A Genome-Wide Regulator–DNA Interaction Network in the Human Pathogen Mycobacterium tuberculosis H37Rv

Jumei Zeng,† Tao Cui,† and Zheng-Guo He*  
National Key Laboratory of Agricultural Microbiology, Center for Proteomics Research, College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070, China

† Supporting Information

ABSTRACT: Transcription regulation translates static genome information to dynamic cell behaviors, making it central to understanding how cells interact with and adapt to their environment. However, only a limited number of transcription regulators and their target genes have been identified in the pathogen Mycobacterium tuberculosis, which has greatly impeded our understanding of its pathogenesis and virulence. In this study, we constructed a genome-wide transcription regulatory network of M. tuberculosis H37Rv using a high-throughput bacterial one-hybrid technique. A transcription factor skeleton network was derived on the basis of the identification of more than 5400 protein–DNA interactions. Our findings further highlight the regulatory mechanism of the mammalian cell entry 1 (mce1) module, which includes mce1R and the mce1 operon. Mce1R was linked to global negative regulation of cell growth, but was found to be positively regulated by the dormancy response regulator DevR. Expression of the mce1 operon was shown to be negatively regulated by the virulence regulator PhoP. These findings provide important new insights into the molecular mechanisms of several mce1 module-related hypervirulence phenotypes of the pathogen. Furthermore, a model of mce1 module-centered signal circuit for dormancy regulation in M. tuberculosis is proposed and discussed.

KEYWORDS: M. tuberculosis, transcription regulation network, Mce1, DevR

INTRODUCTION

Mycobacterium tuberculosis is the causative agent of tuberculosis, a leading cause of mortality worldwide. During infection, the intracellular pathogen confronts a particular host cell environment, causing reduced oxygen tension and restricted access to nutrients, among others. M. tuberculosis can withstand various stress conditions encountered within phagosomes, but the genetic mechanisms behind this are unknown.1 Although the genome of M. tuberculosis encodes ~200 putative transcription regulators, the identity and targets of only a few of them are known to date.2 The gene expression profiles of M. tuberculosis under various environmental stresses, such as hypoxia, nitric oxide release, nutrient starvation, low pH and drug exposure, and during infection of the lungs and macrophages have been extensively studied.3 Despite the abundance of studies on expression profiles, the transcription regulation (TR) mechanisms and regulators involved in mediating bacterial response to stress are largely unknown. Several studies have underscored the importance of having a comprehensive TR network for better understanding of the pathogenesis and survival of M. tuberculosis.4,5 By collecting documented regulatory information and mapping the homologous regulatory relationship of Escherichia coli and Corynebacterium glutamicum, several groups6–8 have assembled the M. tuberculosis TR network, which has improved our understanding of transcription regulation in M. tuberculosis. However, the regulators and target genes considered in such studies remain limited.

TR depends on specific physical interactions between transcription factors (TFs) and the promoter sequences of their target genes. Advances in high-throughput genetic techniques have made it possible to analyze the direct interaction between a trans regulator and a cis-promoter at a whole genome scale. Such innovations include DNA microarray,9 ChIP-Seq10 yeast one-hybrid system11 and bacteria one-hybrid system.12 To identify protein–DNA interactions (PDIs) for known TFs in M. tuberculosis, we screened the regulatory sequence library using a recently developed bacterial one-hybrid (B1H) technique.12 We characterized more than 5400 PDIs, whose overall quality was validated. A genome-wide regulatory network of M. tuberculosis was then constructed and network topology was analyzed. To effectively simplify the original network while preserving its major architecture, a TF–TF skeleton network was extracted, which highlighted the regulatory mechanism of the mammalian cell entry 1 (mce1) module, including mce1R and mce1 operon genes. On the basis of this experimental study and previous reports, we propose a model of mce1 module-centered circuit for regulation of dormancy in M. tuberculosis.

Received: June 9, 2012
MATERIALS AND METHODS

Definition of OPERONS and promoters

The complete genome sequence of *M. tuberculosis* and gene location information was downloaded from NCBI (NC_000962.2). Genes with the same transcriptional directions and small intergenic distances (d) between adjacent genes (d ≤ 20) are hypothesized to compose a hypothetical operon that shares the same promoters, and were cotranscribed in vivo. Upstream 350-bp sequences of the first genes of each hypothetical operon were selected as hypothetical promoters for the determination of confidence levels.

Generation of B1H promoter prey library and TFs bait strains

Using the pBXcmT vector, a pROMOTER vector was constructed and integrated with multiple cloning sites. About 2637 promoters were amplified using their respective primers (Table S1, Supporting Information) and cloned into the pROMOTER. The prey library for *M. tuberculosis* promoters was produced by mixing these recombinant plasmids. About 190 predicted transcription regulatory genes of *M. tuberculosis* were amplified using their primers, and 184 transcription regulators were successfully cloned into the pTRG vector.

