WhiB6 is required for the secretion-dependent regulation of ESX-1 substrates in pathogenic mycobacteria.

Abdallah M. Abdallah[1]*@, E.M. Weerdenburg[2]*, Qingtian Guan[1], R. Ummels[2], S. Borggreve[2], S.A. Adroub[1], Tareq B. Malas[1], Raeece Naeem[1], Huoming Zhang[3], T.D. Otto[4], W. Bitter[2@]** & A. Pain[1@]**

[1]Pathogen Genomics Laboratory, BESE Division, King Abdullah University of Science and Technology (KAUST), Thuwal-Jeddah, Kingdom of Saudi Arabia; [2]Department of Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, The Netherlands; [3]Bioscience Core Laboratory, King Abdullah University of Science and Technology (KAUST), Thuwal-Jeddah, Kingdom of Saudi Arabia; [4]Pathogen Genomics, The Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom

*Joint authors
@for correspondence
Abdallah M. Abdallah (abdallah.abdallah@kaust.edu.sa)
Wilbert Bitter (w.bitter@vumc.nl)
Arnab Pain (arnab.pain@kaust.edu.sa)

Abstract

The mycobacterial type VII secretion system ESX-1 is responsible for secretion of a number of proteins that play an important role during infection of the host. Regulation of expression of secreted proteins is often essential to establish a successful infection. Using transcriptome sequencing, we found that abrogation of ESX-1 function in *Mycobacterium marinum* leads to a pronounced increase in gene expression levels of the *espA* operon during infection of macrophages, suggesting an important role in ESX-1-mediated virulence during the early phase of infection. In addition, we found that disruption of ESX-1-mediated protein secretion leads to a specific down-regulation of the substrates, but not the structural components of this system, in both *M. marinum* and *M. tuberculosis* during growth in culture medium. We established that down-regulation of ESX-1 substrates is the result of a regulatory process, which is influenced by the putative transcriptional regulator *whib6*, located adjacent to the *esx-1* locus. In addition, overexpression of the ESX-1 associated PE35/PPE68 protein pair resulted in significantly increased secretion of the ESX-1 substrate EsxA, demonstrating a functional link between these proteins. Together, these data show that ESX-1 substrates are regulated independently from the structural components, both during infection and as a result of active secretion.

Introduction

Mycobacteria use several different type VII secretion systems (T7S) to transport proteins across their thick and waxy cell envelope. One of these T7S systems, ESX-1, is responsible for the transport of a number of important virulence factors. Disruption of the *esx-1* gene cluster severely reduces virulence of *M. tuberculosis* (1), whereas restoration of *esx-1* in the *Mycobacterium bovis*-derived vaccine strain BCG, which has lost part of the *esx-1* region as a result of continuous passaging, leads to an increase in virulence (2). Many studies have tried to elucidate the function of ESX-1 substrates in virulence. In pathogenic mycobacteria, such as *Mycobacterium tuberculosis* and the fish pathogen *Mycobacterium marinum*, ESX-1 is responsible for the translocation of these bacteria from the phagolysosomal compartment to the cytosol of macrophages (3, 4). This translocation activity has been attributed to the membrane-lysing capacity of the secreted protein EsxA (also called ESAT-6) (5, 6). Interestingly, a close homologue of this protein is also secreted by non-pathogenic and non-translocating mycobacteria such as *Mycobacterium smegmatis*. A recent report indicated that, although EsxA of *M. smegmatis* and *M. tuberculosis* are highly homologous, their membrane-lysing potential is different (7). In *M. smegmatis*, ESX-1 is involved in a completely different process, *i.e.* conjugative DNA transfer (8). The proposed functions of ESX-1 in pathogenic mycobacterial species include host cell entry and intercellular spread (9).

