

SCIENTIFIC REPORTS



OPEN

WhiB4 Regulates the *PE/PPE* Gene Family and is Essential for Virulence of *Mycobacterium marinum*

Jing Wu¹, Huan-wei Ru¹, Zhi-hao Xiang¹, Jun Jiang¹, Yu-chen Wang¹, Lu Zhang^{1,2,3} & Jun Liu^{1,4}

During the course of infection, pathogenic mycobacteria including *Mycobacterium tuberculosis* (*M. tb*) encounter host environments of variable oxygen tension, ranging from the hypoxic center of granulomas to the most oxygenated region in the lung cavities. Mycobacterial responses to changes of oxygen tension are critically related to infection outcomes, such as latency and reactivation. WhiB4 is an iron-sulfur containing transcription factor that is highly sensitive to oxygen exposure. In this study, we found that WhiB4 of *Mycobacterium marinum* (*M. marinum*), a pathogenic mycobacterial species that is closely related to *M. tb*, is required for its virulence. *M. marinum* Δ whiB4 exhibited defective intracellular replication in macrophages and diminished virulence in zebrafish. Histology analysis revealed that the host had successfully controlled Δ whiB4 bacteria, forming well-organized granulomas. RNA-seq analysis identified a large number of *pe/ppe* genes that were regulated by WhiB4, which provides an explanation for the essential role of WhiB4 in *M. marinum* virulence. Several antioxidant enzymes were also upregulated in Δ whiB4, supporting its role in modulation of oxidative stress response. Taken together, we have provided new insight into and proposed a model to explain the physiological role of WhiB4.

Tuberculosis, caused by *Mycobacterium tuberculosis* (*M. tb*), continues to be a major global health problem, causing 1.5 million deaths and 9.6 million new infections in 2014. In the majority of individuals infected with *M. tb*, the bacteria establish a latent, asymptomatic infection that can persist for decades^{1,2}. About 5–10% of these individuals will develop active disease in their lifetime and the risk of reactivation is markedly increased by host immunosuppression³. About one-third of the world's population is latently infected, representing a large reservoir for reactivation⁴. Understanding the genetic programs that facilitate the entry of *M. tb* into and emergence from latency will facilitate the development of novel therapeutic strategies.

During the course of infection, *M. tb* must encounter heterogeneous host environments, including the high oxygen tension in the lung alveolus⁵, the reactive oxygen and nitrogen species generated in activated macrophages⁶, as well as the nutrient depleted and hypoxic center of tuberculous granulomas^{7,8}. How *M. tb* survives these diverse environments is not fully understood. Of particular interest is the response of *M. tb* to changes of oxygen tension. Studies from human and animals suggest an intimate link between oxygen tension and the outcome of *M. tb* infection^{9–14}. As such, hypoxia is considered a major stimulus that triggers *M. tb* latency^{9,15} and is one of the most frequently used *in vitro* conditions to mimic the environment of human granulomas^{15–17}. Accordingly, sets of *M. tb* genes that respond to hypoxia have been identified, such as the extensively studied heme-based DosR/S/T regulon^{18–21}.

Little is known about the reactivation of *M. tb* from the persistent state, and presumably access to oxygen is an important requirement. Consistent with this notion, reactivation of latent TB in humans occurs most frequently in the upper lobes of the lung, the most oxygenated region of the body⁹. In the later stage of active pulmonary TB, lung cavities that connect to airways provide oxygen-rich environments, allowing *M. tb* to reach high density and subsequent spread²². Therefore, knowledge on the response of *M. tb* to oxygen exposure is critically important for understanding the pathogenesis of *M. tb*.

¹State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Science, Fudan University, Shanghai, China. ²Key Laboratory of Medical Molecular Virology of Ministries of Education and Health, Fudan University, Shanghai, China. ³Shanghai Engineering Research Center Of Industrial Microorganisms, Shanghai, China. ⁴Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. Correspondence and requests for materials should be addressed to L.Z. (email: zhangu407@fudan.edu.cn) or J.L. (email: jun.liu@utoronto.ca)

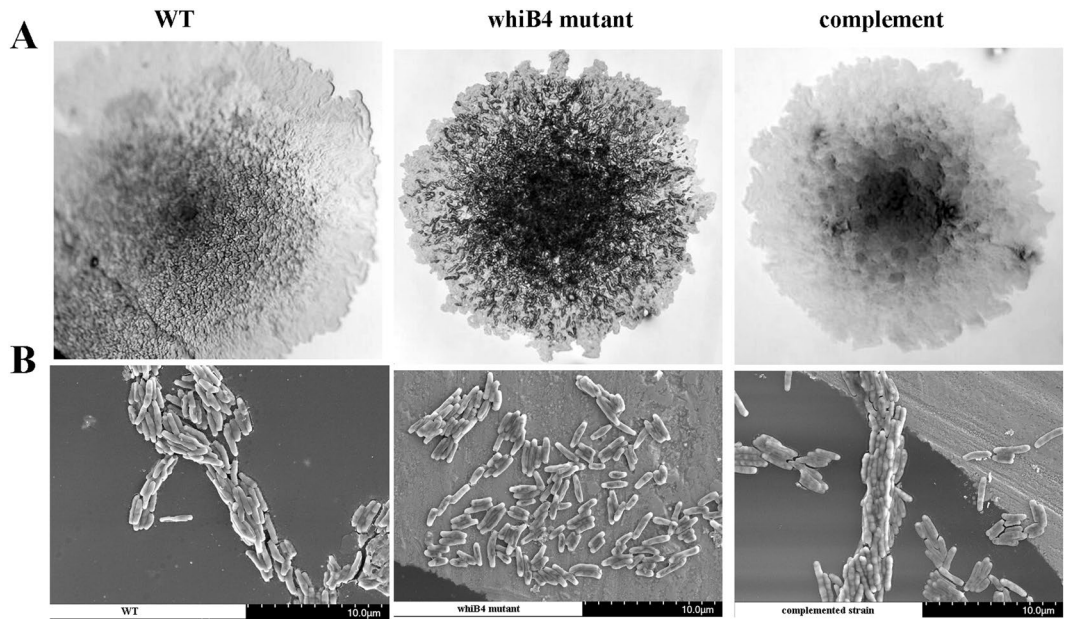


Figure 1. *M. marinum* $\Delta whiB4$ exhibited altered colony morphology and was defective in cording formation. (A) *M. marinum* $\Delta whiB4$ exhibited dry and rough colony morphology on 7H11 agar, as compared to as compared to the smoother, glossier, morphology of WT or the complemented strain. (B) Cells of WT or the complemented strain formed the serpentine cords, which were not observed for $\Delta whiB4$.

