



Transcriptome Changes of *Mycobacterium marinum* in the Process of Resuscitation From Hypoxia-Induced Dormancy

Jun Jiang^{1†}, Chen Lin^{1†}, Junli Zhang¹, Yuchen Wang¹, Lifang Shen¹, Kunpeng Yang¹, Wenxuan Xiao¹, Yao Li^{1*}, Lu Zhang^{1,2*} and Jun Liu^{1,3*}

OPEN ACCESS

Edited by:

Tao Huang,
Shanghai Institutes for Biological
Sciences (CAS),
China

Reviewed by:

Li Zhang,
East China Normal University,
China
Yongyong Yang,
Northwestern University,
United States

*Correspondence:

Yao Li
yaoli@fudan.edu.cn
Lu Zhang
zhanglu407@fudan.edu.cn
Jun Liu
jun.liu@utoronto.ca

†These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Bioinformatics and
Computational Biology,
a section of the journal
Frontiers in Genetics

Received: 14 October 2019

Accepted: 11 December 2019

Published: 07 February 2020

Citation:

Jiang J, Lin C, Zhang J, Wang Y,
Shen L, Yang K, Xiao W, Li Y, Zhang L
and Liu J (2020) Transcriptome
Changes of *Mycobacterium marinum*
in the Process of Resuscitation From
Hypoxia-Induced Dormancy.
Front. Genet. 10:1359.
doi: 10.3389/fgene.2019.01359

¹ State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Science, Fudan University, Shanghai, China, ² State Key Laboratory of Genetic Engineering, MOE Engineering Research Center of Gene Technology, School of Life Science, Fudan University, Shanghai, China, ³ Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada

Nearly one-third of the world's population is latently infected with *Mycobacterium tuberculosis* (*M. tb*), which represents a huge disease reservoir for reactivation and a major obstacle for effective control of tuberculosis. During latent infection, *M. tb* is thought to enter nonreplicative dormant states by virtue of its response to hypoxia and nutrient-deprived conditions. Knowledge of the genetic programs used to facilitate entry into and exit from the nonreplicative dormant states remains incomplete. In this study, we examined the transcriptional changes of *Mycobacterium marinum* (*M. marinum*), a pathogenic mycobacterial species closely related to *M. tb*, at different stages of resuscitation from hypoxia-induced dormancy. RNA-seq analyses were performed on *M. marinum* cultures recovered at multiple time points after resuscitation. Differentially expressed genes (DEGs) at each time period were identified and analyzed. Co-expression networks of transcription factors and DEGs in each period were constructed. In addition, we performed a weighted gene co-expression network analysis (WGCNA) on all genes and obtained 12 distinct gene modules. Collectively, these data provided valuable insight into the transcriptome changes of *M. marinum* upon resuscitation as well as gene module function of the bacteria during active metabolism and growth.

Keywords: transcriptional regulation, resuscitation, *M. marinum*, hypoxia, latency

INTRODUCTION

Mycobacterium tuberculosis (*M. tb*), the causative agent of tuberculosis (TB), is the leading cause of death due to an infectious disease globally, with an estimated 10 million new cases and 1.3 million deaths in 2017. There were an additional 300,000 deaths from TB among HIV-positive people. The success of *M. tb* as a leading pathogen is associated with its ability to infect and persist in the host. About 1.7 billion people, 23% of the world's population, are estimated to have a latent TB infection (LTBI), which is asymptomatic but can persist for decades (Stewart et al., 2003; North and Jung, 2004). About 5–10% of LTBI will eventually develop active disease, and host immunosuppression (e.g., HIV coinfection) markedly increases the risk of reactivation (Corbett et al., 2007). LTBI poses a

major challenge to the effective control of TB because of the difficulty in treatment and the fact that LTBI represents a huge disease reservoir.

During LTBI, *M. tb* is thought to enter nonreplicative 'dormant' states by virtue of its lowered or altered metabolism in response to hypoxia, nitrosative stress, and/or nutrient deprivation (Boshoff and Barry, 2005). Accordingly, much research has been focused on environmental conditions and genetic programs that induce bacteriostasis, and the most extensively studied culture condition is hypoxia (Wayne and Sohaskey, 2001; Rustad et al., 2009). It was shown that an immediate bacterial response (2 hr) was the coordinated upregulation of 47 *M. tb* genes under the control of the response regulator (DosR) and two sensor kinases (DosS and DosT), known as the DosR regulon (Sherman et al., 2001; Boon and Dick, 2002; Park et al., 2003). A second set of 230 genes, induced by longer hypoxia exposure (7 days), was also identified (Rustad et al., 2008). These genes, collectively known as the enduring hypoxic response (EHR), were DosR-independent genes (Rustad et al., 2008).

During the reactivation of LTBI, the dormant bacteria are believed to resuscitate and resume active growth and metabolism. A few recent studies have used re-aeration of hypoxic cultures for *in vitro* modeling of reactivation or resuscitation (Veatch and Kaushal, 2018). Several regulatory proteins, such as transcription factor ClgR and sigma factors SigH and SigE, were found to play a role in *M. tb* resuscitation from hypoxia (Mcgillivray et al., 2015; Iona et al., 2016; Veatch et al., 2016).

Despite the progress, knowledge of the genetic programs used to facilitate entry into and exit from the nonreplicative dormant states remains incomplete. In this study, we examined the transcriptional changes of *Mycobacterium marinum* (*M. marinum*) at different stages of resuscitation from hypoxia-induced dormancy. *M. marinum* is a pathogenic *Mycobacterium* and the closest genetic relative of the *M. tb* complex. *M. marinum* is an excellent model through which to understand various aspects of host-pathogen interactions in *M. tb* pathogenesis. For example, *M. marinum* and *M. tb* share many virulence determinants, such as the ESX-1 secretion system (Tobin and Ramakrishnan, 2008) and lipid virulence factors phthiocerol dimycocerosates and phenolic glycolipids (Yu et al., 2012). As such, findings from the current study of *M. marinum* may be applicable to *M. tb*.

RESULT

RNA-Seq Analysis of *M. marinum* Recovered From Hypoxia

Larry Wayne and co-workers were the first to develop an *in vitro* model to mimic the hypoxic environment of the human granuloma (Wayne, 1977; Wayne and Hayes, 1996; Wayne and Sohaskey, 2001). In the Wayne model, a sealed, standing culture is incubated over an extended period while the bacteria deplete the available oxygen. The gradual depletion of oxygen leads to nonreplicating persistence states with a concomitant shift in gene expression and metabolism.

To gain insight into the genetic mechanisms that facilitate the exit of mycobacteria from the nonreplicative state, we grew *M. marinum* under hypoxia for 7 days using the Wayne model and then re-aerated the cultures. At different time points thereafter (0, 0.5, 4, 12, 24, and 48 hr), *M. marinum* cultures were collected and subjected to RNA-seq analysis. The growth curve of the bacteria is shown in **Supplementary Figure 1**. A total of 18 samples were collected (three biological replicates at each time point) and analyzed.

The RNA-seq reads showed a high mapping ratio for all samples (>96%) (**Table 1**), supporting the overall sequencing accuracy. Transcripts of more than 4,900 genes were detected in each sample. We compared the RNA-seq data of cultures recovered at different time points under aerobic conditions. As expected, results showed that the correlation coefficient decreased as the interval between two samples increased (**Figure 1**). This result also suggested that the recovery from hypoxia is a gradual but dynamic process.

Dynamic Changes of Gene Expression at Different Stage of Resuscitation

To analyze transcriptome changes of *M. marinum*, we focused on genes with RPKM ≥ 10 and compared samples from adjacent intervals: between 0.5 and 0 hr, 4 and 0.5 hr, 12 and 4 hr, 24 and 12 hr, as well as 48 and 24 hr. Differentially expressed genes (DEGs) were identified, and this was defined as a fold change greater than 2 and false discovery rate *P* value less than 0.05.

At the earliest time point after resuscitation (0.5 hr), 136 DEGs were detected, of which 71 were upregulated and 65 were downregulated. Between 4 and 0.5 hr, most of the DGEs were downregulated (81 out of 88 total DEGs). The numbers of DGEs found between 12 and 4 hr, 24 and 12 hr, as well as 48 and 24 hr

TABLE 1 | RNA-seq reads of 18 samples (three biological replicates at each time point).