The ESX-1 substrates identified so far are mostly encoded by genes of the *esx-1* locus, such as EsxA, EsxB, EspE and EspB. The exceptions are EspA and EspC (10, 11) that are both part of the *espA* operon, which is located elsewhere in the genome. However, these genes are homologous to genes of the *esx-1* locus, i.e. *espE* and *espF*, respectively. A peculiar characteristic of ESX-1 substrates is that they are mutually dependent, meaning that their secretion is dependent on secretion of the other substrates (10). The secreted ESX proteins

contain a conserved WxG amino acid motif located between two α-helices (12). Recently, another conserved secretion signal present in all secreted protein pairs was identified. This C-terminal YxxxD/E motif can target proteins for secretion, but does not determine specificity to a particular type VII system (13). Therefore, it is still difficult to predict novel ESX-1 substrates.

In order to establish a successful infection, mycobacteria need regulatory mechanisms to express the right proteins at the right time. Different environments require specific transcriptional responses to successfully deal with the stress conditions encountered. During the first stages of infection, ESX-1-mediated protein secretion is one of the most important virulence mechanisms of pathogenic mycobacteria (4, 9, 10, 14, 15). Consequently, transcriptional regulation of *esx-1* and associated genes needs to be orchestrated tightly. The transcriptional regulator PhoP of the two-component system PhoPR positively regulates transcription of many *esx-1* associated genes, including the *espA* operon (16, 17). It has been proposed that PhoP regulation is dependent on environmental pH (18), which could indicate that the acidic environment of the phagosome induces *esx-1* gene transcription via PhoP leading to bacterial escape from this compartment. Other studies have shown that the *espA* operon is, in addition to PhoP, also regulated by the transcription factors EspR, MprAB and the repressors CRP and Lsr2, indicating that tight regulation of this operon is essential and, furthermore, suggesting that the *espA* operon may be regulated separately from the other ESX-1 substrates (19-22).

Since ESX-1 is crucial for virulence, inactivation of this secretion system would be expected to have a large impact on gene regulation processes in mycobacteria. Here, we apply RNA-seq and quantitative proteomics to determine gene expression and proteomic profiles of the pathogenic mycobacteria *M. marinum* and *M. tuberculosis* in absence of a functional ESX-1 secretion system. During short-term infection of macrophages, we observed highly

91 increased transcript levels of the *espA* operon. In contrast, during *in vitro* growth in culture
92 medium, transcription of most ESX-1 substrates and some putative new substrates was
93 decreased. Based on these gene transcription levels, we could confirm a regulatory role for the
94 putative transcriptional regulator WhiB6 in gene expression of ESX-1 substrates.
95

Results

Global features of the M. marinum esx-1 mutant transcriptome and proteome

In order to investigate the effect of ESX-1 disruption on gene expression and protein production, RNA and protein were extracted from three independent exponential phase cultures of *M. marinum* E11 strain and their isogenic *esx-1*-mutant during growth in 7H9 culture medium to characterize the transcriptome and proteome. Using transcriptomics (RNA-seq) and mass-spectrometry (MS) based proteomics with isobaric labeling for quantification, we captured the expression dynamics of transcripts and proteomes of the *esx-1* mutant. Data quality was assessed using Euclidean distance matrices for RNA (Figure S1) and principal component analysis (PCA) for protein (Figure S2), which demonstrated the degree of reproducibility between biological replicates. The total transcriptome and proteome data are recorded as an average of the clustered samples. A total of 823 genes passed our filter as differentially expressed (DE) as messenger RNA, of which 525 were classified as down-regulated and 298 as up-regulated (Figure 1A, Table S1). To determine parallel change in the protein level, 1,657 proteins were identified by the presence of 2 or more peptides, of which 576 proteins passed our filter and classified as differentially expressed. Of these, 412 proteins were found to be down-regulated and 164 were up-regulated (Table S2) and 482 protein-coding genes are shared and identified in both RNA-seq and quantitative proteomics datasets (Figure 1C).

The degree of global correlation between gene expression and protein abundance scores among the shared gene was relatively low (Figure S3A), which have been noted in other bacterial studies (23). However, within certain class of *M. marinum* functional categories (http://mycobrowser.epfl.ch/marinolist.html), the degree of correlation was much higher, with

R^2 exceeding 0.8 for lipid metabolism (Figure 1D), regulatory (Figure 1E) and conserved hypotheticals categories (Figure S3F). 28% of genes exhibiting differential expression at the RNA and protein level fell into Intermediary metabolism and respiration category, 18% for cell wall and cell process category, 15% for information pathways and 14% for lipid metabolism (Figure S4).