Several recent studies demonstrate that mycobacterial WhiB4 is an oxygen-sensitive transcription factor. WhiB4 is a member of the WhiB superfamily transcription factors that are conserved in actinomycetes, including many mycobacterial species such as *M. tb* and *M. marinum*^{23–28}. WhiB4 contains the highly conserved Cys-X14-22-Cys-X2-Cys-X5-Cys motif, which forms a [4Fe-4S] cluster in the holoenzyme^{29,30}. Upon exposure to O₂, WhiB4 rapidly (within minutes) loses the [4Fe-4S] cluster and leads to an oxidized apo-WhiB4 containing disulfide bonds. Importantly, the oxidized form of WhiB4 has the strongest DNA binding activity³⁰. The response of WhiB4 to oxygen exposure occurred much more rapidly than other WhiB proteins (e.g., WhiB1 and WhiB3), making it an oxygen sensor^{30–32}. The ability to rapidly respond to oxygen exposure and activate DNA binding activity implies an active role of WhiB4 in mycobacterial virulence. However, deletion of *whiB4* from *M. tb* resulted in hypervirulence in the lungs of guinea pigs, which was attributed to the increased expression of antioxidant enzymes including alkyl hydroperoxide reductases AhpC and AhpD in this strain³⁰. Therefore, WhiB4 in *M. tb* is primarily involved in modulation of oxidative stress response and is apparently not required for virulence. In this study, we found that WhiB4 plays a different role in *M. marinum*. Unlike in *M. tb*, inactivation of *whiB4* in *M. marinum* led to loss of virulence in zebrafish, a natural host of *M. marinum*, and impaired replication in macrophages. RNA-seq data revealed that WhiB4 positively regulates a large number of *pel/ppe* gene family in *M. marinum*, providing an explanation for the essential role of WhiB4 in *M. marinum* virulence. Our study provides new insight into the biological function of WhiB4 and also indicates a different role for WhiB4 in different pathogenic mycobacteria.

Results

***M. marinum* $\Delta whiB4$ is defective in cording formation.** We have isolated a transposon inactivated *whiB4* mutant strain of *M. marinum*, which was identified in a screen of mutation library for altered colony morphology. The transposon was inserted at a TA site in the promoter region of *whiB4*, 28 bp upstream of the start codon. Since the transposon inactivated transcription of *whiB4*, we referred this strain $\Delta whiB4$. The mutant displayed aberrant crustose colony morphology as compared to the smoother, glossier morphology of the WT strain (Fig. 1A). Previously, we showed that changes of colony morphology of *M. marinum* are often associated with defective biosynthesis of lipooligosaccharides (LOSs) in the cell wall³³. Consistently, a recent study found that deletion of *whiB4* from *M. marinum* E11 strain resulted in rough colony morphology and diminished levels of LOSs³⁴.

Importantly, we found that *M. marinum* $\Delta whiB4$ is defective in cording formation (Fig. 1B), a phenotype not previously described. When grown in liquid media, the WT strain formed serpentine cords that are characteristic of pathogenic mycobacteria, including *M. marinum*. This phenotype was abolished in $\Delta whiB4$. Complementation of $\Delta whiB4$ with a plasmid pNBV1 containing intact *whiB4* gene restored both the colony morphology and cording formation, indicating that inactivation of *whiB4* gene is responsible for these phenotypes.

***M. marinum* $\Delta whiB4$ is defective in intracellular replication.** Cording morphology was first described by Robert Koch, and previous studies have revealed a correlation between cording and virulence^{35–37}. The defective cording phenotype of the $\Delta whiB4$ strain prompted us to examine if inactivation of *whiB4* affects

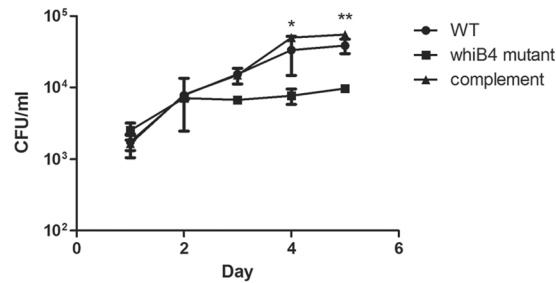


Figure 2. *M. marinum* $\Delta whiB4$ exhibited reduced intracellular survival. RAW macrophage cells were infected with *M. marinum* strains and the intracellular survival of each strain was determined by determining bacterial CFU at indicated time points post infection. Data are plotted as mean \pm SEM ($n = 3$), and are representative of two independent experiments. The CFU of $\Delta whiB4$ at day 4 and 5 were significantly lower than that of WT (* $p < 0.05$; ** $p < 0.01$), or the complemented strain ($p < 0.001$) as analyzed by two-way ANOVA.

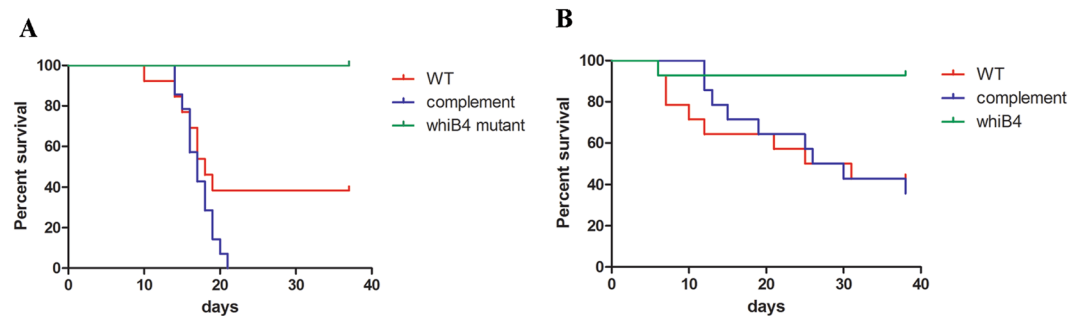


Figure 3. *M. marinum* $\Delta whiB4$ was attenuated in zebrafish. (A) Zebrafish (15 per group) were infected with indicated strain at 10^4 CFU bacteria per fish, and were monitored for mortalities over a 40-day period. The survival curves were plotted using the Kaplan-Meier method and differences between curves were analyzed using the log-rank test. All fish infected with $\Delta whiB4$ or PBS survived. The survival curve of the $\Delta whiB4$ group was significantly different to those of WT ($p = 0.0005$) or the complemented strain ($p < 0.0001$). A similar result was found in another independent experiment (B).

virulence of *M. marinum*. We first performed macrophage infection experiments to determine the intracellular growth of $\Delta whiB4$. RAW macrophages were infected with WT, $\Delta whiB4$, or the complemented strain of *M. marinum* and the intracellular growth of bacteria were assayed by enumerating the bacterial number (colony forming unit, CFU) at different time points post-infection. Both WT and the complemented strain grew substantially inside macrophages. By contrast, the growth of $\Delta whiB4$ inside macrophages was significantly reduced (Fig. 2). All three strains exhibited equivalent growth in 7H9 media. Taken together, these data suggest that *whiB4* is required for the intracellular replication of *M. marinum*.

***M. marinum* $\Delta whiB4$ is attenuated in zebrafish.** Zebrafish are a natural host of *M. marinum* and have been widely used as a laboratory model for studying *M. marinum* infection, which manifests both acute disseminated disease and chronic persistent infection³⁸. To assess the role of *whiB4* in virulence, adult zebrafish (15 per group) were infected with WT, $\Delta whiB4$, or the complemented strain of *M. marinum*, and were monitored for survival. The median survival time of fish infected with WT and the complemented strain was 18- and 17-days, respectively (Fig. 3A). By contrast, all zebrafish infected with $\Delta whiB4$ survived during the 40-day course of experiments. Log-rank analysis indicated that the survival curve of the fish group infected with $\Delta whiB4$ was significantly different to the groups of fish infected with WT ($p = 0.0005$) or the complemented strain ($p < 0.0001$). There was no significant difference between the survival curves of fish infected with WT and the complemented strain ($p = 0.08$).