Bacterial one-hybrid screens

A pair of pROMOTER/pTRG plasmids was cotransformed into the report strain as described previously. Its growth was then tested together with the self-activation control on selective medium containing 3-AT, Kan+, Str-, and Chlor. Positive growth cotransformants were selected on selective screening medium containing 20 mM 3-AT, 16 μg/mL of streptomycin, 15 μg/mL of tetracycline, 34 μg/mL of chloramphenicol, and 50 μg/mL of kanamycin. The plates were incubated at 30 °C for 3–4 d. Potential positive clones were chosen and verified by sequencing, and interactors were identified by BLAST alignment.

In silico DNA-binding motif analysis

Conserved binding motifs for TFs were analyzed in silico with the MEME suite. The latest MEME Suite Software (4.3.0) was downloaded from the MEME Web site (http://meme.sdsc.edu/) and run on a local Linux platform. Minimum and maximum motif widths were set to 8 and 12, respectively. The distribution of motifs was set to OOPS mode. The remaining algorithm-related parameters were kept at default settings.

GO annotation and GO enrichment

The GO terms’ annotation of proteins encoded by the *M. tuberculosis* H37Rv genome was obtained from TBDB. The GO enrichment of target genes was performed by the online enrichment tool provided by TBDB (http://genome.tbdb.org/annotation/genome/tbdb/BatchSelect.html?target=GeneEnrichment.html).

Network visualization and analysis of topological properties

Networks were visualized using Cytoscape. The topological properties of networks were analyzed using the Network-Analysis plug-in of Cytoscape. Networks were visualized using Cytoscape. The topological properties of networks were analyzed using the NetworkAnalysis plug-in of Cytoscape.20 Edges in the regulatory networks were analyzed as directed.

Expression and purification of recombinant TF proteins of *M. tuberculosis*

The transcriptional regulatory genes of *M. tuberculosis* were cloned into the pET28a overexpression vectors to produce corresponding recombinant vectors. *E. coli* BL21 (DE3) cells transformed with the recombinant plasmids were grown at 37 °C in 200 mL of LB medium containing 50 mg/mL of kanamycin. Protein purification was carried out as described previously and verified by SDS-PAGE. Protein concentrations were determined by spectrophotometric absorbance at 260 nm.

**RNA POLYMERASE II-CHROMATIN IMMUNOPRECIPITATION (RNAPOL-CHIP) ASSAY**

Each cotransformant containing both pROMOTER and pTRG plasmids was grown up to OD600 = 0.5, then IPTG was added to a final concentration of 0.5 mM, and the culture was incubated for 4 h at 37 °C with shaking. The culture was fixed with 1% formaldehyde at room temperature and stopped by adding glycerol (0.125 M final concentration). Cross-linked cells were pelleted by centrifugation and washed in PBS and TBSTT (150 mM NaCl, 20 mM Tris-HCl pH 7.5, 10 mM EDTA, 0.1% Triton X-100, and 0.1% Tween 20). After resuspending in TBSTT, the sample was sonicated on ice. The average DNA fragment size obtained was approximately 0.5 kb. After sonication, insoluble debris was removed by centrifugation. The resulting extract was incubated under rotation for 3 h at 4 °C, with 10 μL of antibodies against RNA pol-α subunit. Subsequently, the complexes were immunoprecipitated with 20 μL of 50% protein A-Sepharose for 1 h under rotation at 4 °C. The immunocomplex was recovered by centrifugation, washed five times with 1× TBSTT, and resuspended with 100 μL of TE (20 mM Tris-HCl pH 7.8, 10 mM EDTA, and 0.5% SDS). Cross-linking was reversed for 1 h at 65 °C. The immunoprecipitated DNA was purified and resuspended in 50 μL of TE. DNA in the samples (1:10 dilution) were analyzed by quantitative PCR with corresponding primers (Table S1, Supporting Information), as described previously.

**ELECTrophoretic Mobility Shift ASSay (EMSa)**

DNA-binding assays of *M. tuberculosis* TF proteins were performed using a modified EMSA, as previously described. When required, acetylphosphate (ACP) (Sigma) was used to phosphorylate PhoP (Rv0757) or DevR (Rv3133c, also known as DosR) in vitro. Briefly, the reactions (10 μL) for measuring the mobility shift contained 200 fmol 32P-labeled DNA and various TF proteins diluted in a buffer containing 20 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.5 mM MgCl2, 10 μg/mL of sonicated salmon sperm DNA, 0.7 mM 2-mercaptoethanol and 5% glycerol. Gels were exposed to a storage-phosphor screen and the images were acquired using a Typhoon scanner (GE Healthcare).

**Surface Plasmon Resonance (SPR) Analysis**

Interactions between regulator proteins of *M. tuberculosis* and their target DNA were assayed using SPR. Biotin-labeled promoter DNA was immobilized onto a Streptavidin (SA) chip (BIAcore). Cells were harvested when grown to midlog phase, and ChIP analysis was performed as described previously. DNA–protein interaction assays were performed at 25 °C. An overlay plot was generated to show the interactions.

