media. *J Exp Med* **83**, 409–423.

Fisher, M. A., Plikaytis, B. B. & Shinnick, T. M. (2002). Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes. *J Bacteriol* **184**, 4025–4032.

Garnier, T., Eiglmeier, K., Camus, J. C., Medina, N., Mansoor, H., Pryor, M., Duthoy, S., Grondin, S., Lacroix, C. & other authors (2003). The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A* **100**, 7877–7882.

Geiman, D. E., Raghunand, T. R., Agarwal, N. & Bishai, W. R. (2006). Differential gene expression in response to exposure to antimycobacterial agents and other stress conditions among seven *Mycobacterium tuberculosis whiB*-like genes. *Antimicrob Agents Chemother* **50**, 2836–2841.

Gerdes, K., Christensen, S. K. & Lobner-Olesen, A. (2005). Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol* **3**, 371–382.

Goldstone, D., Baker, E. N. & Metcalf, P. (2005). Crystallization and preliminary diffraction studies of the C-terminal domain of the DipZ homologue from *Mycobacterium tuberculosis*. *Acta Crystallograph Sect F Struct Biol Cryst Commun* **61**, 243–245.

Gonzalo Asensio, J., Maia, C., Ferrer, N. L., Barilone, N., Laval, F., Soto, C. Y., Winter, N., Daffé, M., Gicquel, B. & other authors (2006). The virulence-associated two-component PhoP-PhoR system controls the biosynthesis of polyketide-derived lipids in *Mycobacterium tuberculosis*. *J Biol Chem* **281**, 1313–1316.

Grogan, D. W. & Cronan, J. E., Jr (1997). Cyclopropane ring formation in membrane lipids of bacteria. *Microbiol Mol Biol Rev* **61**, 429–441.

Hewinson, R. G., Michell, S. L., Russell, W. P., McAdam, R. A. & Jacobs, W. R., Jr (1996). Molecular characterization of MPT83: a seroreactive antigen of *Mycobacterium tuberculosis* with homology to MPT70. *Scand J Immunol* **43**, 490–499.

Hoskisson, P. A. & Hobbs, G. (2005). Continuous culture – making a comeback? *Microbiology* **151**, 3153–3159.

Ishii, N., Nakahigashi, K., Baba, T., Robert, M., Soga, T., Kanai, A., Hirasawa, T., Naba, M., Hirai, K. & other authors (2007). Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science* **316**, 593–597.

Jain, S. K., Paul-Satyaseela, M., Lamichhane, G., Kim, K. S. & Bishai, W. R. (2006). *Mycobacterium tuberculosis* invasion and traversal across an *in vitro* human blood-brain barrier as a pathogenic mechanism for central nervous system tuberculosis. *J Infect Dis* **193**, 1287–1295.

Jakimowicz, P., Cheesman, M. R., Bishai, W. R., Chater, K. F., Thomson, A. J. & Buttner, M. J. (2005). Evidence that the *Streptomyces* developmental protein WhiD, a member of the WhiB family, binds a [4Fe–4S] cluster. *J Biol Chem* **280**, 8309–8315.

James, B. W., Williams, A. & Marsh, P. D. (2000). The physiology and pathogenicity of *Mycobacterium tuberculosis* grown under controlled conditions in a defined medium. *J Appl Microbiol* **88**, 669–677.

Jungblut, P. R., Schaible, U. E., Mollenkopf, H. J., Zimny-Arndt, U., Raupach, B., Mattow, J., Halada, P., Lamer, S., Hagens, K., Kaufmann, S. H. & other authors (1999). Comparative proteome analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG strains: towards functional genomics of microbial pathogens. *Mol Microbiol* **33**, 1103–1117.

Kadokura, H., Katzen, F. & Beckwith, J. (2003). Protein disulfide bond formation in prokaryotes. *Annu Rev Biochem* **72**, 111–135.

Keating, L. A., Wheeler, P. R., Mansoor, H., Inwald, J. K., Dale, J., Hewinson, R. G. & Gordon, S. V. (2005). 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* **56**, 163–174.

Keren, I., Shah, D., Spoering, A., Kaldalu, N. & Lewis, K. (2004). Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. *J Bacteriol* **186**, 8172–8180.

Malli, R. & Epstein, W. (1998). Expression of the Kdp ATPase is consistent with regulation by turgor pressure. *J Bacteriol* **180**, 5102–5108.

Maurer, L. M., Yohannes, E., Bondurant, S. S., Radmacher, M. & Slonczewski, J. L. (2005). pH regulates genes for flagellar motility, catabolism, and oxidative stress in *Escherichia coli* K-12. *J Bacteriol* **187**, 304–319.