Fish infected with $\Delta whiB4$ also exhibited significantly less severe pathology than those infected with WT or the complemented strain. At week 3 post infection, fish infected with WT or the complemented strain exhibited severe tissue damage in multiple organs, including livers, kidneys and intestines. Numerous bacteria were detected in these organs. As an example, fish liver infected with WT *M. marinum* is shown (Fig. 4A,B). The bacteria had spread extensively and there were no organized granulomas observed at this stage of infection. By contrast, there was no discernible pathology in livers or other organs of fish infected with $\Delta whiB4$ at the same period (week 3 post infection) (Fig. 4C). A few granulomas were detected in the livers of fish infected with $\Delta whiB4$ at week 4 post infection (Fig. 4D,E). The granulomas were well organized with bacteria contained inside (Fig. 4F), indicating a successful control of bacterial spread by the host. Taken together, these results demonstrated that inactivation of *whiB4* in *M. marinum* greatly reduced its virulence in zebrafish.

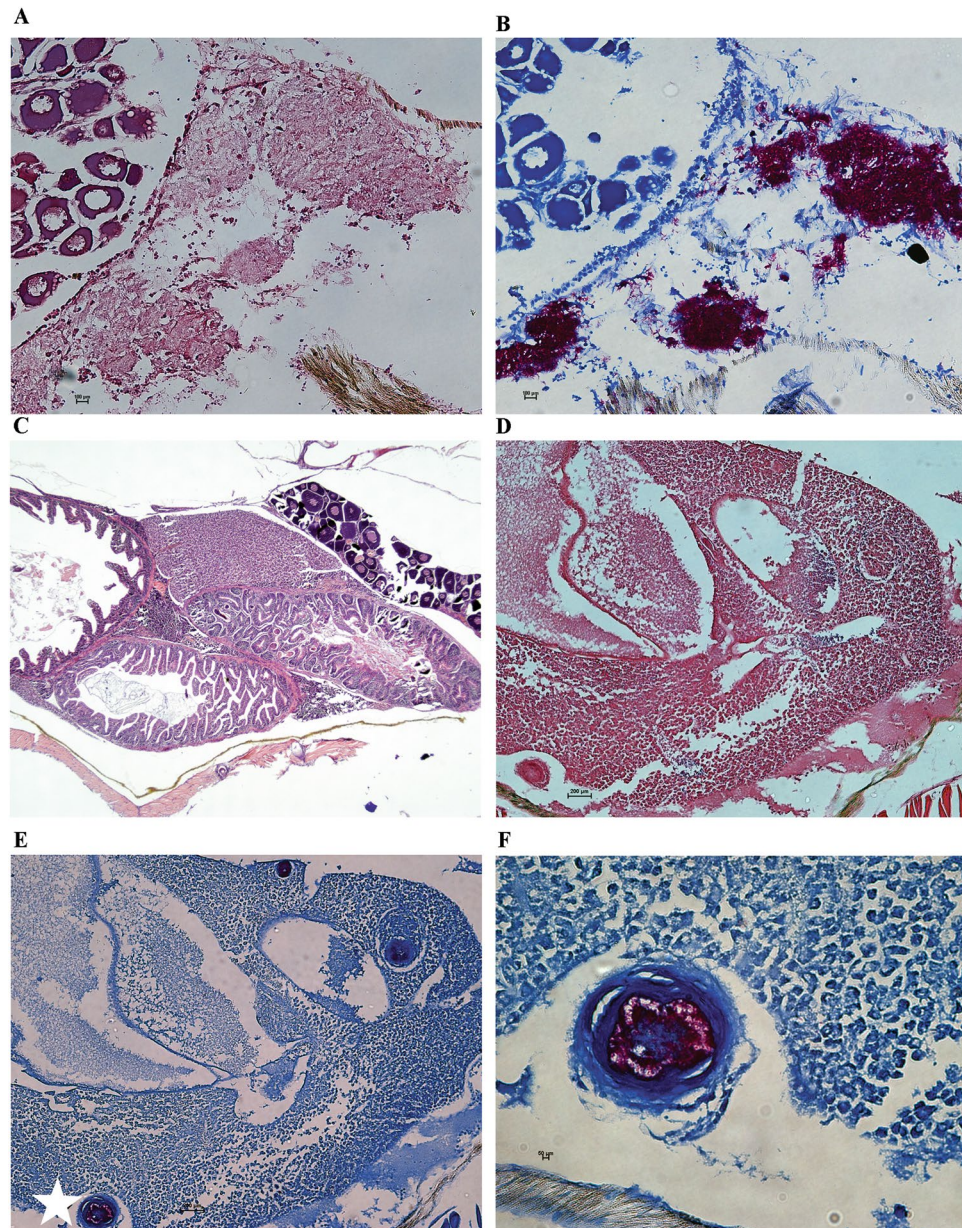


Figure 4. *M. marinum* $\Delta whiB4$ caused less severe pathology. Histological analysis of fish tissues from the experiment described in Fig. 3A. (A,B) Liver sections of fish infected with WT *M. marinum* at 3 weeks post-challenge. Samples were analyzed with hematoxylin and eosin (H&E) staining (A) and Ziehl-Neelsen method of acid fast staining (B). Both are 20 \times magnification. (C) H&E staining of liver sections of fish infected with $\Delta whiB4$ at 3 weeks post-infection (4 \times magnification). (D,E) Liver sections of fish infected with $\Delta whiB4$ at 3 weeks post-infection (10 \times magnification) and stained with H&E (D) or Ziehl-Neelsen (E). The well-organized granuloma observed in (E, white star) was zoomed in for better visualization (F), 40 \times magnification.

A similar result was found in another independent experiment (Fig. 3B). The median survival time of fish infected with WT or the complemented strains was both 28 days, whereas 13 out of 14 fish infected with $\Delta whiB4$ survived. Consistently, there was a significant difference between the survival curve of $\Delta whiB4$ infected group and groups infected with WT ($p = 0.007$, log-rank analysis) or the complemented strain ($p = 0.003$). There is no significant difference between the survival curves of fish infected with WT and the complemented strain ($p = 0.98$).