Sample	Total reads	Mapped reads	Pair mapped reads	Single mapped reads	Mapped ratio(%)
0hr_rep1	25667618	25206648	25071544	135104	98.2
0hr_rep2	24461374	24025951	23875630	150321	98.22
0hr_rep3	24356981	23991626	23851008	140618	98.5
0.5hr_rep1	27750348	27226045	27037288	188757	98.11
0.5hr_rep2	29227286	28693463	28508636	184827	98.17
0.5hr_rep3	30215819	29671934	29509385	162549	98.2
4hr_rep1	29552072	28996782	28795942	200840	98.12
4hr_rep2	25707932	25227133	25058798	168335	98.13
4hr_rep3	26459312	25956585	25781988	174597	98.1
12hr_rep1	20056376	19435701	19294860	140841	96.91
12hr_rep2	25849366	25305478	25143598	161880	97.9
12hr_rep3	24639687	24048335	23903266	145069	97.6
24hr_rep1	28399172	27857378	27662324	195054	98.09
24hr_rep2	31131638	30539305	30334952	204353	98.1
24hr_rep3	29686534	29152176	28950812	201364	98.2
48hr_rep1	28813928	28353514	28145540	207974	98.4
48hr_rep2	27427442	26995636	26807948	187688	98.43
48hr_rep3	28062635	27641695	27444061	197634	98.5

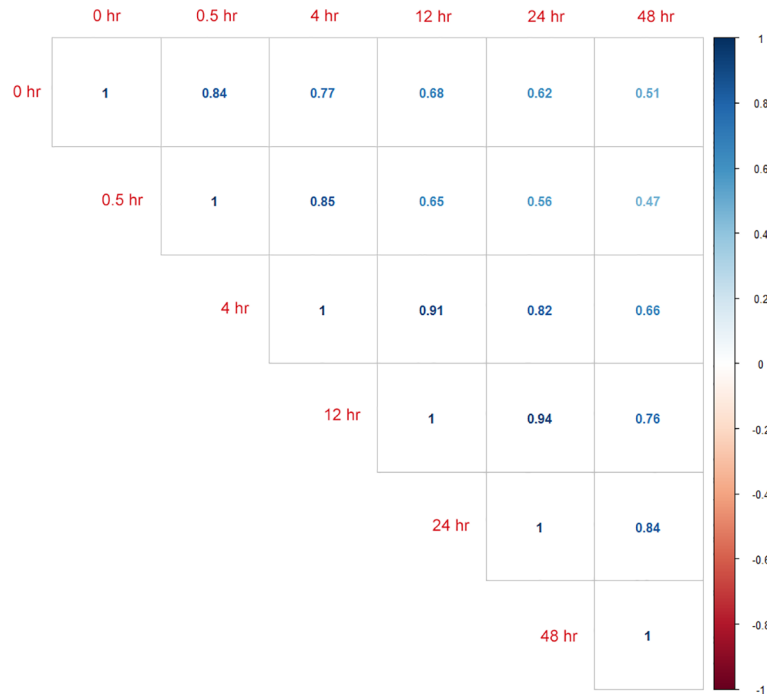


FIGURE 1 | Calculated correlation coefficients between RNA-seq data from samples of different time points (0_hour, 0.5_hour, 4_hour, 12_hour, 24_hour, and 48_hour) during resuscitation.

were 72, 85, and 172, respectively. The heat map of the five groups of DEGs is shown in **Figure 2**.

We performed a Venn analysis of the five DEG groups (**Figure 3A**). The proportion of unique genes in each group was high: 68.4% (93/136, between 0.5 and 0 hr), 60.2% (53/88, between 4 and 0.5 hr), 46.7% (35/72, between 12 and 4 hr), 47.1% (40/85, between 24 and 12 hr), and 74.4% (128/172, between 48 and 24 hr). This suggests that a variety of genes were involved at different stages of resuscitation.

Combing the DEGs of different stages, a total of 440 genes were identified (**Supplementary Table 1**), and their expression underwent dynamic changes during resuscitation (**Figure 3B**). For example, the expression of *MMAR_0922*, *MMAR_3562*, and *MMAR_1654* were significantly changed at the early stage of resuscitation, suggesting that they may play an important role in this period. Some genes had changed in multiple time periods. For example, the expression of *MMAR_3403* was changed in last three periods, suggesting that this gene may be associated with the late stage of resuscitation.

Validation of RNA-seq Results by RT-qPCR

To validate the RNA-seq results, a real-time quantitative (RT-qPCR) analysis was performed. Three biologically independent samples at each time point were used for this experiment. For each DEG group, we selected the top 10 upregulated genes and 10 downregulated genes for this analysis (**Figure 4**). Between 0 and

0.5 hr, *MMAR_4852* and *MMAR_5170* (*whiB4*) were significantly downregulated, and six genes, *MMAR_5122* (*lipX*), *MMAR_0548* (*espG3*), *MMAR_0547* (*esxR*), *MMAR_0551* (*eccE3*), *MMAR_0546* (*esxG*), and *MMAR_0550* (*mycP3*), were significantly upregulated (**Figure 4A**). Between 0.5 to 4 hr, six genes, including *MMAR_1656*, *MMAR_1658* (*hycQ*), *MMAR_1653* (*Rv0081*), *MMAR_1655*, *MMAR_5122* (*lipX*), and *MMAR_5170* (*whiB4*), were significantly downregulated and four genes *MMAR_0845* (*hemB*), *MMAR_5484*, *MMAR_1908* (*ATC1*), and *MMAR_3776* (*rpfE*) were significantly upregulated (**Figure 4B**). Between 4 and 12 hr, *MMAR_2343* (*papA1*), *MMAR_3555*, and *MMAR_2320* (*wecE*) were significantly upregulated and *MMAR_4903* were significantly downregulated (**Figure 4C**). Between 12 and 24 hr, six genes, *MMAR_0335*, *MMAR_0602*, *MMAR_4903*, *MMAR_4899*, *MMAR_2009*, and *MMAR_0615* (*iniA*), were significantly downregulated (**Figure 4D**). Between 24 and 48 hr, six genes, *MMAR_3465* (*PPE51*), *MMAR_4824*, *MMAR_4482* (*cypM*), *MMAR_4750*, *MMAR_2944*, and *MMAR_1790* (*PPE2*), were significantly downregulated, and seven genes, *MMAR_2651*, *MMAR_2914* (*katG*), *MMAR_2649*, *MMAR_5315* (*lpqH*), *MMAR_5319*, *MMAR_2839* (*mpt63*), and *MMAR_0656*, were significantly upregulated (**Figure 4E**).

There is a good agreement between the RNA-seq and qPCR data, evident from the scatter plot using the expression levels of all 97 genes that were analyzed by both RNA-seq and qPCR ($R^2 = 0.784$) (**Figure 4F**). Based on this result, we consider that the RNA-seq data is reliable.

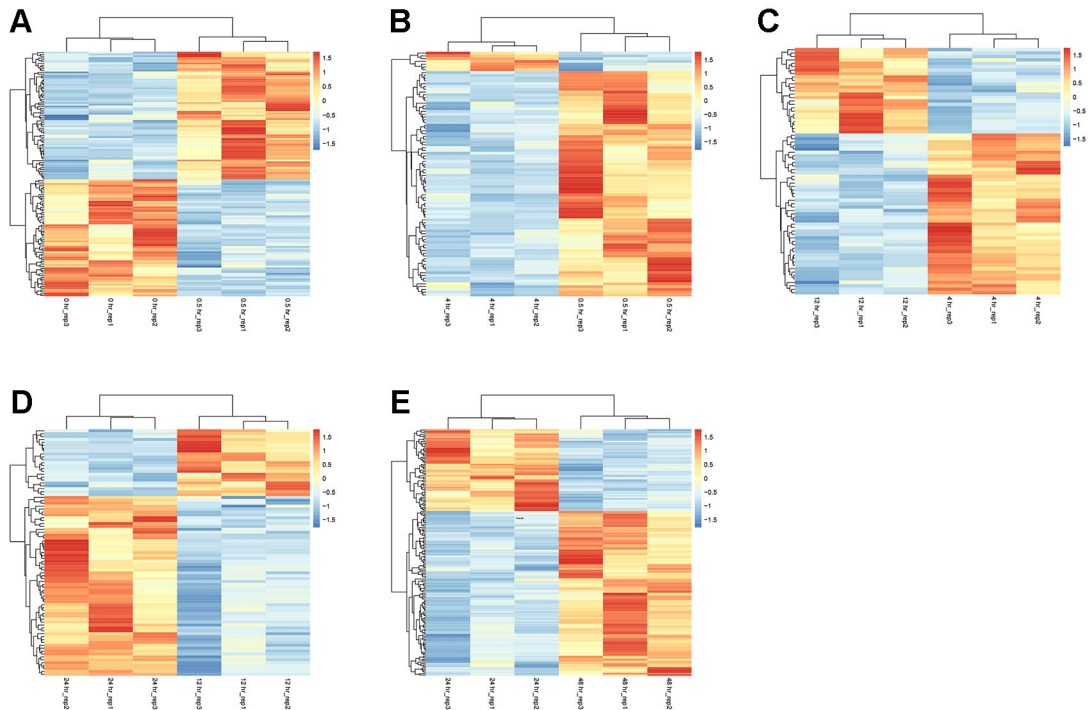


FIGURE 2 | Heatmaps of DEGs between adjacent time points. At each time point, data from three biologically independent samples were included. **(A)** 0.5 vs. 0 hr; **(B)** 4 vs. 0.5 hr; **(C)** 12 vs. 4 hr; **(D)** 24 vs. 12 hr; and **(E)** 48 vs. 24 hr. The red color indicates upregulation. The blue color indicates downregulation.