Mostowy, S., Cleto, C., Sherman, D. R. & Behr, M. A. (2004). The *Mycobacterium tuberculosis* complex transcriptome of attenuation. *Tuberculosis (Edinb)* **84**, 197–204.

Mostowy, S., Inwald, J., Gordon, S., Martin, C., Warren, R., Kremer, K., Cousins, D. & Behr, M. A. (2005). Revisiting the evolution of *Mycobacterium bovis*. *J Bacteriol* **187**, 6386–6395.

Pandey, D. P. & Gerdes, K. (2005). Toxin–antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res* **33**, 966–976.

Rehren, G., Walters, S., Fontan, P., Smith, I. & Zarraga, A. M. (2007). Differential gene expression between *Mycobacterium bovis* and *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* **87**, 347–359.

Rodrigue, S., Brodeur, J., Jacques, P. E., Gervais, A. L., Brzezinski, R. & Gaudreau, L. (2007). Identification of mycobacterial sigma factor binding sites by chromatin immunoprecipitation assays. *J Bacteriol* **189**, 1505–1513.

Rosas-Magallanes, V., Stadthagen-Gomez, G., Rauzier, J., Barreiro, L. B., Tailleux, L., Boudou, F., Griffin, R., Nigou, J., Jackson, M. & other authors (2007). Signature-tagged transposon mutagenesis identifies novel *Mycobacterium tuberculosis* genes involved in the parasitism of human macrophages. *Infect Immun* **75**, 504–507.

Said-Salim, B., Mostowy, S., Kristof, A. S. & Behr, M. A. (2006). Mutations in *Mycobacterium tuberculosis* Rv0444c, the gene encoding anti-SigK, explain high level expression of MPB70 and MPB83 in *Mycobacterium bovis*. *Mol Microbiol* **62**, 1251–1263.

Sassetti, C. M. & Rubin, E. J. (2003). Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A* **100**, 12989–12994.

Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M., Dolganov, G., Efron, B., Butcher, P. A. & other authors (2003). Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. *J Exp Med* **198**, 693–704.

Singh, A., Jain, S., Gupta, S., Das, T. & Tyagi, A. K. (2003). *mymA* operon of *Mycobacterium tuberculosis*: its regulation and importance in the cell envelope. *FEMS Microbiol Lett* **227**, 53–63.

Sirakova, T. D., Thirumala, A. K., Dubey, V. S., Sprecher, H. & Kolattukudy, P. E. (2001). The *Mycobacterium tuberculosis pks2* gene encodes the synthase for the hepta- and octamethyl-branched fatty acids required for sulfolipid synthesis. *J Biol Chem* **276**, 16833–16839.

Smith, N. H., Kremer, K., Inwald, J., Dale, J., Driscoll, J. R., Gordon, S. V., van Soolingen, D., Hewinson, R. G. & Smith, J. M. (2006). Ecotypes of the *Mycobacterium tuberculosis* complex. *J Theor Biol* **239**, 220–225.

Stermann, M., Sedlacek, L., Maass, S. & Bange, F. C. (2004). A promoter mutation causes differential nitrate reductase activity of *Mycobacterium tuberculosis* and *Mycobacterium bovis*. *J Bacteriol* **186**, 2856–2861.

Stewart, G. R., Wernisch, L., Stabler, R., Mangan, J. A., Hinds, J., Laing, K. G., Young, D. B. & Butcher, P. D. (2002). Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. *Microbiology* **148**, 3129–3138.

Sulzenbacher, G., Canaan, S., Bordat, Y., Neyrolles, O., Stadthagen, G., Roig-Zamboni, V., Rauzier, J., Maurin, D., Laval, F. & other authors (2006). LppX is a lipoprotein required for the translocation of phthiocerol dimycocerosates to the surface of *Mycobacterium tuberculosis*. *EMBO J* **25**, 1436–1444.

van Soolingen, D., de Haas, P. E., Haagsma, J., Eger, T., Hermans, P. W., Ritacco, V., Alito, A. & van Embden, J. D. (1994). Use of various genetic markers in differentiation of *Mycobacterium bovis* strains from animals and humans and for studying epidemiology of bovine tuberculosis. *J Clin Microbiol* **32**, 2425–2433.

Walters, S. B., Dubnau, E., Kolesnikova, I., Laval, F., Daffe, M. & Smith, I. (2006). The *Mycobacterium tuberculosis* PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. *Mol Microbiol* **60**, 312–330.

Weeks, D. L., Eskandari, S., Scott, D. R. & Sachs, G. (2000). A H^+-gated urea channel: the link between *Helicobacter pylori* urease and gastric colonization. *Science* **287**, 482–485.

Edited by: J. Parkhill