Transitional profiling analysis of the double auxotrophic *M. tuberculosis* $mc^2$6020 mutant strains and their isogenic *esx-1* mutants during growth was carried out to identify genes whose expression was dependent on the ESX-1 disruption (Figure 1B, Table S3).

Major effects of ESX1 mutation on genes encoding ESX-1 substrates and biosynthetic pathways

Analysis of differential expression (DE) identified changes in genes involved in a variety of cellular processes (Figure 2), although the majority of these top differentially regulated genes were associated with cell wall and cell processes and lipid metabolism functional categories. We noted that a substantial number of *esx-1*-associated genes were down-regulated in the mutant strains during growth in culture medium including 11 genes were located within or directly adjacent to the *esx-1* gene cluster (Figure 2, 3A). Among these down-regulated genes were those coding for known ESX-1 substrates, such as EsxA, EsxB, EspE and EspB. Remarkably, mRNA levels of core components of the ESX-1 secretion system, *i.e.* encoding members of the type VII secretion complex such as $EccB_1$, $EccD_1$, $EccE_1$ and $MycP_1$, remained unchanged, even though their respective genes are interspersed with genes encoding the ESX-1 substrates. In contrast to mRNA level, we noted strong increase of EsxA and EsxB at the protein level, probably reflecting the accumulation of these proteins in the cell due to the secretion defect (Figure 1F, 2). Our data also indicate a significant effect of *esx-1* disruption on genes associated with lipid metabolism, including synthesis of mycolic acids

(Figure 2). Strong down-regulation was observed at mRNA and protein levels for several polyketide synthases including genes involved in phthiocerol dimycocerosate synthesis and mycolic acid biosynthesis such as, *umaA*, *mmaA3, accD5, accD6,* and *pks15/1*, which are components of lipid biosynthesis (Figure 2, Table S1, S2). The change in *esx-1* and lipid metabolism-associated genes at mRNA and protein levels that we observed was not completely unexpected, since it has been reported previously that ESX-1-dependent protein secretion and mycolic acid synthesis are critically linked (24). However, we also noted a surprisingly broad impact of ESX-1 mutation on major biosynthetic pathways including ribosomal protein synthesis and DNA biosynthesis (Figure 2, table S1, S2). Down-regulation was observed at mRNA and protein levels for several ribosomal protein genes, DNA gyrase and a ribonucleotide-diphosphate reductase, which are components of protein and DNA biosynthesis, respectively. We identified also changes at both mRNA and protein levels in genes involved in general stress responses (*grpE, dnaK, groES, groEL1*), stress response regulation (*sigA, sigB, devS*), members of the WhiB family (*whiB2, whiB4, whiB6*) and several PE_PGRS genes (Figure 2). For *M. tuberculosis*, a similar trend was observed (Figure 2). In fact, of all genes, the *esx-1* genes encoding substrates EsxA, EsxB and EspK were the most significantly down-regulated in the mutant strain (Figure 2, Table S3). In contrast to *M. marinum* mutant, gene expression levels of $eccB_1$ and $eccD_1$ were also somewhat decreased in the *M. tuberculosis* mutant (Figure 2) On the other hand, the *esx-1* mutation did not seem to exert a significant effect on the expression of genes involved in lipid metabolism compared to *M. marinum* (Figure 2, Table S1, S3). Finally, a significant number of genes that are associated with information pathways including ribosomal protein genes were up-regulated at the mRNA level in the *esx-1* mutant (Figure 2). Taken together, the observed changes in the *esx-1* mutant transcriptome and proteome reflect the role of *esx-1* cluster employed by mycobacteria for major biosynthetic pathways.