WhiB4 regulates the *pelpe* gene family in *M. marinum*. To understand the potential mechanisms of WhiB4-mediated virulence in *M. marinum*, we performed RNA-seq analysis and compared global transcripts between WT and $\Delta whiB4$ grown in 7H9 culture media. Two biologically independent experiments were performed and there was excellent agreement between the experiments. The correlation coefficient of total FPKM (Fragments Per Kilobase of exon model per Million mapped reads) counts between two biologically independent

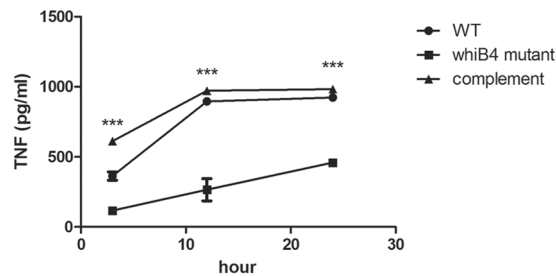


Figure 5. *M. marinum* $\Delta whiB4$ was less potent inducing TNF. RAW 264.7 macrophage cells were infected with mycobacteria at MOI of 10. At 3, 12, and 24-hr post infection, cell-free supernatants were collected and the levels of TNF were determined by enzyme-linked immunosorbent assay (ELISA). Data are plotted as mean \pm SEM ($n = 3$), and are representative of two independent experiments. The level of TNF secretion induced by $\Delta whiB4$ was significantly lower than that of WT or the complemented strain ($***p < 0.001$; two-way ANOVA).

samples of WT was 0.99. Similarly, the correlation coefficient between two independent samples of $\Delta whiB4$ was 0.98. Averaged FPKM data of each strain from the two independent experiments were used for comparative gene expression analysis. RNA-seq revealed that the expression levels of 387 genes were altered by ≥ 2 fold in $\Delta whiB4$, and the differences were statistically significant (Supplementary Table S1). Of these, 257 genes were downregulated and 130 genes were upregulated in $\Delta whiB4$ compared to WT. The *whiB4* transcript in $\Delta whiB4$ was 15-fold less than in WT, confirming the inactivation of *whiB4* in the transposon insertion mutant. A number of genes (21 genes) in the biosynthetic locus of LOSs (MMAR_2307–MMAR_2354) were downregulated³³, providing additional genetic evidence for the diminished LOSs in *M. marinum* $\Delta whiB4$ observed in a previous study³⁴.

Remarkably, 98 genes that are differentially expressed in $\Delta whiB4$ encode the PE/PPE family proteins (Supplementary Table S2), suggesting that the *pe/ppe* gene family are significantly enriched in the WhiB4 regulon (the hypergeometric $p = 1.8e^{-45}$). PE family proteins are characterized by the presence of a proline-glutamic acid (PE) motif at positions 8 and 9 within a highly conserved N-terminal domain consisting of around 110 amino acids. Similarly, PPE family proteins contain a proline-proline-glutamic acid (PPE) at positions 7–9 in a highly conserved N-terminal domain of approximately 180 amino acids³⁹. The C-terminal domains of both PE and PPE family proteins are highly variable in both size and sequence, and often contain repetitive sequences. Based on the C-terminal sequence, PE/PPE family proteins can be divided into subfamilies^{39,40}. PE-PGRS is the largest PE subfamily and is characterized by a C-terminal domain that contains multiple tandem repeats of a glycine-glycine-alanine (GGA) or a glycine-glycine-asparagine (GGN) motif. Likewise, the PPE-MPTR subfamily contains major polymorphic tandem repeats at the C-terminal. The *M. marinum* genome contains 281 *pe/ppe* genes, which encodes 27 PE, 148 PE-PGRS, and 106 PPE proteins⁴¹. About 25.3% (98 out of 387) of genes whose expression changed in $\Delta whiB4$ are *pe/ppe* genes, which is ~ 5 fold of the percentage (5.2%, 281 out of 5424) of total *pe/ppe* genes present in the *M. marinum* genome, suggesting that WhiB4 specifically regulates the expression of these multigenic families of genes.

WhiB4 primarily acts as a positively regulator of the *pe/ppe* gene family. Of the 98 *pe/ppe* genes differentially expressed in $\Delta whiB4$, 79 were downregulated and 19 were upregulated. Moreover, WhiB4 predominantly regulates the *pe-pgrs* gene subfamily. Sixty-seven *pe-pgrs* genes were differentially expressed in $\Delta whiB4$, including 55 downregulated and 12 upregulated. Taken together, these results indicated that WhiB4 is a major regulator of the *pe/ppe* gene family, especially *pe-pgrs* genes.

Accumulated evidence suggests that the PE/PPE family proteins play a critical role in mycobacterial pathogenesis (reviewed in ref. 42). PE-PGRS30 and PE-PGRS62 were shown to inhibit macrophage phagosomal maturation^{43,44}, and PE-PGRS30 was required for the full virulence of *M. tb* in mice⁴⁴. Interestingly, the expression levels of *pe-pgrs62* and *pe-pgrs30* were reduced 18.6- and 3.3-fold, respectively in $\Delta whiB4$ compared to that in WT, which provides a partial explanation for the reduced intracellular replication and diminished virulence of $\Delta whiB4$.

Many PE/PPE proteins have been shown to modulate host immune responses (reviewed in refs 45 and 46). Excessive proinflammatory responses can lead to deleterious effects for the host. In active TB, emerging evidence suggests that a high level of TNF could accelerate the disease progression^{47–52}. Interestingly, we found that $\Delta whiB4$ induced lower levels of pro-inflammatory response, decreasing the production of TNF by macrophages at early time points (Fig. 5). The dampening immune response by $\Delta whiB4$ is consistent with the reduced histopathology in fish infected with this mutant (Fig. 4).

WhiB4 differentially regulates antioxidant systems in *M. marinum*. The expression of several antioxidant enzymes were upregulated in $\Delta whiB4$ of *M. marinum* (Supplementary Table S1), including superoxide dismutase SodA, catalase/peroxidase KatE, and peroxidoxin BcpB. This is consistent with a previous finding that several genes associated with antioxidant systems were upregulated in *M. tb* $\Delta whiB4$, including *ahpC* and *ahpD* encoding alkyl hydroperoxide reductases³⁰. The finding that WhiB4 negatively regulates the expression of antioxidant enzymes is somewhat difficult to reconcile with the positive requirement of WhiB4 in *M. marinum* virulence. Interestingly, a recently identified mycothiol-dependent oxidoreductase (mycothiol-1, Mrx1, MMAR_1363)⁵³ was downregulated in $\Delta whiB4$ of *M. marinum* (Supplementary Table S1). Mycobacteria do not

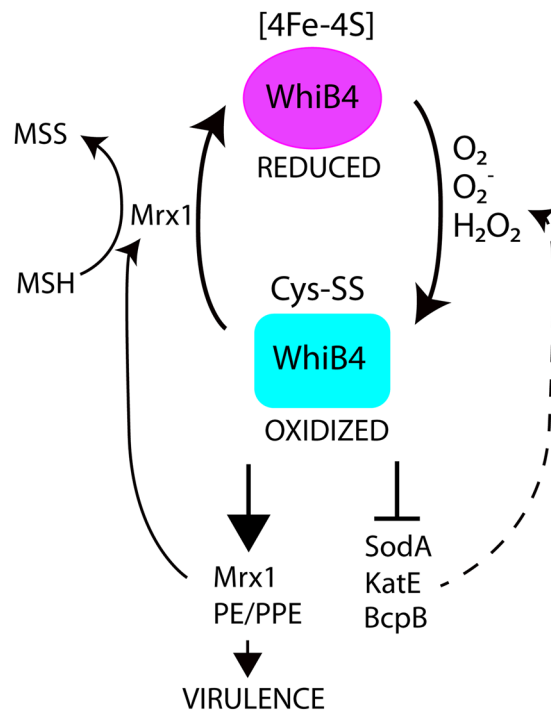


Figure 6. A proposed model for the physiological role of WhiB4 in *M. marinum*. See text for explanations.