**M. tuberculosis Growth, Gene Overexpression and Antisense Inhibition Assays**

*M. tuberculosis* H37Rv was incubated at 37 °C in 7H9 media with 0.05% Tween 80, 0.2% glycerol, and OADC supplement (Becton Dickinson). Cells were harvested when grown to midlog phase, and ChIP analysis was performed as described previously. Vectors for overexpressing Mce1R (Rv0165c), DevR, or PhoP were constructed using pMV261. To analyze the effects of an inhibitory expression of mce1R on its target genes, antisense mce1R was cloned into pMIND (a TetR-controlled expression vector).
RNA isolation

Total RNA was isolated from *M. tuberculosis* culture pellets using a Pure RNA mini kit (Biotake) as described previously. Briefly, the pellet was resuspended in an appropriate volume of the lysis buffer provided in the kit. Then the cell debris was removed by centrifugation at 2000g for 5 min at 4 °C. The supernatant was collected and processed according to the manufacturer’s instructions. To remove residual DNA from the RNA preparation, the RNA sample was incubated at 37 °C for 1 h with 100 units of DNase in the presence of DNase buffer, and then extracted with phenol/chloroform and precipitated with ethanol. The precipitated RNA was recovered by centrifugation at 13000g for 30 min at 4 °C and resuspended in 50 μL of RNase-free water. The A_{260} values of 1 μL samples were checked to assess yield.

Quantitative Real-Time PCR (qRT-PCR)

For qRT-PCR analysis, gene-specific primers (Table S2, Supporting Information) were used, and first-strand cDNAs were synthesized using SuperScript II reverse transcriptase (Invitrogen) according to the manufacturer’s instructions. Each PCR reaction (20 μL) contained 10 μL of 2X SYBR Green Master Mix Reagent (Applied Biosystems), 1.0 μL of cDNA samples, and 200 nM primers. The reactions were performed in a Bio-Rad IQ5 RT-PCR machine. The thermocycling conditions were as follows: 95 °C for 5 min, and 40 cycles at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s. Amplification specificity was assessed by conducting a melting curve analysis. Gene expression levels were normalized to the levels of 16S rRNA gene transcripts. The degrees of change in expression were calculated using the 2^{-ΔΔC_t} method.

RESULTS

Identification of Regulator—Promoter DNA Interactions in *M. tuberculosis*

We employed the well-established B1H system to generate a genome-wide PDI network of *M. tuberculosis* (Figure 1). In accordance with the defined operon map of *M. tuberculosis* H37Rv (see the Materials and Methods), 2637 promoters were amplified using their respective primers (Table S1, Supporting Information) and were cloned into the modified pBXcmT vector, designated as PROMOTER. A promoter library of *M. tuberculosis* was then generated by mixing these promoter recombinant plasmids. About 190 genes encoding TFs were cloned into the pTRG vector, resulting in 1298 TF-promoter interactions (Table S3, Supporting Information).

Construction and Topological Properties Analysis of *M. tuberculosis* TR Network

The *M. tuberculosis* TR network was then built by mapping each promoter in the PDI database to the genes whose expression was controlled by the promoter. The final regulatory network contains 184 TFs, 1116 target genes, and 8510 TF-gene interactions (Table S4, Supporting Information). The entire network is highly interconnected without any isolated nodes (Figure 1). The average path length of the network is 2.619, which suggests the “small-world” effect of the network. In the directed TR network, in-degree of a gene refers to the number of transcriptional factors regulating its expression, while out-degree of a regulator shows the number of genes under its direct control. The in-degree of *M. tuberculosis* B1H TR network ranges from 1 to 142 and follows a power-law distribution (Figure S1A, Supporting Information). The number of genes decline rapidly when in-degree increases. A total of 712 genes (63.79% of the network) are regulated by a single TF, which reflects specific regulation. The out-degree of regulators ranges from 1 to 130 and follows an almost Gaussian-like distribution (Figure S1B, Supporting Information). Rx2021c has the highest out-degree and regulates the expression of 130 genes.

Degree analysis of a TR network can reflect its global topological properties, while network motifs can reveal local properties of a network. Motifs are small connected subnetworks that a network displays in significantly higher frequencies than would be expected for a random network. Motifs in TR networks represent the basic building blocks of the network and thus are essential for uncovering structural design principles that control gene expression in cells. Using famod to compare 100 derived random networks, we calculated the significance profiles for the *M. tuberculosis* B1H TR network focusing on the motifs of 3 nodes. Out of 13 possible directed 3-nodes motifs, motif 13 was not found in the B1H TR network. The first six motifs, which correspond to open motifs, were found more frequently than the random version of the TR network, while...
282 closed motifs were mostly underrepresented in the B1H TR
283 network (Figure S1C, Supporting Information).