171

172

173

174 ***Global Transcriptional Profiling of intraphagosomal M. marinum and the esx-1 mutant***

175 We next determined the effect of ESX-1 abrogation in *M. marinum* on gene transcription
176 during infection of primary macrophages. Using PMA-differentiated THP-1 cell line as a
177 model of primary macrophages, we analyzed the global gene expression of *M. marinum* wild-
178 type and the *esx-1* mutant after 6 h of infection. Wild-type mycobacteria can escape the
179 phagosome within a couple of hours after infection, whereas *esx-1* mutants are known to be
180 limited to the phagosomal compartment. The intraphagosomal transcriptome of *esx-1* mutant
181 was compared with the intracellular transcriptome of *M. marinum* wild type. Furthermore,
182 these intracellular transcriptomes were also compared with the transcriptome of *M. marinum*
183 wild-type grown in standard broth culture. We identified 720 ($p<0.05$) genes with significant
184 changes in expression after THP-1 infection in *esx-1* mutant compared to the wild type strain.
185 Of these, 465 genes were down-regulated and 255 genes were up-regulated (Table S4, Figure
186 S5). Remarkably, none of the genes within the *esx-1* region were significantly differentially
187 expressed in the *esx-1*-mutant as compared to the wild-type strain. However, we did find a
188 specific and pronounced increase in transcript levels of the *espA* operon in the
189 intraphagosomal transcriptome of *esx-1* mutant as compared with the in vitro transcriptomes
190 (Figure 3A). During growth in culture medium, mRNA levels of *espA* did not differ between
191 the wild type and *esx-1*-deficient *M. marinum*, which was confirmed by q-RT-PCR (Figure
192 3B). Therefore, these data suggest that proteins encoded by the *espA* operon, i.e. EspA, EspC
193 and EspD, play an important role in ESX-1-specific processes during the first stages of
194 macrophage infection. The *espA* operon was also somewhat induced in the wild-type bacteria
195 inside macrophages, albeit at a lower level. Perhaps this difference is due to the fact that wild-

196 type bacteria are able to escape from the phagosome, whereas the *esx-1* mutant population are
197 not.
198 Further analysis showed that a significant number of genes that code for functions of cell wall
199 and cell process were differentially regulated by intracellular *M. marinum* wild-type as well
200 ESX-1-deficient strain (Table S5, S6). *M. marinum* genes involved in mycolic acid synthesis,
201 phthiocerol dimycocerosate (PDIM) synthesis and transport to the cell surface, such as *fabG1*,
202 *accDs*, *ppsC, ppsD, pks11_1, pks13*, as well as genes coding for the polyketide synthases, and
203 the mycolic acid methyltransferase *umaA*, were differentially expressed during infection of
204 THP-1 cells (Figure 3C, D). Furthermore, *cpsY* a gene that encodes UDP-glucose-4-
205 epimerases and essential for linking of peptidoglycan and mycolic acid (25) had a pronounced
206 increase of its message in the intracellular *esx-1* mutant (Table S4, S5, S6, S7). We also found
207 many genes, such as *ftsE*, *ftsH*, *ftsW*, *murC*, and *murG* that are associated with cell division
208 and peptidoglycan assembly (26, 27), were down-regulated by intracellular bacteria (Table
209 S4, S5, S6, S7).
210 A significant number of genes that code for functions of lipid metabolism and metabolic
211 adaptation were differentially regulated in macrophages (Figure S6A). This subset includes
212 genes involved in fatty acid metabolism, such as isocitrate lyase (*icl*), an enzyme necessary
213 for the glyoxylate cycle and required for intracellular survival (28, 29), *pckA*, which encodes
214 the phosphoenolpyruvate carboxykinase and essential for mycobacterial survival in both
215 macrophages and mice (30, 31), energy metabolism (Figure S6B) and KstR-dependent
216 cholesterol regulon (Figure S6C), which is involved in lipid degradation and carbon
217 metabolism (32). We also observed a significant effect for a number of genes involved in
218 general stress reponses (*groES*, *groEL1*, *hsp*, *ahp*, *dnaK*), stress response regulation (*sigB,*
219 *devR, devS, hspR, kstR*), including members of the WhiB family (*whiB2*, *whiB3*, *whiB4*, *whiB,*
220 *whiB6*, *whiB7*) and alternative sigma factors (*sigE*, *sigL*, *sigM*), during the infection in *esx-1*

mutant. This pattern is illustrated in figure S6D and is probably linked with to stressful intraphagosomal conditions.