contain glutathione but produce functionally equivalent mycothiol that is present in millimolar concentrations. Mycothiol is therefore a major redox buffer against oxidative stress⁵⁴. The ability of Mrx1 to couple the reducing power of mycothiol to reduce oxidized cellular proteins was recently demonstrated in AhpE of *M. tb*⁵⁵. It is conceivable that Mrx-1 may use mycothiol as a cofactor and reduce WhiB4. Based on these observations, we propose a model to explain the physiological function of WhiB4 in *M. marinum* (Fig. 6). In the presence of oxygen, the [4Fe-4S] cluster in WhiB4 is quickly destroyed and oxidized to disulfide bonds, which activates the DNA binding activity of WhiB4³⁰. Oxidized WhiB4 would activate the transcription of *pe/ppe* genes, which would facilitate bacterial intracellular replication and survival in the host. At the same time, oxidized WhiB4 would repress the expression of antioxidant enzymes and consequently maintain a relatively oxidizing environment. As a result, more oxidized WhiB4 would be generated. As these processes continue, the accumulation of Mrx1, which is dependent on oxidized WhiB4, would form a negative feedback regulation. Mrx1 would use mycothiol as a cofactor to reduce oxidized WhiB4 and eventually lead to the formation of [4Fe-4S]-containing WhiB4. The holo-WhiB4 would then lose its DNA binding activity and antioxidant enzymes (SodA, KatE) would be de-repressed, which would reduce the oxidative stress and return the system to its original state.

Discussion

In this study, we found that WhiB4 plays similar yet distinct roles in *M. marinum* and *M. tb*. WhiB4 modulates the oxidative stress response in *M. marinum* by repressing the expression of antioxidant enzymes including SodA and KatE, which is similar to a previous finding of WhiB4 in *M. tb*³⁰. However, WhiB4 is required for virulence in *M. marinum*. *M. marinum* Δ *whiB4* exhibited impaired replication in macrophages and diminished virulence in zebrafish. By contrast, *M. tb* Δ *whiB4* exhibited enhanced replication in macrophages and increased virulence in lungs of guinea pigs, although its dissemination to spleens was reduced³⁰. Sequence analysis reveals that WhiB4 from *M. tb* and *M. marinum* shares an overall 83% sequence identity (amino acids), and 100% identity in the Cys-X14-22-Cys-X2-Cys-X5-Cys motif. Therefore, the small sequence variation between WhiB4 of these two organisms cannot account for the difference in physiological function. Our finding that WhiB4 plays a different role in virulence of different pathogenic mycobacteria is not without precedent; a previous study found that deletion of *whiB3*, another member of the *whiB* gene family, in two nearly identical organisms *M. tb* and *M. bovis*, resulted in different virulence phenotypes in guinea pigs⁵⁶. Taken together, these studies suggest that although the WhiB family proteins are highly conserved among mycobacteria, they could play distinct physiological roles in specific mycobacterial species.

The different role of WhiB4 in virulence of *M. marinum* and *M. tb* likely arises from the difference of the WhiB4 regulon in these organisms. Strikingly, we found that a large number of *pe/ppe* genes (98) were regulated by WhiB4 in *M. marinum*. By contrast, only 3 *pe/ppe* genes (*pe35*, *ppe19*, *ppe68*) were regulated by WhiB4 in *M. tb*³⁰. The *pe/ppe* gene family represent one of the most intriguing aspects of mycobacterial genomes. The *pe/ppe* gene family was initially discovered in *M. tb*, comprising approximately 10% of *M. tb* genome³⁹. Since then, *pe/ppe* genes have been identified in many pathogenic mycobacteria including the *M. tb* complex (MTBC), including *M. tb*, *M. africanum*, *M. canettii*, *M. bovis* and *M. microti*), *M. leprae*, *M. avium*, *M. marinum* and *M. ulcerans*. *M. marinum* has the most extensive repertoire of *pe/ppe* genes reported (281 in *M. marinum* compared with 169 in

M. tb and 115 in *M. ulcerans*)⁴¹. Nonpathogenic mycobacteria have few *pe/ppe* genes (e.g., 10 in *M. smegmatis*). Although the origin of the *pe/ppe* gene family is unknown, the ancestral *pe/ppe* genes are associated with the type VII secretion system, including the virulence determinant ESX-1 secretion system⁴⁰. Phylogenetic analysis suggests a slow initial expansion of *pe/ppe* genes, which accompanies the duplication of *esx* gene clusters, and that the *pe-pgrs* and *ppe-mptr* gene subfamilies are the most recent and expansive sublineage, sublineage V⁴⁰. Interestingly, sublineage V is only present in the MTBC and closely related *M. marinum* and *M. ulcerans*, suggesting that it may play a role in host-pathogen interactions⁴⁰. Within this context, it is striking to see that majorities of *pe/ppe* genes (79 out of 98) regulated by WhiB4 belong to sublineage V, including 67 *pe-pgrs* and 12 *ppe-mptr* genes (Supplementary Table S2). Furthermore, most of these *pe-pgrs* and *ppe-mptr* genes (55 out of 67) were positively regulated by WhiB4, including those that have been experimentally demonstrated to be critical for *M. marinum* virulence (e.g., *pe-pgrs30*, *pe-pgrs62*, and *ppe38*)^{43,44,57}. Taken together, these data provide a plausible explanation for the essential role of WhiB4 in *M. marinum* virulence. We should point out that since the complemented strain was not included in the RNA-seq analysis, we cannot fully exclude the possibility that some differences in the gene expression could be linked to possible WhiB4-independent genetic modification(s) than may occur in a given selected transposon mutant.

At present it is not clear why WhiB4 did not perform a similar function in *M. tb* as in *M. marinum*, considering that a large number of *pe-pgrs* and *ppe-mptr* genes are also present in *M. tb* genome. Of the 98 *pe/ppe* genes regulated by WhiB4 in *M. marinum*, 59 genes have an ortholog in *M. tb*, 20 genes are restricted to *M. marinum* and 15 are present in both *M. marinum* and *M. ulcerans* (Supplementary Table S2). Previously, it was found that WhiB4 binds DNA non-specifically but has a preference for GC rich DNA sequences³⁰. This is consistent with our finding that a large number (67) of *pe-pgrs* genes were regulated by WhiB4. The high glycine content (up to 50%) of PE-PGRS proteins makes their corresponding genes GC rich. On the other hand, the fact that there are still 81 *pe-pgrs* genes in *M. marinum* that are apparently not regulated by WhiB4 implies that cofactors may be required for WhiB4 mediated regulation of *pe/ppe* genes, and that these factors may be missing in *M. tb*. Several transcriptional regulators including sigma factors SigB and SigD, two-component systems PhoPR and MprAB have been shown to regulate specific *pe/ppe* genes (reviewed in ref. 42). However, compared to WhiB4, the number of *pe/ppe* genes controlled by these regulators is much lower. Whether these regulators interact with WhiB4 or whether additional factors are required for WhiB4 mediated regulation of *pe/ppe* genes will be addressed in future studies.