 Modules in the network are formed by groups of genes that
285 play roles in related cellular function and are relatively isolated
286 from other parts of the network. As full connectivity feature, no
287 clearly physically separated modules could be directly identi-
288 fi ed in our current TR network. Several tools have been used in the
289 past to find highly interconnected regions and identify modules
290 in protein interaction networks. We performed a modules search
291 in our current TR network using the MCODE tool that was
292 originally designed for analysis of protein interaction networks.30

 Many clusters formed by either relatively intensive TF–TF
293 interaction or groups of genes coregulated by groups of
294 regulators were identified (Figure S2, Supporting Information).
Uncovering the biological function of these clusters will require
295 wide range information combined with experimental con-
296 firma-
297 tion in the future. In the present study, the regulatory mechanism
298 of the mammalian cell entry 1 (mce1) module was further studied
299 as a model of TF–TF interaction.

Quality Evaluation of the PDI Network

Two independent methods, RNA polymerase II-chromatin
immunoprecipitation (RNAPol-ChIP)33 and electrophoretic
mobility shift assay (EMSA) were chosen for evaluating the
quality of the PDI data. The high-degree feature in our network is

---

Figure 2. Analysis of the TF–TF skeleton network. (A) In-degree distribution of the TF–TF network. (B) Out-degree distribution of the TF–TF network. (C) High in-degree nodes in the TF–TF network. (D) High in-degree nodes in the TF–TF network had a high degree of cross-regulation. (E) A simplified framework of the TF–TF network. Nodes in the TF–TF network were divided into five groups according to their in-degree. Each group is indicated with a different color. The sizes of the nodes correspond to the number of regulators in the groups. The height of the edges corresponds to the number of total interactions between two groups. (F) Loop paths in the M. tuberculosis TR network. (G) Rv2034-centered loop paths.

---

<table>
<thead>
<tr>
<th>Regulator</th>
<th>Gene Name</th>
<th>In-degree</th>
<th>Out-degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rv1219c</td>
<td>-</td>
<td>142</td>
<td>3</td>
</tr>
<tr>
<td>Rv3260c</td>
<td>whiB2</td>
<td>105</td>
<td>0</td>
</tr>
<tr>
<td>Rv1990c</td>
<td>-</td>
<td>77</td>
<td>4</td>
</tr>
<tr>
<td>Rv2034</td>
<td>-</td>
<td>77</td>
<td>9</td>
</tr>
<tr>
<td>Rv3160c</td>
<td>-</td>
<td>68</td>
<td>1</td>
</tr>
<tr>
<td>Rv0232</td>
<td>-</td>
<td>63</td>
<td>5</td>
</tr>
<tr>
<td>Rv0001</td>
<td>dnaA</td>
<td>51</td>
<td>8</td>
</tr>
<tr>
<td>Rv1909c</td>
<td>furA</td>
<td>49</td>
<td>0</td>
</tr>
<tr>
<td>Rv1049</td>
<td>-</td>
<td>39</td>
<td>7</td>
</tr>
<tr>
<td>Rv0792c</td>
<td>-</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>Rv0474</td>
<td>-</td>
<td>29</td>
<td>9</td>
</tr>
<tr>
<td>Rv0022c</td>
<td>whiB5</td>
<td>25</td>
<td>9</td>
</tr>
<tr>
<td>Rv0328</td>
<td>-</td>
<td>24</td>
<td>7</td>
</tr>
<tr>
<td>Rv1534</td>
<td>-</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Rv3082c</td>
<td>virS</td>
<td>18</td>
<td>5</td>
</tr>
<tr>
<td>Rv1994c</td>
<td>-</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>Rv3240c</td>
<td>-</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Rv0586</td>
<td>mce2R</td>
<td>10</td>
<td>3</td>
</tr>
</tbody>
</table>
similar to that of previously published networks. However, high out-degree regulators and high in-degree promoters likely contain potential false positive interactions. Therefore, 10 high out-degree regulators, including Mce2R (Rv0586c), IdeR (Rv2711c), DevR, CarD (Rv3583c), Rv0275c, and Rv1219c, and 3 high in-degree promoters, including dnaAp (Rv0001p), dnaKp (Rv0350p), and fadE26p (Rv3504p), were first selected for validation by the RNAPol-ChIP approach (Figure S3, Supporting Information). In addition, 33 pairs of interactions involved in continuous promoters were randomly selected and evaluated. In

Figure 3. Mce1R as a global negative regulator. (A) ChIP assays for the binding of Mce1R with its putative target promoters in M. tuberculosis. ChIP using preimmune or anti-Mce1R antibody raised against Mce1R. DNA samples of the input and ChIP were purified and resuspended in 50 μL of TE. DNA recovered from immunoprecipitates was amplified with primers specific for target genes or to an unrelated mycobacterial promoter SigM. (B) The sequence motif predicted from the promoters of Mce1R target regulators. (C) qRT-PCR assays for the levels of target gene expression in M. tuberculosis overexpressing the antisense mce1R. Relative expression levels of genes were normalized using the 16S rRNA gene as an invariant transcript. Fold change in expression was calculated using the 2ΔΔCt method. All target genes were amplified using specific primers. Representative data are shown. (D) qRT-PCR assays for target gene expression in M. tuberculosis overexpressing mce1R. (E) EMSA assays for the binding of Mce1R with its promoter DNA. 200 fmol 32P-labeled DNA substrates were coincubated with various amounts of the Mce1R protein (100–400 nM). Heat-denatured Mce1R protein was used as the negative control. (F) SPR assays for the binding of Mce1R with its own promoter. Biotin-labeled mce1R promoter was immobilized on an SA chip (BIAcore). Following a period of stabilization, Mce1R, heated-Mce1R or GST was passed over the chip. An overlay plot was then produced to look for interactions.
total, 536 of 598 interactions were confirmed, and the rate of false positives was less than 10% (Figure S3, Supporting Information). The transcriptional regulators KmtR (Rv0827c), IdeR, WhiB3 (Rv3416), and Rv3066, as well as the 33 pair interactions were also independently validated by EMSA experiments. Additionally, 230 of 261 DNA-binding activities were confirmed, and the average false rate was appropriate 12%. A summary of the regulator-gene interactions validated by the two different methods described here is provided in Table S5 (Supporting Information).