Different M. marinum esx-1 transposon mutants have similar gene transcription profiles

The ESX-1-deficient strain of *M. marinum* used for RNA sequencing contains a transposon in the *eccCb*$_1$ gene. To confirm that the gene transcription effects we observed were due to a defective ESX-1 system and not due to a side effect of this particular mutation, we analyzed several mutants containing transposon insertions in different genes from the *esx-1* gene cluster and compared mRNA levels of selected genes by quantitative RT-PCR. Our results show decreased transcript levels of the known ESX-1 substrate *esxA* and other *esx-1* secretion associated (*esp*) genes *espL*, *espK* and *espJ* for all tested *esx-1* mutants, whereas transcript levels of *eccD*$_1$, which encodes a structural component of the ESX-1 system, did not differ from wild-type *M. marinum* (Figure 4). These gene expression patterns in the *eccB*$_1$, *eccCa*$_1$, *eccD*$_1$ and *eccE*$_1$ transposon mutants are similar to the RNA sequencing results obtained for the *eccCb*$_1$ mutant. The only exception was that for the mutant containing a transposon insertion in *eccD*$_1$, we observed an increase of *eccD*$_1$ transcription itself and to a lesser extent of the adjacent gene *espJ* (Figure 4). However, this increase is most likely due to the presence of a strong promoter on the transposon that transcribes the kanamycin resistance cassette, as the measured mRNA is transcribed from sequences directly downstream of this promoter. Altogether, our results demonstrate that inactivation of the ESX-1 secretion system leads to a down-regulation in transcription of ESX-1 substrates and associated proteins.

ESX-1 substrate gene transcription is reduced by a regulatory mechanism

We next sought to determine the molecular mechanism underlying down-regulation of specific transcripts in *esx-1* mutant strains of *M. marinum*. It is possible that the decrease in

mRNA levels is due to a regulatory effect at the level of transcription. Alternatively, mRNA derived from specific sequences may be degraded via a post-transcriptional mechanism. To investigate these possibilities, we expressed an extra copy of the *espL* gene under control of a constitutively active promoter in the *M. marinum* wild type and *eccCb*$_1$ mutant strains and determined *espL* gene transcript levels. We found a similar increase in *espL* transcripts in both wild type and *eccCb*$_1$ mutant strains, indicating that degradation of specific mRNA is probably not the cause of the decreased mRNA levels in the mutant strain (Figure 5A). Expression levels of the downstream gene *espK* were not affected by the introduction of *espL*. These results indicate that there is a regulatory mechanism that prevents transcription of genes encoding ESX-1 substrates and associated proteins in absence of a functionally active ESX-1.

PE35 and PPE68 play an important role in ESX-1 secretion but not in gene regulation

Previously, PE35, which is located within the *esx-1* gene cluster, has been implicated in the regulation of *esxA/esxB* gene expression in *M. tuberculosis* (33). In contrast to this proposed function, the PE35/PPE68_1 protein pair in *M. marinum*, which coding genes have been duplicated from the *esx-1* cluster, is secreted via ESX-1 (34, 35). To determine whether PE35 plays a role in regulation of ESX-1 substrates, we overexpressed the *pe35/ppe68_1* operon in *M. marinum*. Interestingly, although there was no effect on gene transcription (Figure 5C), we did notice a substantial increase of EsxA secretion in the wild-type strain (Figure 5B). This increased EsxA secretion does not seem to represent a general increase in ESX-1 secretion, as protein levels of the cell surface localized EspE remained similar (Figure 5B). To study this effect in more detail, we introduced PE35 with a truncated version of PPE68_1 that only contained the PPE domain and was devoid of the C-terminal part. Although the introduced PE35 protein was expressed and secreted efficiently by ESX-1 (Figure 5B), levels of secreted EsxA were not increased, indicating that the C-terminal part of PPE68_1 plays a role in EsxA