Methods

Bacterial strains and culture conditions. *M. marinum* strain 1218R (ATCC 927) and the transposon insertion mutant *M. marinum* Δ *whiB4* were routinely grown in Middlebrook 7H9 broth or 7H11 agar ((Difco™) supplemented with 10% OADC (oleic acid-albumin-dextrose-catalase) and 0.05% Tween80. *Escherichia coli* strain DH5 α was used for routine manipulation and propagation of plasmid DNA. *E. coli* DH5 α was grown in LB broth or agar. Antibiotics were added as required: kanamycin, 50 μ g/ml for *E. coli* and 25 μ g/ml for mycobacteria; hygromycin, 150 μ g/ml for *E. coli* and 75 μ g/ml for mycobacteria.

Isolation of transposon insertion *whiB4* mutant. The *mariner*-based transposon system was used to generate a transposon insertion mutant library of *M. marinum* 1218R as described previously^{33,58}. Colonies with unusual morphology were identified by visual inspection. The method used to localize and identify the transposon-disrupted gene has been previously described^{33,58}. Briefly, total chromosomal DNA of the transposon insertion mutant was cleaved with *Bam*HI, then self-ligated with T4 DNA ligase and transformed into competent *E. coli* DH5 α λ *pir*116 cells. Plasmid DNA was isolated from kanamycin resistant *E. coli* transformants and MycoMar-specific primers were used to determine the DNA sequence at the transposon/chromosomal junction.

Molecular cloning. A DNA fragment containing intact *whiB4* gene and 400 bp upstream of its start codon was amplified by PCR using the forward primer 5'-GCTACGCCAGCCCGTTAGGA-3' and the reverse primer 5'-GCTTGAGCAGGGTGAGTGCG-3'. *M. marinum* genomic DNA was used as template. The fragment was ligated into pDRIVE and the resulting construct was isolated and digested with *Bam*HI and *Hin*DIII. The *whiB4* containing fragment was ligated into pNBV1 which contains hygromycin resistance gene to generate pWhiB4. Standard electroporation protocols were used for transformation of pWhiB4 into *M. marinum* Δ *whiB4*. Transformants were selected on Middlebrook 7H11 agar containing hygromycin (75 μ g/ml) and kanamycin (25 μ g/ml). Three randomly selected clones were tested.

Scanning electron microscopy analysis. Mycobacterial cultures (5 ml each) grown to mid-log phase were collected and added with two droplets of heparin sodium. The cultures were then fixed, dehydrated and observed using scanning electron microscope SU5000 (Hitachi, Japan).

Analysis of *M. marinum* intracellular growth. Infection of the marine RAW 264.7 macrophage cells and enumeration of intracellular *M. marinum* were performed as previously described⁵⁹. Briefly, RAW cells were seeded into 24-well plates at a density of 10⁵ cells per well in DMEM medium supplemented with 10% FBS for 24 hr. RAW cells were then infected with *M. marinum* at multiplicity of infection (MOI) of 10 for 2 hr at 30 °C in 5% CO₂. Wells were washed twice with sterile PBS and extracellular *M. marinum* were killed by incubating the tissue culture with 200 μ g/ml gentamycin for 2 hr. Cells were again washed twice with PBS and subsequently, incubated at 32 °C in 5% CO₂ in fresh media with 20 μ g/ml gentamycin. At different time points, the infected macrophage monolayers (3 wells per strain) were washed 3 times with fresh media and then lysed with 0.1 ml of 1% Triton X-100 to release intracellular mycobacteria. The number of intracellular mycobacteria was enumerated by plating appropriate dilutions on Middlebrook 7H10 agar plates containing appropriate antibiotics. Two independent experiments were performed.

Cytokine assay. RAW 264.7 macrophage cells were infected with mycobacteria at MOI of 10 as described above. At 3, 12, and 24-hr post infection, cell-free supernatants were collected and the levels of TNF were determined by enzyme-linked immunosorbent assay (ELISA) according to the manufacturer's instructions (Dakew, China). Two independent experiments were performed.

Ethics statement. All of the animal procedures were approved by the local animal care committees at Fudan University. All methods were performed in accordance with the relevant guidelines and regulations.

Zebrafish infection. Zebrafish infection with *M. marinum* was performed as previously described⁵⁷. Briefly, adult zebrafish (15 per group) were infected by intraperitoneal injection with a dosage of 10⁴ CFU bacteria per fish or PBS as the negative control and monitored for their survivals.

For histopathological study, two fish were sacrificed at the indicated time. Fish were fixed and then dehydrated with ethanol. After paraffin embedding and sectioning, serial paraffin sections were prepared and subjected to hematoxylin and eosin (H&E) staining and Ziehl-Neelsen acid fast staining. Sections were examined under a Nikon Eclipse Ni-U Microscope (Japan), and images were collected with a digital camera. Two independent zebrafish infection and survival experiments were performed.

RNA extraction and illumine sequencing. *M. marinum* WT and Δ *whiB4* each were grown in 5 duplicate cultures, each containing 5 ml 7H9 media to OD600 ~1.5. 1-ml cells from 5 duplicate cultures were collected and mixed together. The bacterial cultures were pelleted and washed twice with PBS. Total RNA (1–3 μ g) was isolated using the RNeasy Mini Kit (Qiagen) and purified using the RNAClean XP Kit (Beckman Coulter) and RNase-Free DNase Set (Qiagen) according to the manufacturer's instructions. Purified RNA was used to construct cDNA library according to the TruSeq Stranded RNA LT Guide from Illumina. The concentration and size distribution of cDNA library was analyzed by Agilent 2100 Bioanalyzer and the average library size was approximately 350 bp. High-throughput sequencing was carried out on an Illumina HiSeq 2500 system according to the manufacturer's instructions (Illumina HiSeq 2500 User Guide) and 150-bp paired-end reads were obtained. The raw reads were filtered by Seqtk and then mapped to the *M. marinum* M strain reference sequence (GenBank GCA_000018345.1) using Bowtie2 (version: 2–2.0.5)⁶⁰. Counting of reads per gene was performed using HTSeq followed by TMM (trimmed mean of M-values) normalization^{61,62}. Differentially expressed genes were defined as those with a false discovery rate <0.05 and fold-change >2 using the edgeR software⁶³. Two independent experiments were performed.

Statistical analysis. Two-Way ANOVA were performed on data of 3 groups at multiple time points (macrophage infection experiments). Zebrafish survival curves were plotted using the Kaplan-Meier method and differences between curves were analyzed using the log-rank test.