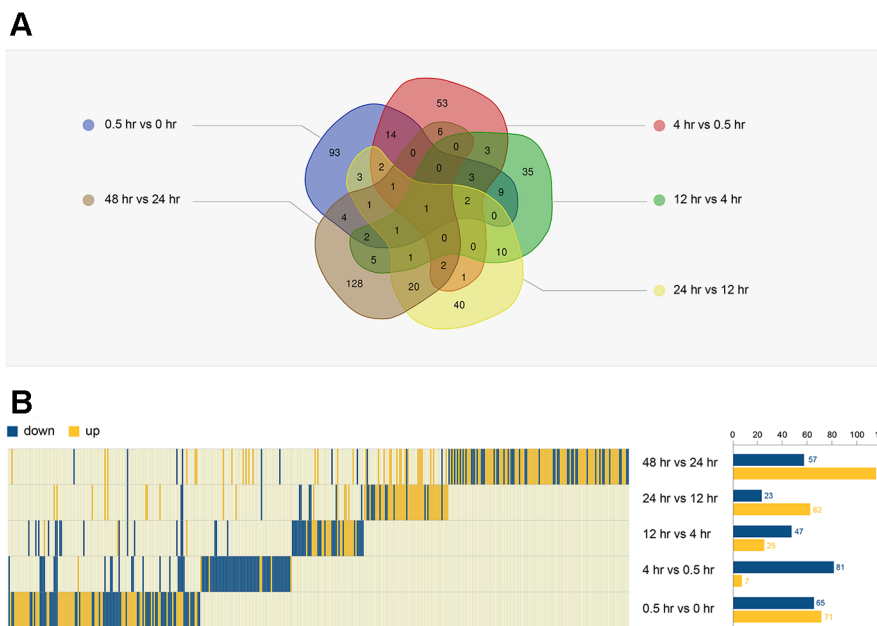
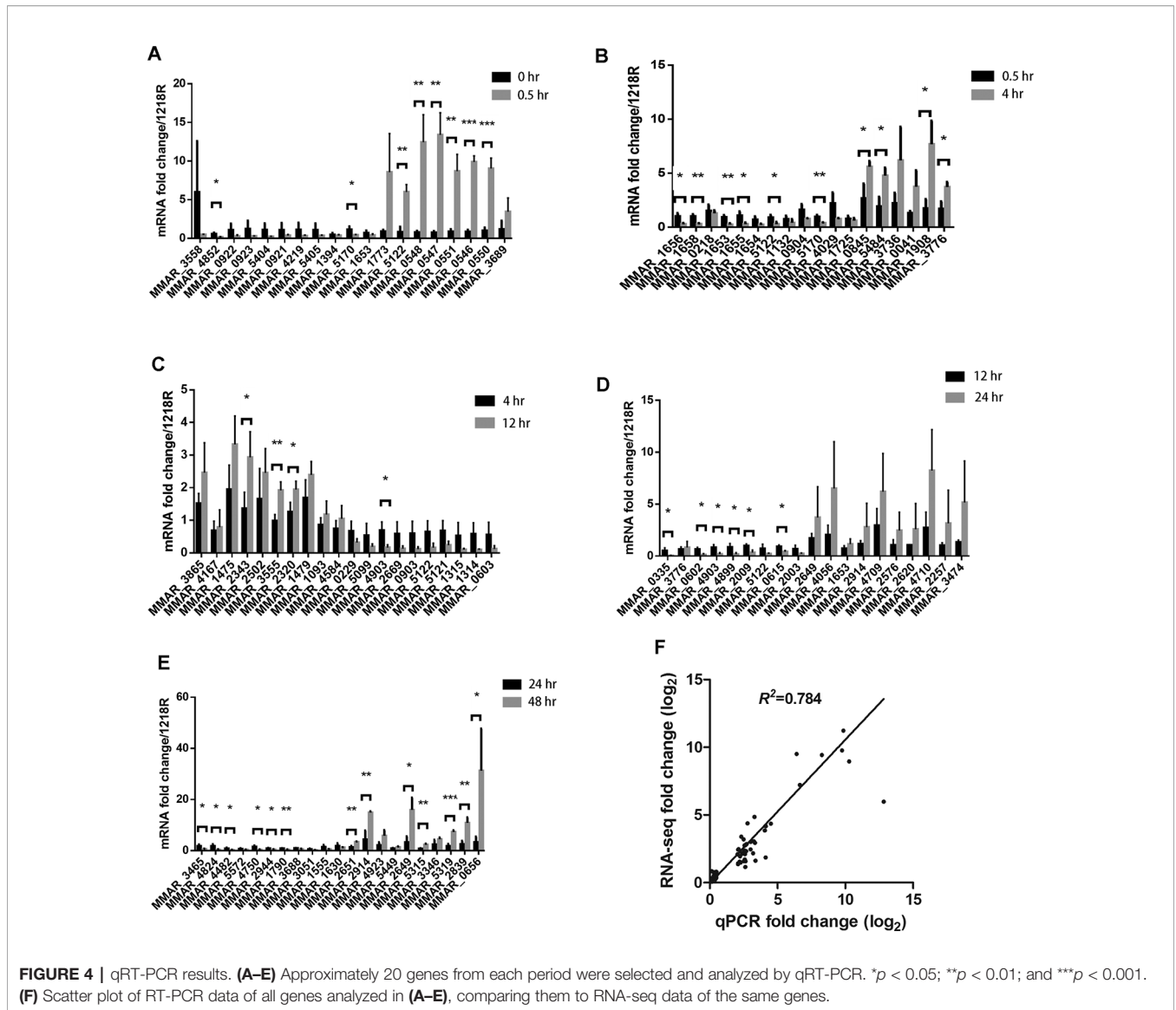


FIGURE 3 | Analysis of DEGs in five groups. **(A)** The Venn diagram of five DEGs. **(B)** The expression trend of all DEGs (440 genes) in the five periods (left). The number of DEGs in each period is shown on the right.



Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Enrichment Analysis

To gain insight into the biological consequence of the observed transcriptome changes, we performed GO and KEGG pathway analyses of the five DEG groups (Figure 5). A GO analysis was applied to identify the functional categories of DEGs. Between 0.5 and 0 hr, more than half of DEGs were involved in membrane and cell wall processes and were significantly enriched (Figure 5). This is consistent with the notion that, upon reoxygenation, the bacteria resumed cell division, which involved cell wall and membrane biogenesis. A KEGG analysis consistently revealed that DEGs involved in metabolic pathways and biosynthesis were significantly enriched and accounted for the largest number. A similar trend was observed for later periods, in which genes involved in cell wall and membrane biogenesis were highly enriched and accounted for the majority of DEGs at these stages. These results

provide snapshots into the recovery of the bacteria from hypoxia and active growth under aerobic conditions.

Co-Expression Analysis Between Transcriptional Regulators and mRNAs

Co-expression networks can show relationships between genes. To explore the regulatory mechanisms at different stages of resuscitation, we constructed a co-expression network between transcriptional regulators and mRNAs. For this, we selected known transcriptional regulators, including transcription factors and sigma factors from the five DEGs, and calculated the correlation coefficients between these transcription factors and the remaining DEGs in the same group. We considered that a relationship existed between a given transcriptional regulator and other genes if the absolute value of the correlation coefficient was greater than 0.9, which included both positive and negative correlation. Based on these results, we constructed five