Figure 4. Regulation analysis of mce1R and the mce1 operon. (A) EMSA assays for the binding of DevR with the promoter of mce1R. 200 fmol [32P]-labeled DNA substrates were coincubated with various amounts of the DevR protein (100−400 nM). The effect of phosphorylation of DevR on binding with the mce1R promoter was assayed by adding 50 mM ACP to the reactions. (B) SPR assays for binding of DevR with the mce1R promoter. Biotin-labeled mce1R promoter was immobilized on an SA chip (BIAcore). DevR, heat-denatured DevR, or GST was passed over the chip. An overlay plot was then produced to look for interactions. Assays for the effects of phosphorylation of DevR on binding with the mce1R promoter were carried out by adding 50 mM ACP to the reactions. (C) ChIP assays for the binding of DevR and PhoP with their target promoters in M. tuberculosis. ChIP using preimmune or specific antibody raised against DevR or PhoP. DNA recovered from immunoprecipitates was amplified with primers specific for target genes or to an unrelated mycobacterial promoter SigM. (D) qRT-PCR assays for changes in gene expression levels upon overexpression of DevR or PhoP. (E) EMSA assays for binding of PhoP with the mce1R promoter. (F) SPR assays for binding of PhoP with the mce1R promoter. Biotin-labeled mce1R promoter was immobilized on an SA chip. The effects of phosphorylation of PhoP on binding with the mce1R promoter were assayed by adding 50 mM ACP to the reactions.
Taken together, about 800 pairs of PDIs were evaluated, and more than 80% of the interactions were confirmed (Table S6, Supporting Information).

**Extraction of the TF–TF Skeleton Network and Its Characteristics**

The extraction of a TF–TF skeleton network from the TR network can effectively simplify it while preserving its architecture and information on the regulatory signal stream. A TF–TF skeleton network of *M. tuberculosis* was produced, and the degree of each TF was calculated. Interestingly, it followed a power-law and had Gaussian distribution (Figure 2A,B). There were about 18 TFs with in-degree higher than 10 and comprised the high in-degree nodes (Figure 2C). A certain number of input regulation of these high in-degree group TFs came from members of the group itself; these nodes were thus highly regulated internally by each other (Figure 2D). For a more intuitive exploration of network architecture and signal stream, a simplified framework of the TF–TF network was drawn by group TFs according to their in-degree (Figure 2E). The framework showed that direct regulation of targets was the most common mode of regulation in the network. The multicascade portion decreased along with increase in cascade hierarchy (Figure 2E). Further path analysis of the TF–TF network identified the number of loop paths (Figure 2F). In general, more than one loop path exists for a TF. For example, when the maximum path length was limited to three, there were a total of 12 loop paths for the ArsR-type repressor protein Rv2034c (Figure 2G). These regulators are possibly responsible for the multifeedback paths and stability of the global network. Figure S4 (Supporting Information) shows the individual loop path network for each regulator.

**Mce1R Functions as a Global Negative Regulator in the Network**

In the TF–TF skeleton network, a GntR family repressor, Mce1R, directly regulated 12 regulators including Mce1R itself, VirS (Rv3082c), Lsr2 (Rv3597c), FurA (Rv1909c), three WhiB family regulators (Whib2 (Rv3260c), Whib5 (Rv0022c) and Whib7 (Rv3197A)), DNA replication initiation protein DnaA, and four other TFs (Rv1049, Rv1909c, Rv1219c, Rv2034). This suggests that Mce1R is a potential high-level regulator. The interactions between Mce1R and most of these target promoters were confirmed by B1H spot assay and validated by EMSA using purified Mce1R protein (Figure S5A,B, Supporting Information). In addition, using ChIP with anti-Mce1R antibody, we confirmed the recognition by Mce1R regulator of these target promoter DNA in *M. tuberculosis* H37Rv (Figure 3A), except for the promoter of Rv2034. The sequence of these promoters was further analyzed in silico with MEME, and an 8-bp conserved DNA motif was identified (Figure 3B).