References

- Stewart, G. R., Robertson, B. D. & Young, D. B. Tuberculosis: a problem with persistence. *Nat Rev Microbiol* **1**, 97–105, doi:10.1038/nrmicro749 (2003).
- North, R. J. & Jung, Y. J. Immunity to tuberculosis. *Annu Rev Immunol* **22**, 599–623, doi:10.1146/annurev.immunol.22.012703.104635 (2004).
- Corbett, E. L. *et al.* Epidemiology of tuberculosis in a high HIV prevalence population provided with enhanced diagnosis of symptomatic disease. *PLoS Med* **4**, e22, doi:06-PLME-RA-0369R2 (2007).
- Corbett, E. L. *et al.* The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Arch Intern Med* **163**, 1009–1021, doi:10.1001/archinte.163.9.1009 (2003).
- Barry, C. E. 3rd *et al.* The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat Rev Microbiol* **7**, 845–855, doi:10.1038/nrmicro2236 (2009).
- Weiss, G. & Schaible, U. E. Macrophage defense mechanisms against intracellular bacteria. *Immunol Rev* **264**, 182–203, doi:10.1111/imr.12266 (2015).
- Flynn, J. L., Chan, J. & Lin, P. L. Macrophages and control of granulomatous inflammation in tuberculosis. *Mucosal Immunol* **4**, 271–278, doi:mi201114 (2011).
- Russell, D. G. *et al.* Mycobacterium tuberculosis wears what it eats. *Cell Host Microbe* **8**, 68–76, doi:S1931-3128(10)00180-0 (2010).
- Rustad, T. R., Sherrid, A. M., Minch, K. J. & Sherman, D. R. Hypoxia: a window into Mycobacterium tuberculosis latency. *Cell Microbiol* **11**, 1151–1159, doi:CMI1325 (2009).
- Chao, M. C. & Rubin, E. J. Letting sleeping dogs lie: does dormancy play a role in tuberculosis? *Annu Rev Microbiol* **64**, 293–311, doi:10.1146/annurev.micro.112408.134043 (2010).
- Via, L. E. *et al.* Tuberculous granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates. *Infect Immun* **76**, 2333–2340, doi:IAI.01515-07 (2008).
- Gomez, J. E. & McKinney, J. D. M. tuberculosis persistence, latency, and drug tolerance. *Tuberculosis (Edinb)* **84**, 29–44, doi:S1472979203000866 (2004).
- Boshoff, H. I. & Barry, C. E. 3rd Tuberculosis - metabolism and respiration in the absence of growth. *Nat Rev Microbiol* **3**, 70–80, doi:nrmicro1065 (2005).
- Kesavan, A. K., Brooks, M., Tufariello, J., Chan, J. & Manabe, Y. C. Tuberculosis genes expressed during persistence and reactivation in the resistant rabbit model. *Tuberculosis (Edinb)* **89**, 17–21, doi:S1472-9792(08)00095-4 (2009).
- Wayne, L. G. & Sohaskey, C. D. Nonreplicating persistence of mycobacterium tuberculosis. *Annu Rev Microbiol* **55**, 139–163, doi:10.1146/annurev.micro.55.1.139 (2001).
- Wayne, L. G. Synchronized replication of Mycobacterium tuberculosis. *Infect Immun* **17**, 528–530 (1977).
- Wayne, L. G. & Hayes, L. G. An *in vitro* model for sequential study of shiftdown of Mycobacterium tuberculosis through two stages of nonreplicating persistence. *Infect Immun* **64**, 2062–2069 (1996).
- Boon, C. & Dick, T. Mycobacterium bovis BCG response regulator essential for hypoxic dormancy. *J Bacteriol* **184**, 6760–6767 (2002).
- Sherman, D. R. *et al.* Regulation of the Mycobacterium tuberculosis hypoxic response gene encoding alpha-crystallin. *Proc Natl Acad Sci USA* **98**, 7534–7539, doi:10.1073/pnas.121172498 (2001).
- Park, H. D. *et al.* Rv3133c/dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis. *Mol Microbiol* **48**, 833–843, doi:3474 (2003).