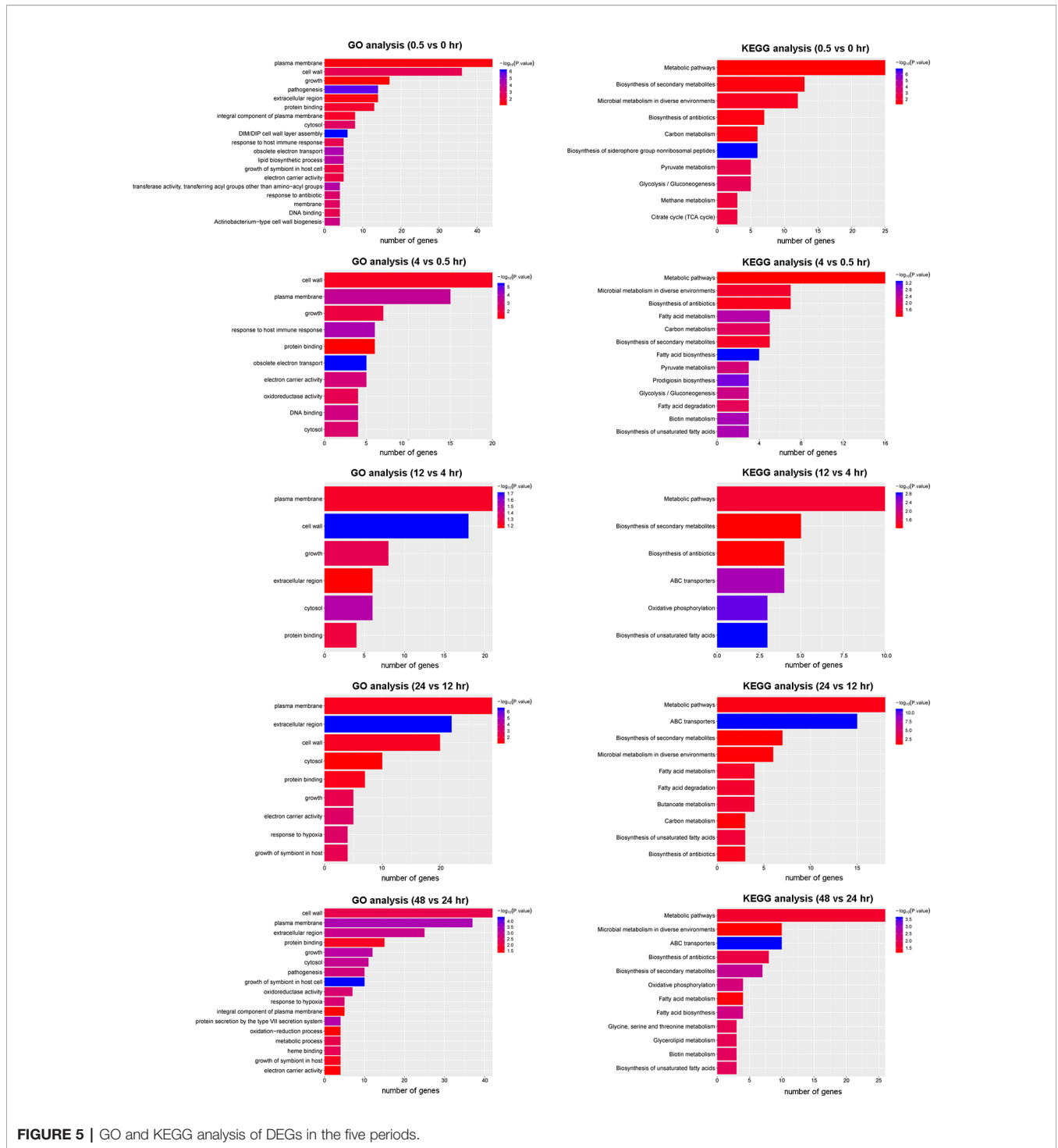


FIGURE 5 | GO and KEGG analysis of DEGs in the five periods.

co-expression modules and integrated them into a large network (**Figure 6**). The dark blue nodes in the figure represent transcription factors, and the green nodes denote DEGs in the same period. The size of the node is determined by the degree of connectivity. Greater degrees of connectivity are indicated by larger points. If there is a line between two nodes, then there is a relationship between them.

In the first co-expression module (between 0 and 0.5 hr), three transcription factors, MMAR_4874 (CosR), MMAR_1653 (Rv0081), and MMAR_4852 (KmtR), formed the major regulatory hubs, and MMAR_4874 (CosR) was the largest hub and interacted with other hubs in the network (**Figure 6**). The MMAR_4874 (CosR) and MMAR_1653 (Rv0081) hubs remained in the second co-expression module (between 0.5

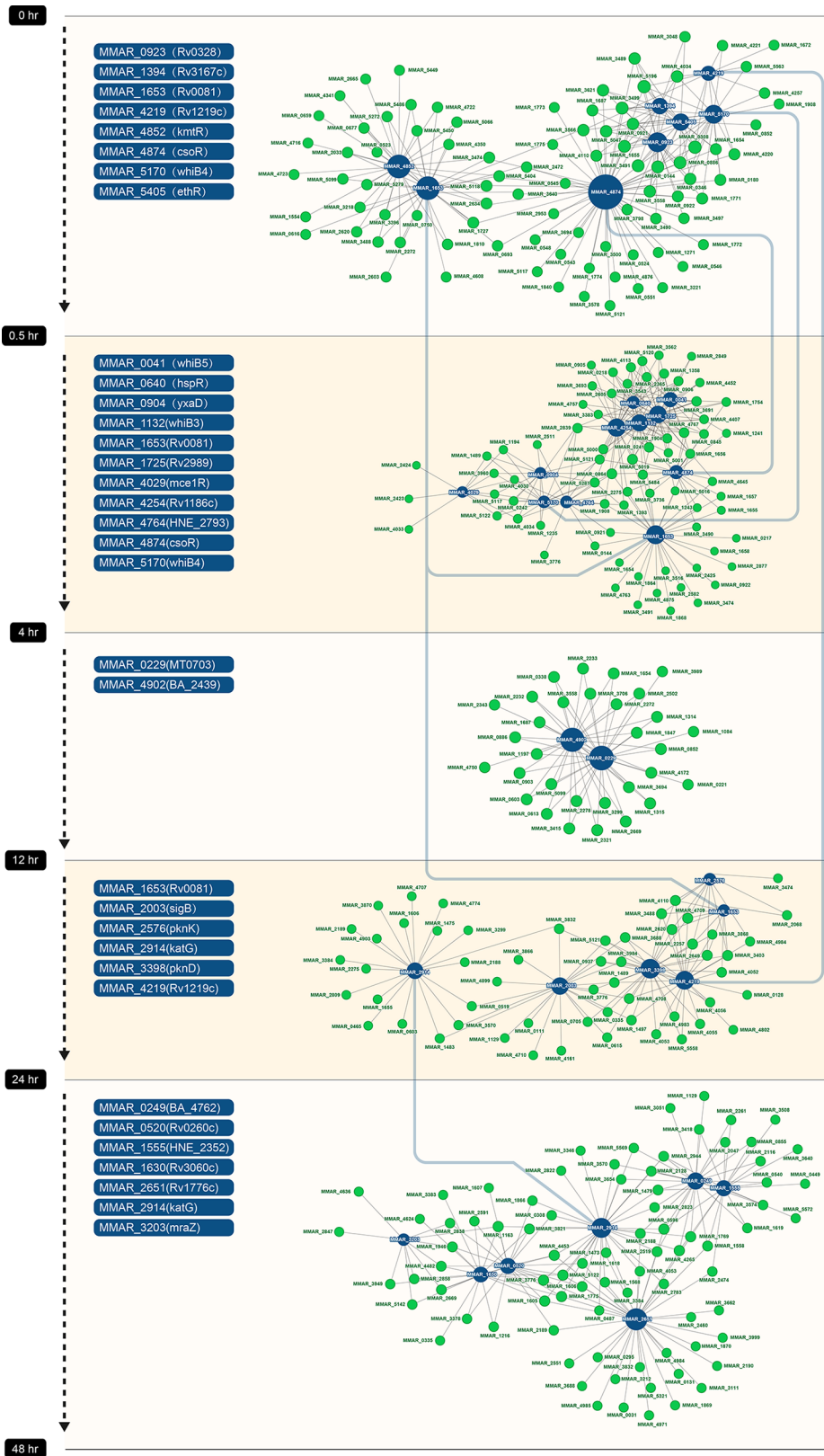


FIGURE 6 | Co-expression networks in five periods. Each dot represents a gene. Transcription factors are labeled in blue, and other genes are labeled in green. Line between dots represents co-expression relationship between genes. The size of dot is proportional to the level of connectivity.

and 4 hr), in addition to three new hubs formed by MMAR_4254, MMAR_1725 and MMAR_1132. In the third period (between 4 and 12 hr), MMAR_0229 and MMAR_4902 formed the hubs. In the fourth period (between 12 and 24 hr), MMAR_2003 (SigB) and MMAR_4219 formed the hubs. In the final period (24 to 48 hr), MMAR_2651, MMAR_1555, and MMAR_0249 formed the hubs.

Weighted Gene Co-Expression Network Analysis (WGCNA)

The DEG analysis focused on partial dynamic changes during resuscitation. While the co-expression network of transcription factors provides an overview of the regulatory programs enabling the resuscitation of *M. marinum*, our knowledge on the overall genetic changes is still missing. Therefore, in this section, we analyzed the expression of all genes from cultures at different stages of resuscitation.

Weighted gene co-expression network analysis (WGCNA) (Langfelder and Horvath, 2008) is a method for analyzing the

gene expression patterns of multiple samples. It clusters genes into modules by similar expression trends and reveals the relationship between gene modules and specific traits or phenotypes. We applied this method to analyze the RNA-seq data of *M. marinum* at different stages of resuscitation.

The gene modules were identified by the WGCNA package in R software. We first determined the appropriate “soft-thresholding” value, which emphasizes strong gene–gene correlations at the expense of weak correlations. An optimal parameter (power = 20) was determined by plotting the strength of correlation against a series (range 2 to 20) of soft threshold powers (Figure 7A).