By RT-PCR assays, up-regulation in the expression of target regulator genes was clearly observed when the antisense mRNA of the mce1R gene was expressed through a pMind recombinant plasmid (Figure 3C). Conversely, a significant down-regulation in the expression of these target genes was observed when Mce1R was overexpressed using the pMV expression vector (Figure 3D). Interestingly, a significant up-regulation of mce1R (about 4.5-fold) was observed one day after induction. However, down-regulation of most of the target regulators did not occur in that time frame (Figure 3D, upper panel). At a later time point (after seven days), the mce1R gene itself was also down-regulated (Figure 3D, lower panel). This phenomenon can be attributed to feedback regulation of Mce1R because the specific DNA-binding activity of Mce1R with its own promoter was confirmed by EMSA and surface plasmon resonance (SPR) assays (Figure 3E,F).

To further investigate the potential function of Mce1R, all of its target genes were analyzed in the context of the GO catalog (Table S7, Supporting Information). None of the biological processes were found to be significantly enriched, which suggested a wide regulatory role of Mce1R in cells. However, it is notable that many target genes (Rv0282, Rv0283, Rv0284, Rv0285, Rv0286, Rv0674, Rv0675, Rv0700, Rv0701, Rv0702, Rv0703, Rv1194c, Rv2500c, Rv2827c, Rv3177, Rv3597c) are involved in the process of cell “growth” (GO: 0040007). Notably, one of the target proteins Rv0284 is homologous to the FtsK/SpoIIE family of proteins, suggesting that Rv0284 may play important roles during cell division.

Taken together, our network analysis and expression assays indicated that Mce1R represents a global repressor in *M. tuberculosis*, likely by regulating genes involved in cell growth and cell cycle.

**Dormancy Regulator DevR Positively Regulates the Expression of mce1R**

The upstream regulator of *mce1R* was analyzed further in the current regulatory network to explore the potential actions of Mce1R in cell processes. The *mce1R* gene was found to be regulated by a two-component response regulator (DevR) involved in the regulation of *M. tuberculosis* dormancy (Table S4, Supporting Information). This interaction was confirmed by B1H spot assay (Figure S5C, Supporting Information). An EMSA assay was performed to further validate the specific DNA-binding activity of DevR in vitro. As shown in Figure 4A, a band of retarded mobility, indicating protein bound to the DNA, was evident when mce1R promoter DNA was mixed with DevR.

DevR is the response regulator of the DevRS two-component system and can be phosphorylated by the cognate histidine sensor kinase DevS. To test the effect of phosphorylation on the ability of DevR to bind to the mce1R promoter, an EMSA assay was performed using acetylphosphate (ACP)-preincubated DevR protein. Interestingly, phosphorylation of DevR resulted in a weaker DNA-binding activity as evidenced by the fewer DevR/ DNA complexes observed on the gel compared with the reaction with unphosphorylated DevR (Figure 4A). The interaction of DevR with the mce1R promoter and the effect of phosphorylation were further confirmed by chip assay (Figure 4B). When an increasing concentration of DevR protein (100–500 nM) was passed over the chip surface, an increasing trend in response values was observed (Figure 4B, upper panel). By contrast, both heated inactive protein and GST showed no response (Figure 4B, upper panel). This result is consistent with those of the EMSA assays. More interestingly, a conserved binding motif (TTTGGGACCTWWAGTC_CCSAA) of DevR was also identified in the promoter of mce1R. This suggests that the potential binding region possesses a sequence motif that specifically interacts with unphosphorylated DevR (Table S8, Supporting Information).

DevR was also found bind to the mce1R promoter in *M. tuberculosis* in a ChIP assay. Figure 4C shows that the mce1R promoter DNA was specifically recovered by immunoprecipitation using the specific anti-DevR antiserum. In contrast, the preimmune serum failed to precipitate significant amounts of DNA. Furthermore, overexpression of DevR (by about 10-fold) increased the expression of mce1R by about 5-fold in *M. tuberculosis* (Figure 4D, left panel).
Our results suggest that DevR can specifically bind with the promoter of mce1R both in vitro and in vivo and positively regulates mce1R expression.

Virulence Regulator PhoP Regulates the Expression of mce1 Operon

In our network, the two-component response regulator PhoP, which is a key virulence regulator, participates in the regulation of the mce1 operon (Table S4, Supporting Information), which is tightly associated with the activation of host immune response during infection. Binding of PhoP to the mce1 promoter was confirmed by gel-shift assays in which a characteristic protein/DNA complex appeared while the amount of free promoter DNA decreased with increasing PhoP concentration (Figure 4E). The negative effect of phosphorylation on the ability of PhoP to bind to mce1p was clearly observed when protein samples were preincubated with ACP before using for assays. As shown in Figure 4E, phosphorylated PhoP had very low DNA binding ability. The interaction between PhoP and mce1p DNA as well as the effect of PhoP phosphorylation on their interaction was confirmed by SPR assay. As shown in Figure 4F, when an increasing concentration of the PhoP protein (100–500 nM) was passed over the chip surface, a corresponding increase in response values was observed (Figure 4F, upper panel). By contrast, ACP-preincubated PhoP protein showed almost no response when passed over the chip, which is similar to the result obtained for the heat-denatured protein or GST (Figure 4F, lower panel).