secretion. To determine if secretion of the PE35/PPE68_1 protein pair itself was important for this process, we also determined the effect of removing the last 15 amino acids of the PE protein, containing the general secretion signal. This small deletion not only abolished secretion of the introduced PE35 protein, it also abolished EsxA secretion completely, despite the presence of an intact chromosomal copy of the *pe35/ppe68_1* operon (Figure 5B). This suggests that the truncated form of PE35 somehow interferes with EsxA secretion. Together these data show that, although PE35 and PPE68_1 do not seem to regulate the transcription of genes encoding ESX-1 substrates, they have a strong effect on EsxA, as was also observed previously (33).

Increasing EspI and EspG$_1$ does not lead to altered esx-1 gene expression

A second candidate protein that might regulate gene expression levels of ESX-1 substrates is EspI. The gene encoding this *esx-1* secretion associated protein of unknown function is located within the *esx-1* region and is down-regulated in *esx-1* mutants of both *M. marinum* and *M. tuberculosis* (Figure 2). In contrast to the other Esp proteins, EspI contains a putative nucleotide-binding domain. However, when we overexpressed this protein, we did not observe a change in down-regulation of *esx-1* associated gene transcription in the *M. marinum* *eccCb$_1$* transposon mutant, suggesting that EspI does not regulate this process (Figure 5C). We next focused on EspG$_1$ as a candidate *esx-1* gene regulator. EspG$_1$, which is a cytosolic protein that is not part of the membrane-bound secretion machinery, has recently been shown to interact specifically with PE35/PPE68_1 in *M. marinum* (34). It is conceivable that EspG$_1$ might function as a sensor that measures protein levels of intracellular ESX-1 substrates. When substrate levels are low, unbound EspG$_1$ may signal to induce gene expression. In absence of a functional ESX-1 system, accumulated PE35/PPE68_1 or other substrates may occupy EspG$_1$ leading to reduced transcription of *esx-1* associated genes. In order to

296 investigate the effect of $EspG_1$ on *esx-1* associated gene expression and protein levels, we
297 increased $EspG_1$ levels by overexpressing the *espG₁* gene in wild-type and ESX-1-deficient
298 *M. marinum*. However, this did not result in altered gene transcription (Figure 5C), nor ESX-1
299 protein secretion (data not shown). Together, our data shows that EspI and $EspG_1$ do not
300 appear to play a key role in *esx-1* associated gene regulation.
301
302 ***WhiB6 plays a role in transcription of ESX-1 substrates***
303 In addition to *espI*, also another gene encoding a putative regulatory protein was down-
304 regulated in *esx-1* mutant strains of both *M. marinum* and *M. tuberculosis*, *i.e. whiB6* (Figure
305 2). WhiB proteins are actinobacteria-specific regulators that contain iron-sulfur clusters and
306 are thought to act as redox-sensing transcription factors that can result in both gene activation
307 and repression (36). WhiB6 has been suggested to be involved in the regulation of EsxA
308 secretion (37), and later studies have confirmed this suggestion (38-40). In order to determine
309 whether Whib6 had an effect on expression levels of *esx-1* associated genes, we
310 overexpressed this protein in the ESX-1-deficient *M. marinum eccCb₁* transposon mutant
311 strain. We found that specifically those genes that were already down-regulated in the mutant
312 strain, such as *esxA* and *espK*, showed an even further transcriptional inhibition when *whib6*
313 levels were increased (Figure 5D). Furthermore, expression of *eccD1* was unaltered by *whib6*
314 overexpression, indicating that *whib6* is involved in transcription of ESX-1 substrates and
315 associated genes, but not of the system components. Surprisingly, *whiB6* is one of the genes
316 that is down-regulated upon abrogation of ESX-1-mediated protein secretion. Possibly,
317 WhiB6 is activated when the ESX-1 machinery is blocked and represses genes encoding
318 ESX-1 substrates as well as its own gene. Together, our data suggests that ESX-1 regulation is
319 even more complex than previously thought.
320