An unsigned pairwise correlation matrix was calculated, and the WGCNA algorithm was used to transfer the correlation coefficient between genes into the adjacent coefficient. Then, the dissimilarity of the topological overlap matrix was calculated based on the adjacent coefficient. Using the calculated dissimilarity, we carried out a hierarchical analysis by using agglomerative hierarchical clustering, also known as the bottom-up method. Other assumptions were made: (i) distances between

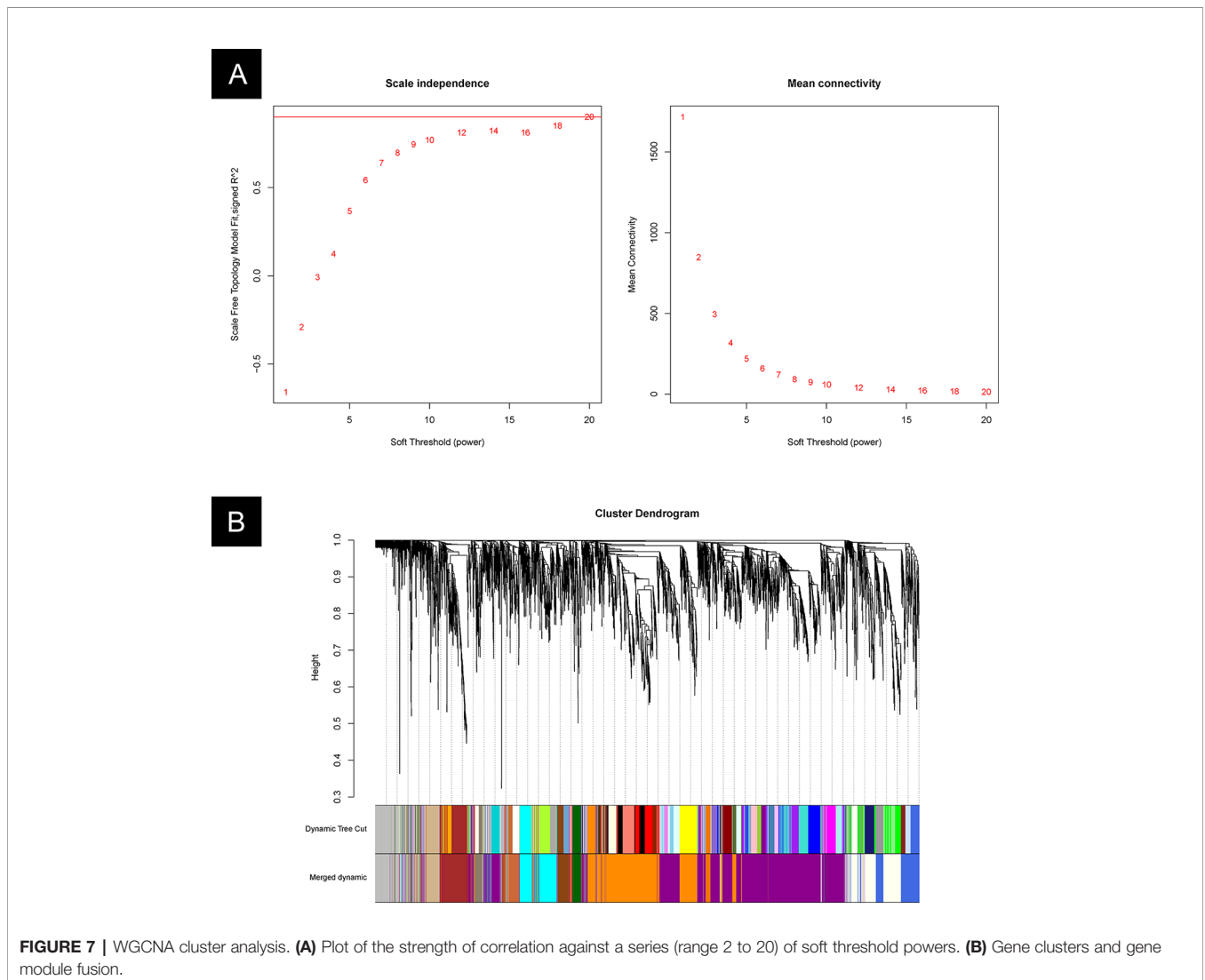


FIGURE 7 | WGCNA cluster analysis. **(A)** Plot of the strength of correlation against a series (range 2 to 20) of soft threshold powers. **(B)** Gene clusters and gene module fusion.

different classes were measured by the average connectivity, and (ii) there should be at least 30 genes in each gene module.

Based on these analyses, we initially obtained 48 gene modules. The hierarchical cluster tree was then treated using the dynamic tree cut algorithm in the WGCNA package. A total of 13 gene modules were obtained. The “gray” module was the default module, which included discarded genes that could not be clustered. Thus, we focused on the analysis in the remaining 12 gene modules. The process of fusion is shown in **Figure 7B**. The number of genes varied among these 12 modules, and the detailed information is listed in **Table 2** and **Supplementary Table 2**.

The first principal component analysis (PCA) was performed on the 12 gene modules (**Figure 8**). The PCA results reflected the main trend of gene expression in the modules. Module 20 played an important role in the early stage (0 to 0.5 hr) of recovery, module 35 played a role mostly in the middle stage (4 to 12 hr), and module 12 was only involved in the last stage (24 to 48 hr). Other modules appeared to play roles in more extended periods.

Identification of Key Gene Modules Associated With Different Stages of Resuscitation

Thus far, we have identified 5 DEGs and 12 gene modules. We then performed an enrichment analysis between them. When the *P* value of Fisher's exact test was less than 0.001, we considered that these gene modules were significantly enriched in the DEG sets. The results are shown in **Figure 9**. Interestingly, module 20 was significantly enriched in DEGs of the early stage (0 to 0.5 hr) and module 12 was significantly enriched in DEGs of the last stage (24 to 48 hr) of the recovery. This is consistent with the result that these two modules were only involved in the early and last stages of resuscitation, respectively (**Figure 8**).

DISCUSSION

In this study, we examined the transcriptome changes of *M. marinum* recovered from hypoxia-induced dormancy. To gain a comprehensive view, multiple time points, including shortly after resuscitation (0.5 hr) to more extended periods up to 48 hr, were included. For each time point, three biologically independent

samples were analyzed. Transcripts of the whole genome were analyzed by RNA-seq, and the quality of the RNA-seq data was reflected by the high genome mapping ratio and further validated by qPCR analysis of close to 100 genes. With these high-quality sequence data, we performed in-depth analyses, which included the identification of DEGs and the construction of co-expression network of transcription factors in each period. The availability of transcriptomes of independent samples at multiple time points also allowed us to employ a weighted gene co-expression network analysis to identify gene modules of *M. marinum*. Collectively, these data provide valuable insight into not only the genetic changes of the bacteria upon resuscitation but also the gene module function of *M. marinum* during active metabolism and growth.

A total of 136 DEGs were identified in *M. marinum* upon resuscitation from dormancy (0 to 0.5 hr), including eight transcription factors (**Figure 6**). Interestingly, all of these transcription factors were significantly downregulated. Among them, MMAR_1653 is a homolog of Rv0081 in *M. tb*. Previous studies have shown that Rv0081 is a member of the DosR regulon and is induced at the early stage of hypoxia (Sherman et al., 2001). Rv0081 is a major regulator of *M. tb* response to hypoxia and forms a large regulatory hub (Galagan et al., 2013; Chawla et al., 2018; Sun et al., 2018). Rv0081 plays an important role connecting the early and enduring hypoxic responses (Sun et al., 2018). WhiB4 (MMAR_5170) is an oxygen-sensitive transcription factor and has been shown to regulate PE/PPE family proteins, and it plays a role in *M. marinum* virulence (Chawla et al., 2012; Wu et al., 2017). CosR (MMAR_4874) is a copper-inducible transcriptional regulator, and the loss of *cosR* resulted in a hypoxia-type response with the induction of the DosR regulon (Talaat et al., 2004; Ward et al., 2008; Marcus et al., 2016). Given their roles in the hypoxic response, it is not surprising that Rv0081, WhiB4, and CosR underwent dynamic changes in expression upon resuscitation by reaeration. Two other transcription factors that were downregulated at this stage, MMAR_5405 (EthR) and MMAR_1394 (Rv3176c), belong to the TetR family transcription factors (Leiba et al., 2014; Sharma et al., 2017).

We also found that multiple members of the ESAT-6 family proteins were upregulated upon resuscitation (Harboe et al., 1996; Priscille et al., 2004), including EsxA (Sandra et al., 2010; Zhang et al., 2016), EsxB (Sandra et al., 2010), EsxG (Sweeney et al., 2011), EsxH (Alka et al., 2013; Portal-Celhay et al., 2016), EsxK, and EsxN (Zhang et al., 2017). These proteins are components of the Type VII secretion systems, and many of them are important T cell antigens and play a critical role modulating the host-pathogen interactions (Abdallah et al., 2007).

From 0.5 to 4 hr after reaeration, 11 transcription factors were downregulated (**Figure 6**), including Rv0081, CsoR, and WhiB4 as mentioned above. Notably, the expression of two other WhiB family proteins (Averina et al., 2012), WhiB3 and WhiB5, were also significantly altered. WhiB3 responds to dormancy signals, including hypoxia and NO, and controls redox homeostasis of the bacteria (Priscille et al., 2004). WhiB5 responds to oxygen and controls the expression type VII secretion systems (Priscille et al., 2004). Consistently, we found that *whiB3* was downregulated while *whiB5* was upregulated at this stage of resuscitation.

TABLE 2 | Information of gene modules identified by WGCNA.