The binding of PhoP with the promoter of the mce1 operon in M. tuberculosis was also confirmed by ChIP assays. As shown in Figure 4C, the mce1 operon promoter DNA was specifically recovered by immunoprecipitation using anti-PhoP antiserum. In contrast, the preimmune serum failed to precipitate significant amounts of DNA. The physiological significance of the interaction between PhoP and mce1p was studied further by overexpression experiments in M. tuberculosis. When phoP was cloned into the pMV261 vector and overexpressed (about 7-fold) in M. tuberculosis (Figure 4D, right panel), about 2-fold down-regulation of expression of the mce1 operon was observed (Figure 4D, right panel).

These results strongly suggest that PhoP can physically interact with the mce1 promoter, and this interaction can be inhibited by phosphorylation. Consistent with this, we observed that PhoP can negatively regulate expression of the mce1 operon in M. tuberculosis.

DISCUSSION

TR translates static genome information to dynamic cell behaviors, making it central to understanding how cells interact and adapt to their environment. In the case of pathogens, a comprehensive understanding of their transcription regulatory network would enable the identification of new regulators, and aid the tracing of regulatory circuits that allow them to respond to...
host environments. In the present study, we constructed a genome-wide transcription regulatory network of the human pathogen *M. tuberculosis* on the basis of the B1H assay. Following a TF-centered screen strategy, almost all the TFs encoded by *M.* tuberculosi were included in the global network, covering about one-third of the defined hypothetical promoters.

Studies on specific PDIs and large-scale TR networks have become central topics in both prokaryotes and eukaryotes in the postgenomic era. However, most studies focus primarily on changes in the expression of individual genes, single regulons, or distinct subunits of the complex TR network of *M. tuberculosis*. In addition, most interactions derived from expression profiles are indirect protein–DNA interactions, and only a handful of interactions have been confirmed experimentally. Some previously or recently reported interactions were found in our global regulatory network (Table S9, Supporting Information). Importantly, a significant proportion of our interaction data was confirmed through various independent methods (80% success rate). The success rate in our network is similar to that of other published networks.

Several resources such as TBDB6 and MycoRegNet systems have been designed as data repositories and provide bioinformatic platforms for transcriptional regulatory network visualization, analysis, and reconstruction. We will release the data reported in this study for public use by integrating them into existing platforms in the near future.

The entire network is highly interconnected without any isolated nodes. Two kinds of regulation tend to facilitate this connectivity feature of a TR network. Direct regulation between regulators could link different regulators and genes under their control, while indirect links between different regulators could be established if these regulators shared common target genes. In our analysis of the B1H TR network, both of these possibilities were found to have contributed significantly to this connectivity feature on the basis of successful identification of many genes regulated by more than one regulator and considerable number of TF–TF interactions (see module analysis part). The peculiarly high in-degrees were found in the *M. tuberculosis* B1H TR network and a single *M. tuberculosis* gene or gene cluster could be regulated by as many as 142 regulators (Figure 2). While similar in-degrees have been observed for human PDI network, several features of dormant bacteria. *M. tuberculosis* enters dormancy when it encounters various stressful conditions exerted by the host immune system. The regulator DevR can respond to host environments and induce dormancy of *M. tuberculosis*. The *mce1* operon has been shown to be tightly associated with the activation of host immune response during infection. In the current study, Mce1R was found to act as a global negative regulator of cell growth. Additionally, Mce1R was identified as a downstream target of DevR, and PhoP was determined to be a negative regulator of the *mce1* operon (Figure 5A). These findings reveal a signaling pathway that links the initial dormancy signal to the final dormancy-associated cellular processes (Figure 5B). In this model, PhoP regulates expression of the *M. tuberculosis* *mce1* operon, which stimulates immune response. DevR detects the immune response signals such as NO or hypoxic stress. If a robust immune system exists, devR is activated and dormancy is induced via activation of a global negative regulator Mce1R, which participates in regulating dormancy by repressing the expression of cell cycle and metabolism-associated genes. In this signaling model, mutation of the *mce1* operon, devR, or mce1R breaks the signal stream and prevents the activation of dormancy, leading to active bacterial growth and replication.

Several *M. tuberculosis* mutants exhibit rapidly growing hypervirulence phenotypes. However, the mechanism of such phenotypes remains poorly understood. Interestingly, the components of the signaling pathway identified here, namely devR, mce1R and the *mce1* operon, have previously been reported to be associated with hypervirulence. According to our model, hypervirulence is caused by the failure of regulatory circuits to cooperate in controlling *M. tuberculosis* dormancy lifestyles, thereby causing uncontrolled replication. This is consistent with the known function of DevR, which is to respond to hypoxia and induce dormancy in bacteria. Deletion of devR generates a hypervirulent mutant that grows more rapidly in immunocompetent mice and in gamma interferon-activated macrophages.