21. Galagan, J. E. *et al.* The Mycobacterium tuberculosis regulatory network and hypoxia. *Nature* **499**, 178–183, doi:10.1038/nature12337 (2013).
22. Canetti, G. Present aspects of bacterial resistance in tuberculosis. *Am Rev Respir Dis* **92**, 687–703, doi:10.1164/arrd.1965.92.5.687 (1965).
23. Soliveri, J. A., Gomez, J., Bishai, W. R. & Chater, K. F. Multiple paralogous genes related to the Streptomyces coelicolor developmental regulatory gene whiB are present in Streptomyces and other actinomycetes. *Microbiology* **146**(Pt 2), 333–343, doi:10.1099/0021287-146-2-333 (2000).
24. Wu, M. L. *et al.* Developmental transcriptome of resting cell formation in Mycobacterium smegmatis. *BMC Genomics* **17**, 837, doi:10.1186/s12864-016-3190-4 (2016).
25. Larsson, C. *et al.* Gene expression of Mycobacterium tuberculosis putative transcription factors whiB1–7 in redox environments. *PLoS ONE* **7**, e37516, doi:10.1371/journal.pone.0037516 (2012).
26. Ma, S. *et al.* Integrated Modeling of Gene Regulatory and Metabolic Networks in Mycobacterium tuberculosis. *PLoS Comput Biol* **11**, e1004543, doi:10.1371/journal.pcbi.1004543 (2015).
27. Humpel, A., Gebhard, S., Cook, G. M. & Berney, M. The SigF regulon in Mycobacterium smegmatis reveals roles in adaptation to stationary phase, heat, and oxidative stress. *J Bacteriol* **192**, 2491–2502, doi:10.1128/JB.00035-10 (2010).
28. Alam, M. S., Garg, S. K. & Agrawal, P. Studies on structural and functional divergence among seven WhiB proteins of Mycobacterium tuberculosis H37Rv. *FEBS J* **276**, 76–93, doi:10.1111/j.1742-4658.2008.06755.x (2009).
29. Alam, M. S., Garg, S. K. & Agrawal, P. Molecular function of WhiB4/Rv3681c of Mycobacterium tuberculosis H37Rv: a [4Fe-4S] cluster co-ordinating protein disulphide reductase. *Mol Microbiol* **63**, 1414–1431, doi:MMI5589 (2007).
30. Chawla, M. *et al.* Mycobacterium tuberculosis WhiB4 regulates oxidative stress response to modulate survival and dissemination *in vivo*. *Mol Microbiol* **85**, 1148–1165, doi:10.1111/j.1365-2958.2012.08165.x (2012).
31. Smith, L. J. *et al.* Mycobacterium tuberculosis WhiB1 is an essential DNA-binding protein with a nitric oxide-sensitive iron-sulfur cluster. *Biochem J* **432**, 417–427, doi:10.1042/BJ20101440 (2010).
32. Singh, A. *et al.* Mycobacterium tuberculosis WhiB3 maintains redox homeostasis by regulating virulence lipid anabolism to modulate macrophage response. *PLoS Pathog* **5**, e1000545, doi:10.1371/journal.ppat.1000545 (2009).
33. Ren, H. *et al.* Identification of the lipooligosaccharide biosynthetic gene cluster from Mycobacterium marinum. *Mol Microbiol* **63**, 1345–1359, doi:MMI5603 (2007).
34. van der Woude, A. D. *et al.* Unexpected link between lipooligosaccharide biosynthesis and surface protein release in Mycobacterium marinum. *J Biol Chem* **287**, 20417–20429, doi:10.1074/jbc.M111.336461 (2012).
35. Middlebrook, W. R. & Phillips, H. The action of formaldehyde on the cystine disulphide linkages of wool: 2. The conversion of subfraction A of the combined cystine into combined lanthionine and djenkolic acid and subfraction B into combined thiazolidine-4-carboxylic acid. *Biochem J* **41**, 218–223 (1947).
36. Koch, R. Classics in infectious diseases. The etiology of tuberculosis: Robert Koch. Berlin, Germany 1882. *Rev Infect Dis* **4**, 1270–1274 (1982).
37. Darzins, E. & Fahr, G. Cord-forming property, lethality and pathogenicity of Mycobacteria. *Dis Chest* **30**, 642–648 (1956).
38. Tobin, D. M. & Ramakrishnan, L. Comparative pathogenesis of Mycobacterium marinum and Mycobacterium tuberculosis. *Cell Microbiol* **10**, 1027–1039, doi:CMI1133 (2008).
39. Cole, S. T. *et al.* Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **393**, 537–544, doi:10.1038/31159 (1998).
40. Gey van Pittius, N. C. *et al.* Evolution and expansion of the Mycobacterium tuberculosis PE and PPE multigene families and their association with the duplication of the ESAT-6 (esx) gene cluster regions. *BMC Evol Biol* **6**, 95, doi:1471-2148-6-95 (2006).
41. Stinear, T. P. *et al.* Insights from the complete genome sequence of Mycobacterium marinum on the evolution of Mycobacterium tuberculosis. *Genome Res* **18**, 729–741, doi:10.1101/gr.075069.107 (2008).
42. Fishbein, S., van Wyk, N., Warren, R. M. & Sampson, S. L. Phylogeny to function: PE/PPE protein evolution and impact on Mycobacterium tuberculosis pathogenicity. *Mol Microbiol* **96**, 901–916, doi:10.1111/mmi.12981 (2015).
43. Huang, Y. *et al.* Phagolysosome maturation of macrophages was reduced by PE_PGRS 62 protein expressing in Mycobacterium smegmatis and induced in IFN-gamma priming. *Vet Microbiol* **160**, 117–125, doi:10.1016/j.vetmic.2012.05.011 (2012).
44. Iantomasi, R. *et al.* PE_PGRS30 is required for the full virulence of Mycobacterium tuberculosis. *Cell Microbiol* **14**, 356–367, doi:10.1111/j.1462-5822.2011.01721.x (2012).
45. Mukhopadhyay, S. & Balaji, K. N. The PE and PPE proteins of Mycobacterium tuberculosis. *Tuberculosis (Edinb)* **91**, 441–447, doi:10.1016/j.tube.2011.04.004 (2011).
46. Sampson, S. L. Mycobacterial PE/PPE proteins at the host-pathogen interface. *Clin Dev Immunol* **2011**, 497203, doi:10.1155/2011/497203 (2011).
47. Bekker, L. G. *et al.* Immunopathologic effects of tumor necrosis factor alpha in murine mycobacterial infection are dose dependent. *Infect Immun* **68**, 6954–6961 (2000).
48. Blackmore, T. K., Manning, L., Taylor, W. J. & Wallis, R. S. Therapeutic use of infliximab in tuberculosis to control severe paradoxical reaction of the brain and lymph nodes. *Clin Infect Dis* **47**, e83–85, doi:10.1086/592695 (2008).
49. Thwaites, G. E. *et al.* Dexamethasone for the treatment of tuberculous meningitis in adolescents and adults. *N Engl J Med* **351**, 1741–1751, doi:351/17/1741 (2004).
50. Goldstone, R. M., Goonesekera, S. D., Bloom, B. R. & Sampson, S. L. The transcriptional regulator Rv0485 modulates the expression of a pe and ppe gene pair and is required for Mycobacterium tuberculosis virulence. *Infect Immun* **77**, 4654–4667, doi:IAI.01495-08 (2009).
51. Tobin, D. M. *et al.* The It4h locus modulates susceptibility to mycobacterial infection in zebrafish and humans. *Cell* **140**, 717–730, doi:S0092-8674(10)00128-5 (2010).
52. Agarwal, N., Lamichhane, G., Gupta, R., Nolan, S. & Bishai, W. R. Cyclic AMP intoxication of macrophages by a Mycobacterium tuberculosis adenylate cyclase. *Nature* **460**, 98–102, doi:nature08123 (2009).
53. Van Laer, K. *et al.* Mycoredoxin-1 is one of the missing links in the oxidative stress defence mechanism of Mycobacteria. *Mol Microbiol* **86**, 787–804, doi:10.1111/mmi.12030 (2012).
54. Newton, G. L. *et al.* Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes. *J Bacteriol* **178**, 1990–1995 (1996).
55. Hugo, M. *et al.* Mycothiol/mycoredoxin 1-dependent reduction of the peroxiredoxin AhpE from Mycobacterium tuberculosis. *J Biol Chem* **289**, 5228–5239, doi:10.1074/jbc.M113.510248 (2014).
56. Steyn, A. J. *et al.* Mycobacterium tuberculosis WhiB3 interacts with RpoV to affect host survival but is dispensable for *in vivo* growth. *Proc Natl Acad Sci USA* **99**, 3147–3152, doi:10.1073/pnas.052705399 (2002).
57. Dong, D. *et al.* PPE38 modulates the innate immune response and is required for Mycobacterium marinum virulence. *Infect Immun* **80**, 43–54, doi:10.1128/IAI.05249-11 (2012).
58. Alexander, D. C., Jones, J. R., Tan, T., Chen, J. M. & Liu, J. PimF, a mannosyltransferase of mycobacteria, is involved in the biosynthesis of phosphatidylinositol mannosides and lipoarabinomannan. *J Biol Chem* **279**, 18824–18833, doi:10.1074/jbc.M400791200 (2004).
59. Tan, T., Lee, W. L., Alexander, D. C., Grinstein, S. & Liu, J. The ESAT-6/CFP-10 secretion system of Mycobacterium marinum modulates phagosome maturation. *Cell Microbiol* **8**, 1417–1429, doi:CMI721 (2006).

60. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).
61. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169, doi:10.1093/bioinformatics/btu638 (2015).
62. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25, doi:10.1186/gb-2010-11-3-r25 (2010).
63. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140, doi:10.1093/bioinformatics/btp616 (2010).

Acknowledgements

This study was supported by grants from the China's 12th Five Year Programs for the prevention and cure of great infectious diseases (No. 2012ZX10003002-002), Shanghai Municipal Natural Science Foundation (No. 16ZR1402800), International cooperation and exchange program of Shanghai Municipal (No. 16430724000) and Shanghai Science and Technology Commission (13DZ2252000), and a grant from Canadian Institutes of Health Research (MOP-15107).

Author Contributions

J.W., H.W.R., Z.H.X., J.J., and Y.C.W. performed experiments. J.W., L.Z., and J.L. wrote the manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-03020-4

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017