Module_ID	Number of genes
module_34	1711
module_26	919
module_43	424
module_14	320
module_20	298
module_3	268
module_35	142
module_29	136
module_12	127
module_22	89
module_47	89
module_45	46

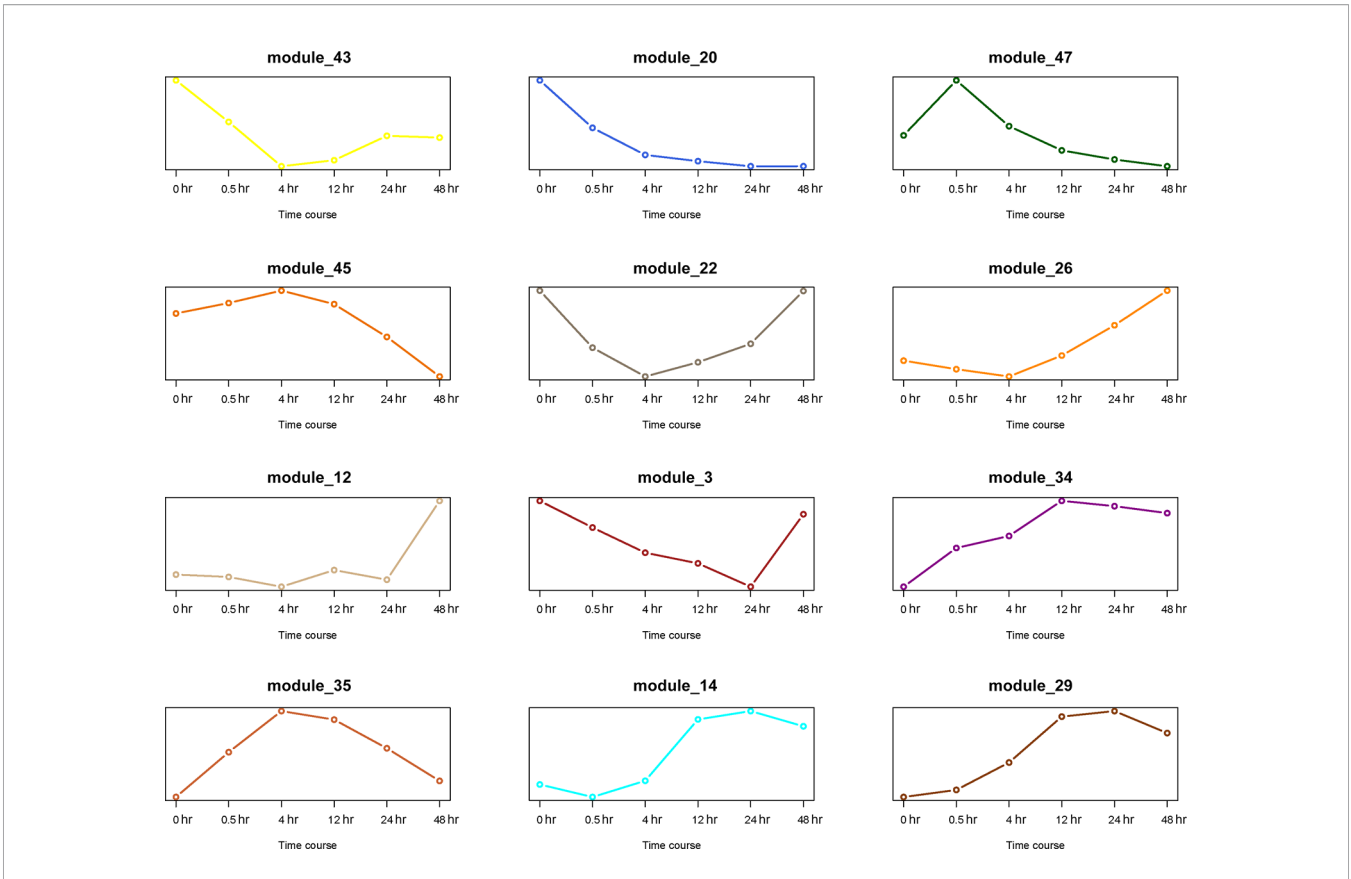


FIGURE 8 | PCA analysis of 12 gene modules. The X axis represents time period, and the Y axis represents expression of first principal component.

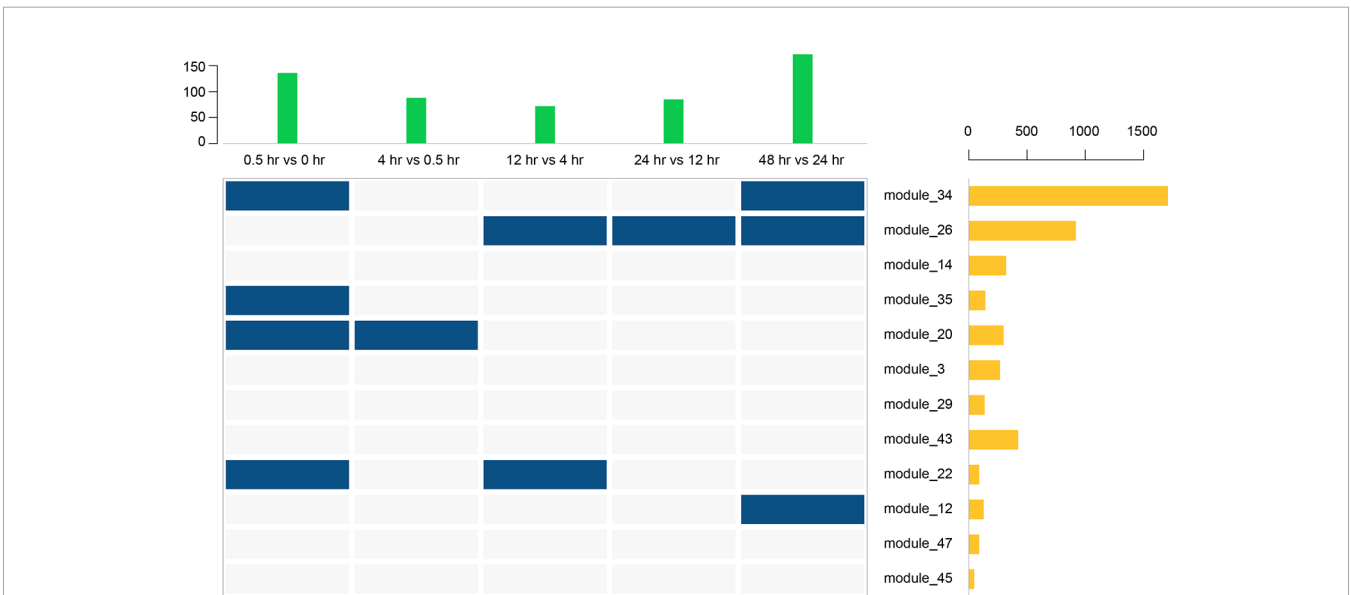


FIGURE 9 | Enrichment analysis between 12 gene modules and 5 DEGs. The above bar chart represents the number of DEGs in each period, and the bar chart on the right represents the number of genes in each gene module. Blue squares represent a significant enrichment between the row set and the column set.

During the period of 4 to 12 hr, we identified 72 DEGs, including 7 PPE family genes (*MMAR_5121*, *MMAR_1095*, *MMAR_1235*, *MMAR_1847*, *MMAR_1905*, *MMAR_2669*, and *MMAR_3989*). The PE/PPE family proteins play a critical role in mycobacterial pathogenesis (Fishbein et al., 2015). In addition, several genes of the Mce family were upregulated, including *MMAR_3865* (*mce2B*), *MMAR_3868* (*mce5E*), *MMAR_3867* (*mce5D*), *MMAR_3866* (*mce5C*). The Mce family genes comprise four mammalian cell invasion factor (*mce*) operons (*mce1-4*), and some of these are involved in the invasion of host cells (Zhang and Xie, 2011).

A picture appears to have emerged from these analyses; in the early stage of resuscitation from the hypoxia-induced dormancy, transcription factors critical for a hypoxia-induced response are downregulated, and, as the recovery continues, genes important for virulence and host interactions are upregulated.

A WGCNA analysis revealed 12 distinct gene modules. Of particular interest is gene module 20, which was involved in the early stage of resuscitation only (Figure 8). This module comprises of ~300 genes, many of which have unknown functions or annotations. Future studies focusing on genes in this module may help to understand the molecular machinery enabling the exit of the bacteria from dormancy.

METHODS AND MATERIALS

Bacterial Strain, Media, and Growth Conditions

M. marinum 1218R (ATCC 927) was grown in Middlebrook 7H9 broth to OD600~0.5, at which point they were aliquoted and cultured in screw-capped conical flasks at 30°C without additional oxygen. The hypoxic culture conditions were described previously by Wayne and Hayes (Wayne and Hayes, 1996). After 7 days in hypoxic conditions, the screw cap was replaced with a permeable membrane, and the rest of the conditions were unchanged. After aeration, samples were taken at 0 h, 0.5 h, 4 h, 12 h, 24 h, and 48 h, and an aliquot was used to measure the OD value (Supplementary Figure 1). The remaining samples were collected and snap frozen in liquid nitrogen for RNA sequencing.