The *mce1* operon is responsible for stimulating host immune responses, and its disruption causes a decrease in host immune response during infection. A robust immune response is required for inducing and maintaining the dormancy of *M.* tuberculosi in vivo. Despite lacking a direct role, the *mce1* operon contributes to dormancy regulation when an immune system exists. The hypervirulence phenotype in mutants of the *mce1* operon may result from failures in stimulating the immune response, thereby causing failure to activate dormancy. Mce1R negatively regulates expression of the *mce1* operon. That mutation of *mce1R* also causes hypervirulence is unexpected considering the hypervirulence phenotype of the *mce1* operon mutation. In the current study, we provide evidence to support that Mce1R negatively regulates bacterial growth and replication as a global regulator. In this model, hypervirulence caused by the *mce1R* mutation may be caused by a different mechanism that is independent of the *mce1* operon. One possible explanation is that mutation in the *mce1* operon reduces the host immune response and inhibits dormancy at an earlier stage. By contrast, Mce1R mutation leads to derepression and tends to activate infection, leading to failure in execution of dormancy signals at a later stage. Another phenotype of the *mce1R* mutation is that it can be more rapidly reactivated from dormancy than wild-type strains in a drug-treated mouse model. A possible explanation for this is that *mce1R* mutants do not actually enter dormancy but simply have temporary inhibition of growth.

PhoP has been shown to be an essential virulence regulator in *M. tuberculosis*. A single nucleotide polymorphism within the PhoPR operon has been confirmed to be the major factor contributing to avirulence in H37Rv. However, mechanisms for the contribution of PhoP to virulence are largely unknown.
because of limited knowledge on its direct target genes. In the present study, PhoP was found to bind to the promoter of the mce1 operon, and overexpression of PhoP caused downstream regulation of expression of the mce1 operon. Thus, a possible mechanism for the contribution of PhoP to virulence of M. tuberculosis is repression of the expression of the mce1 operon. Interestingly, a recent study showed that PhoP controls M. tuberculosis ESAT-6 secretion and specific T cell recognition and may contribute to virulence in part by facilitating pathogen–host interaction. Similar to the ESAT-6 system, production of the mce1 operon is another essential component that participates in interaction with host immune systems. Thus, both our results and those of previous studies support the model that the importance of PhoP in virulence may largely depend on its regulation of genes that interact with host immune systems. In addition, we found that phosphorylation negatively regulates the binding of PhoP to the mce1 operon. It is possible that an unknown physiological condition dissociates phosphorylated PhoP from the mce1 operon, thereby allowing expression of the mce1 operon.

CONCLUSION

In the present study, a genome-wide TF-promoter DNA interaction screen was conducted, and a global TR network of M. tuberculosis was constructed. The network contains more than 5400 pairs of PDIs, covering almost all TFs and one-third of M. tuberculosis genes. Further analysis of the network, in combination with experiments, uncovered a molecular regulatory pathway of the mce1 module. A signal circuit for dormancy regulation of M. tuberculosis is proposed. Our findings provide insight into the molecular mechanisms of several mce1 module-related hypervirulence phenotypes of the pathogen. The TR network reported here should serve as a significant resource for further studies on gene regulatory pathways in M. tuberculosis.

ASSOCIATED CONTENT

Supporting Information
Figures S1–S5 and Tables S1–S9. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: hezhengguo@hotmail.com; hezhengguo@mail.hzau.edu.cn. Tel: +86-27-87284300. Fax: +86-27-87280670.

Author Contributions
The authors equally contributed to this work.

Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The M. tuberculosis H37Rv and MDR10093006 strains were provided by Dr. J. Xiang (Wuhan Medical Treatment Center, Wuhan, China). We thank Prof. G. Zhao (Laboratory of Systems Biology, Chinese Academy of Sciences, Shanghai, China) for his kind help. This study was funded by the National Natural Science Foundation of China (31025002 and 30930003), Creative Research Groups of Hubei, and Hubei Chutian Scholar Program (to Z.-G.H.).

ABBREVIATIONS

ACP, acetylphosphate; PCR, polymerase chain reaction; TB, tuberculosis; SPR, surface plasmon resonance; TR, transcription regulation; TFs, transcription factors; PDIs, protein–DNA interactions; B1H, bacterial one-hybrid; SA, Streptavidin; GO, gene ontology; EMSA, electrophoretic mobility shift assay; GST, glutathione S-transferase; IPTG, isopropyl-β-D-thiogalactopyranoside

REFERENCES


H37Rv. J. Proteome Res. 2010, 9, 6665–77.


(23) Sala, C.; Forti, F.; Magnoni, F.; Ghisotti, D. The katG mRNA of Mycobacterium tuberculosis and Mycobacterium smegmatis is processed at its 5′ end and is stabilized by both a polypurine sequence and translation initiation. BMC Mol. Biol. 2008, 9, 33.


(33) Saini, D. K.; Malhotra, V.; Dey, D.; Pant, N.; Das, T. K.; Tyagi, J. S. DevR-DevS is a bona fide two-component system of Mycobacterium tuberculosis that is hypoxia-responsive in the absence of the DNA-binding domain of DevR. Microbiology 2004, 150, 865–78.