RNA Extraction, Illumina Sequencing, and RT-qPCR

M. marinum cultures were centrifuged at 4,500 × g for 5 min at room temperature and frozen on dry ice. The frozen cell pellets were resuspended in 1 mL TRIzol reagent (CW Bio). RNA extraction and illumina sequencing were performed as previously described (Wu et al., 2017). Raw data of RNA sequencing have been uploaded to the GEO database (BioProject ID : PRJNA588556).

For RT-qPCR validation of RNA-seq data, 1 μg RNA was reversed-transcribed to cDNA, which was then used as the template for RT-PCR analysis. The primers for analyzing the selected genes were listed in Supplementary Table 3.

Transcriptome and Bioinformatics Analysis

The RNA-seq analysis and identification of differentially expressed genes (DEGs) were performed as previously described (Lee et al., 2019).

GO and KEGG Analysis

We used the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Ogata et al., 2000) (www.genome.jp/kegg/) and the Gene Ontology (GO) database (Ashburner et al., 2000) (www.geneontology.org/) for our data analysis.

Construction of the Gene Co-Expression Network

We calculated the correlation coefficient between identified transcription factors and other DEGs in each time period. The correlation coefficient value ranges from -1 to 1, representing a negative and positive correlation, respectively. We considered the expression of two genes as correlated if the absolute value of the correlation coefficient was larger than 0.8. The results were imported into Cytoscape 3.0 to generate the network map (Kohl et al., 2011).

Weighted Gene Co-Expression Network Analysis (WGCNA)

We used the RNA-seq data from multiple time points (three biological independent samples at each time point) for WGCNA analysis. We used the WGCNA package to cluster gene modules as follows.

- Define gene co-expression similarity: calculate the similarity between any two genes using Pearson's correlation coefficient ($S_{ij} = |\text{cor}(i,j)|$), the correlation coefficient of gene *i* and gene *j*), which then forms the correlation matrix ($S = [S_{ij}]$).
- Define the exponential weighted value β : for any gene pair (*i* and *j*), apply the exponential adjacency function in the WGCNA algorithm to measure their relation index, namely, the exponential weighted β square of the correlation coefficient ($a_{ij} = \text{power}(S_{ij}, \beta) = |S_{ij}|^\beta$). Exponential weighted β is the power of the correlation coefficient. We selected $\beta = 5$ after the analysis (fit value R^2 to approximately 0.9).
- Define a measure of node dissimilarity: after determining the adjacency function parameter β , the correlation matrix $S = [S_{ij}]$ is switched into the adjacency matrix $A = [a_{ij}]$ and converted into the topological overlap matrix $\Omega = [\omega_{ij}]$. k_i or k_j indicate the sum of one node's adjacency coefficients. The node is a gene (*i* or *j*).

$$\omega_{ij} = \frac{l_{ij} + a_{ij}}{\min\{k_i, k_j\} + 1 - a_{ij}}$$

$$l_{ij} = \sum_{\mu} a_{i\mu} a_{\mu j} \quad k_i = \sum_{\mu} a_{i\mu} \quad k_j = \sum_{\mu} a_{\mu j}$$

- Build hierarchical clustering tree to identify gene modules: the hierarchical clustering tree built using the dissimilarity coefficient d_{ij}^{ω} ($d_{ij}^{\omega} = 1 - \omega_{ij}$), and the different branches represent the gene modules.

Enrichment Analysis and PCA Analysis

To determine whether one set of genes were more enriched in another set of genes, we used the Chi-square test or Fisher's exact test (Upton, 1992). First, the two sets of genes were taken and used to form a 2*2 contingency table. If there was a value less than or equal to five in the table, the Fisher's exact test was applied; otherwise, the Chi-square test was applied. When the p value was less than 0.05, the two sets were considered significantly enriched to each other. PCA analyses were performed by the princomp function in R software (version 3.5.1)

DATA AVAILABILITY STATEMENT

Raw data of RNA sequencing have been uploaded to the GEO database (BioProject ID:PRJNA588556).

AUTHOR CONTRIBUTIONS

JJ and CL performed the experiment. JJ, CL, and JL wrote the manuscript. YW and JZ prepared **Figures 1–9**. LS, KY, and WX prepared **Tables 1 and 2**. LZ, YL, and JL provided guidance for experiments. All authors reviewed the manuscript.

REFERENCES

- Abdallah, A. M., Gey van Pittius, N. C., Champion, P. A., Cox, J., Luirink, J., Vandembroucke-Grauls, C. M., et al. (2007). Type VII secretion—mycobacteria show the way. *Nat. Rev. Microbiol.* 5 (11), 883–891. doi: 10.1038/nrmicro1773
- Alka, M., Aleena, Z., Victor, T., Natalie, S., Ashley, W., Maura, P., et al. (2013). Mycobacterium tuberculosis Type VII secreted effector EsxH targets host ESCRT to impair trafficking. *PLoS Pathogens* 9 (10), e1003734. doi: 10.1371/journal.ppat.1003734
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. *Gene Ontol. Consortium Nat. Genet.* 25 (1), 25–29. doi: 10.1038/75556
- Averina, O. V., Zakharevich, N. V., and Danilenko, V. N. (2012). Identification and characterization of WhiB-like family proteins of the Bifidobacterium genus. *Anaerobe* 18 (4), 421–429. doi: 10.1016/j.anaerobe.2012.04.011
- Boon, C., and Dick, T. (2002). Mycobacterium bovis BCG response regulator essential for hypoxic dormancy. *J. Bacteriol.* 184 (24), 6760–6767. doi: 10.1128/JB.184.24.6760-6767.2002
- Boshoff, H. I., and Barry, C. E.3rd (2005). Tuberculosis - metabolism and respiration in the absence of growth. *Nat. Rev. Microbiol.* 3 (1), 70–80. doi: 10.1038/nrmicro1065
- Chawla, M., Parikh, P., Saxena, A., Munshi, M., Mehta, M., Mai, D., et al. (2012). Mycobacterium tuberculosis WhiB4 regulates oxidative stress response to modulate survival and dissemination *in vivo*. *Mol. Microbiol.* 85 (6), 1148–1165. doi: 10.1111/j.1365-2958.2012.08165.x
- Chawla, M., Mishra, S., Anand, K., Parikh, P., Mehta, M., Vij, M., et al. (2018). Redox-dependent condensation of the mycobacterial nucleoid by WhiB4. *Redox Biol.* 19, 116–133. doi: 10.1016/j.redox.2018.08.006
- Corbett, E. L., Bandason, T., Cheung, Y. B., Munyati, S., Godfrey-Faussett, P., Hayes, R., et al. (2007). Epidemiology of tuberculosis in a high HIV prevalence population provided with enhanced diagnosis of symptomatic disease. *PLoS Med.* 4 (1), e22. doi: 10.1371/journal.pmed.0040022
- Fishbein, S., van Wyk, N., Warren, R. M., and Sampson, S. L. (2015). Phylogeny to function: PE/PPE protein evolution and impact on Mycobacterium tuberculosis pathogenicity. *Mol. Microbiol.* 96 (5), 901–916. doi: 10.1111/mmi.12981
- Galagan, J. E., Minch, K. J., Peterson, M. W., Lyubetskaya, A., Azizi, E., Sweet, L., et al. (2013). The Mycobacterium tuberculosis regulatory network and hypoxia. *Nature* 499 (7457), 178–183. doi: 10.1038/nature12337

ACKNOWLEDGMENTS

This study was supported by grants from China's 13th Five Year Programs for the prevention and cure of great infectious diseases (2017ZX10201301-005), the National Key R&D Program of China (No.2018YFD0500900), and the Open Fund of Shanghai Key Laboratory of Tuberculosis (2018kf03).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.01359/full#supplementary-material>

SUPPLEMENTARY FIGURE 1 | Growth curve of *M. marinum* after re-aeration from hypoxic cultures.

SUPPLEMENTARY TABLE 1 | Detailed information on the differentially expressed gene list at different time periods.

SUPPLEMENTARY TABLE 2 | List of genes identified in 12 gene modules by WGCNA analysis.

SUPPLEMENTARY TABLE 3 | Primer sequences of genes analyzed by RT-PCR.

- Harboe, M., Oettinger, T., Wiker, H. G., Rosenkrands, I., and Andersen, P. (1996). Evidence for occurrence of the ESAT-6 protein in Mycobacterium tuberculosis and virulent Mycobacterium bovis and for its absence in Mycobacterium bovis BCG. *Infection Immunity* 64 (1), 16–22. doi: 10.1128/IAI.64.1.16-22
- Iona, E., Pardini, M., Mustazzolu, A., Piccaro, G., Nisini, R., Fattorini, L., et al. (2016). Mycobacterium tuberculosis gene expression at different stages of hypoxia-induced dormancy and upon resuscitation. *J. Microbiol.* 54 (8), 565–572. doi: 10.1007/s12275-016-6150-4
- Kohl, M., Wiese, S., and Warscheid, B. (2011). Cytoscape: software for visualization and analysis of biological networks. *Methods Mol. Biol.* 696 (696), 291–303. doi: 10.1007/978-1-60761-987-1_18
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9 (1), 559–559. doi: 10.1186/1471-2105-9-559
- Lee, J., Lee, S.-G., Kim, K. K., Lim, Y.-J., Choi, J.-A., Cho, S.-N., et al. (2019). Characterisation of genes differentially expressed in macrophages by virulent and attenuated Mycobacterium tuberculosis through RNA-Seq analysis. *Sci. Rep.* 9 (1), 4027. doi: 10.1038/s41598-019-40814-0
- Leiba, J., Carrère-Kremer, S., Blondiaux, N., Dimala, M. M., Wohlkönig, A., Baulard, A., et al. (2014). The mycobacterium tuberculosis transcriptional repressor EthR is negatively regulated by Serine/Threonine phosphorylation. *Biochem. Biophys. Res. Commun.* 446 (4), 1132–1138. doi: 10.1016/j.bbrc.2014.03.074
- Marcus, S. A., Sidiropoulos, S. W., Steinberg, H., and Talaat, A. M. (2016). csor is essential for maintaining copper homeostasis in mycobacterium tuberculosis. *PLoS One* 11 (3), e0151816. doi: 10.1371/journal.pone.0151816
- Mcgillivray, A., Golden, N. A., and Kaushal, D. (2015). The mycobacterium tuberculosis Clp gene regulator is required for *in vitro* reactivation from hypoxia-induced dormancy. *J. Biol. Chem.* 290 (4), 2351–2367. doi: 10.1074/jbc.M114.615534
- North, R. J., and Jung, Y. J. (2004). Immunity to tuberculosis. *Annu. Rev. Immunol.* 22, 599–623. doi: 10.1146/annurev.immunol.22.012703.104635
- Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (2000). (KEGG). Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 27 (1), 29–34. doi: 10.1093/nar/28.1.27
- Park, H. D., Guinn, K. M., Harrell, M. I., Liao, R., Voskuil, M. I., Tompa, M., et al. (2003). Rv3133c/dosR is a transcription factor that mediates the hypoxic response of Mycobacterium tuberculosis. *Mol. Microbiol.* 48 (3), 833–843. doi: 10.1046/j.1365-2958.2003.03474.x

- Portal-Celhay, C., Tufariello, J. M., Srivastava, S., Zahra, A., Klevorn, T., Grace, P. S., et al. (2016). Mycobacterium tuberculosis EsxH inhibits ESCRT-dependent CD4+ T-cell activation. *Nat. Microbiol.* 2, 16232. doi: 10.1038/nmicrobiol.2016.232
- Priscille, B., Ida, R., Peter, A., Cole, S. T., and Roland, B. (2004). ESAT-6 proteins: protective antigens and virulence factors? *Trends Microbiol.* 12 (11), 500–508. doi: 10.1016/j.tim.2004.09.007
- Rustad, T. R., Harrell, M. I., Liao, R., and Sherman, D. R. (2008). The enduring hypoxic response of Mycobacterium tuberculosis. *PLoS One* 3 (1), e1502. doi: 10.1371/journal.pone.0001502
- Rustad, T. R., Sherrid, A. M., Minch, K. J., and Sherman, D. R. (2009). Hypoxia: a window into Mycobacterium tuberculosis latency. *Cell Microbiol.* 11 (8), 1151–1159. doi: 10.1111/j.1462-5822.2009.01325.x
- Sandra, A. S. R., Facey, P. D., Lorena, F. M., Caridad, R., Carlos, V., Ricardo, D. S., et al. (2010). A heterodimer of EsxA and EsxB is involved in sporulation and is secreted by a type VII secretion system in *Streptomyces coelicolor*. *Microbiology* 156 (Pt 6), 1719. doi: 10.1099/mic.0.037069-0
- Sharma, A., Singh, K., and Kaur, J. (2017). MesT, a unique epoxide hydrolase, is essential for optimal growth of Mycobacterium tuberculosis in the presence of styrene oxide. *Future Microbiol.* 00 (00), 527–546. doi: 10.2217/fmb-2016-0206
- Sherman, D. R., Voskuil, M., Schnappinger, D., Liao, R., Harrell, M. I., and Schoolnik, G. K. (2001). Regulation of the Mycobacterium tuberculosis hypoxic response gene encoding alpha-crystallin. *Proc. Natl. Acad. Sci. U. S. A.* 98 (13), 7534–7539. doi: 10.1073/pnas.121172498
- Stewart, G. R., Robertson, B. D., and Young, D. B. (2003). Tuberculosis: a problem with persistence. *Nat. Rev. Microbiol.* 1 (2), 97–105. doi: 10.1038/nrmicro749
- Sun, X., Zhang, L., Jiang, J., Ng, M., Cui, Z., Mai, J., et al. (2018). Transcription factors Rv0081 and Rv3334 connect the early and the enduring hypoxic response of Mycobacterium tuberculosis. *Virulence* 9 (1), 1468. doi: 10.1080/21505594.2018.1514237
- Sweeney, K. A., Dao, D. N., Goldberg, M. F., Tsungda, H., Venkataswamy, M. M., Marcela, H. T., et al. (2011). A recombinant Mycobacterium smegmatis induces potent bactericidal immunity against Mycobacterium tuberculosis. *Nat. Med.* 17 (10), 1261. doi: 10.1038/nm.2420
- Talaat, A. M., Rick, L., Howard, S. T., and Stephen Albert, J. (2004). The temporal expression profile of Mycobacterium tuberculosis infection in mice. *Proc. Natl. Acad. Sci. U. S. A.* 101 (13), 4602–4607. doi: 10.1073/pnas.0306023101
- Tobin, D. M., and Ramakrishnan, L. (2008). Comparative pathogenesis of Mycobacterium marinum and Mycobacterium tuberculosis. *Cell Microbiol.* 10 (5), 1027–1039. doi: 10.1111/j.1462-5822.2008.01133.x
- Upton, G. J. G. (1992). Fisher's exact test. *J. Royal Stat. Soc.* 155 (3), 395–402.
- Veatch, A. V., and Kaushal, D. (2018). Opening Pandora's box: mechanisms of mycobacterium tuberculosis resuscitation. *Trends Microbiol.* 26 (2), 145. doi: 10.2307/2982890
- Veatch, A. V., Niu, T., Caskey, J., McGillivray, A., Gautam, U. S., Subramanian, R., et al. (2016). Sequencing-relative to hybridization-based transcriptomics approaches better define Mycobacterium tuberculosis stress-response regulons. *Tuberculosis* 101S, S9–S17.
- Ward, S. K., Hoye, E. A., and Talaat, A. M. (2008). The Global responses of mycobacterium tuberculosis to physiological levels of copper. *J. Bacteriol.* 190 (8), 2939–2946. doi: 10.1128/JB.01847-07
- Wayne, L. G., and Hayes, L. G. (1996). An *in vitro* model for sequential study of shiftdown of Mycobacterium tuberculosis through two stages of nonreplicating persistence. *Infect. Immun.* 64 (6), 2062–2069. doi: 10.1016/1380-2933(96)00040-1
- Wayne, L. G., and Sohaskey, C. D. (2001). Nonreplicating persistence of Mycobacterium tuberculosis. *Annu. Rev. Microbiol.* 55, 139–163. doi: 10.1146/annurev.micro.55.1.139
- Wayne, L. G. (1977). Synchronized replication of Mycobacterium tuberculosis. *Infect. Immun.* 17 (3), 528–530. doi: 10.1128/IAI.17.3.528-530.1977
- Wu, J., Ru, H., Xiang, Z., Jiang, J., Wang, Y., Zhang, L., et al. (2017). WhiB4 regulates the PE/PPE gene family and is essential for virulence of mycobacterium marinum. *Sci. Rep.* 7 (1), 3007. doi: 10.1038/s41598-017-03020-4
- Yu, J., Tran, V., Li, M., Huang, X., Niu, C., Wang, D., et al. (2012). Both phthiocerol dimycocerosates and phenolic glycolipids are required for virulence of Mycobacterium marinum. *Infect. Immun.* 80 (4), 1381–1389. doi: 10.1128/IAI.06370-11
- Zhang, F., and Xie, J. P. (2011). Mammalian cell entry gene family of Mycobacterium tuberculosis. *Mol. Cell. Biochem.* 352 (1-2), 1–10. doi: 10.1007/s11010-011-0733-5
- Zhang, Q., Wang, D., Jiang, G., Liu, W., Deng, Q., Li, X., et al. (2016). EsxA membrane-permeabilizing activity plays a key role in mycobacterial cytosolic translocation and virulence: effects of single-residue mutations at glutamine 5. *Sci. Rep.* 6 (1), 32618. doi: 10.1038/srep32618
- Zhigang, X. U., Dezhou, L. I., Chen, X., Duan, Z., Mao, J., and Wen., J. (2017). Identification and application of Mycobacterium tuberculosis esxN-specific cell epitopes in the diagnosis of pulmonary tuberculosis. *Chin. J. Cell. Mol. Immunol.* 33 (12), 1686–1691.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Jiang, Lin, Zhang, Wang, Shen, Yang, Xiao, Li, Zhang and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.