WhiB6 Is Required for Regulation of ESX-1 System

To determine whether WhiB6 is required for ESX-1 regulation, we constructed a deletion mutant of *whiB6* (*M. marinum* M^{USA} *-⊿whiB6*–and *M. marinum* M^{VU} *–⊿whiB6*). Analysis of gene expression identified 32 genes ($p<0.05$) exclude the *esx-1* locus genes showed a clear pattern that is significantly downregulated after *whiB6* gene has been knocked out (Figure 6A). Complementation of *M. marinum* M^{USA} *-⊿whiB6*–and *M. marinum* M^{VU} *–⊿whiB6* with the whiB6 gene on a mycobacterial shuttle plasmid reversed the downregulation of these genes to high expression level (Figure 6A, B). As expected, several genes that are associated with oxidative stress (*ahpC, ahpD, rebU*) were found in the DE gene pool. Also, the enrichment analysis of the associated Gene Ontology terms for the DE genes (*dnaB, dinP*) reveal that whiB6 may also regulate DNA replication or repair through regulating DNA-directed DNA polymerase and DNA helicase (Figure S7). However, many of the DE genes are hypothetical proteins and needed to be further characterized. Interestingly, *whiB7* is within the *whiB6-active* gene set, which implies that WhiB7 is active by, or works with WhiB6. Other than the *whiB6*-active gene set, 13 genes, which are involved in iron-sulfur cluster binding, cellular lipid metabolic processes are downregulated.

Remarkably, most of the genes within ESX-1 locus show an apparent co-expression with WhiB6 protein when *whiB6* has been knocked out (Figure 6C). These genes show a clear down-regulation while the expression was recovered when *whiB6* has been complemented in both strains (Figure 6C). The substrates of ESX-1 including EsxA, EsxB, EspE and EspF are highly regulated by WhiB6 which corroborates our previous result (Figure 6C).

Discussion

346 In this study, we have determined the transcriptome of the *M. marinum* E11 wild type and the
347 double auxotrophic *M. tuberculosis* $mc^2$6020 mutant strains and compared them with their
348 isogenic *esx-1* mutants. We found that during growth in 7H9 culture medium, genes encoding
349 ESX-1 substrates such as EsxA and other ESX-1-associated proteins were down regulated in
350 the mutant strains, whereas transcription of genes encoding structural components of the
351 ESX-1 system remained unaffected. This specific decrease in transcription might function as a
352 mechanism to avoid toxic accumulation of substrates. Interestingly, similar decrease of
353 substrate production has been shown for the ESX-5 secretion system, where the PE_PGRS
354 substrates do not accumulate intracellularly when secretion is blocked (41, 42). However, for
355 these PE_PGRS substrates the regulation was shown to be post-transcriptionally (42),
356 implying that a different mechanism is involved.

357 The most prominent change in gene expression that was observed upon host cell
358 infection by the *M. marinum esx-1* mutant strain was the increase in transcription of the *espA*
359 operon. The specific and pronounced transcriptional increase of this operon, and not of any
360 other *esx-1* associated gene, indicates that transcription of the *espA* operon is regulated
361 independently of the other substrates during infection. Previously, it has been shown that the
362 *espA* operon is regulated by different transcription and regulation factors, including EspR,
363 MprAB and PhoPR (20, 43, 44). Our new finding also suggests that EspA, EspC and EspD
364 are highly important for the bacteria during the early phase of infection. Since ESX-1 has
365 been shown to be responsible for mycobacterial escape from the phagosome, which occurs
366 within the first few hours of infection with *M. marinum* (6), the proteins produced by the *espA*
367 operon may play an important role in this process. Consequently, the avirulent phenotype of
368 ESX-1-deficient mycobacteria might be partly attributable to the inability to secrete EspA
369 and/or EspC early in infection.

In order to determine how ESX-1 substrate regulation is mediated, we overexpressed proteins that may have a regulatory function. Overexpression of the *esx-1* encoded EspI and $EspG_1$ did not have an effect on the lowered transcription of ESX-1 substrates in ESX-1-deficient *M. marinum*. The putative regulatory protein WhiB6 however, did affect transcription of those genes. While transcript levels of *whib6* itself was decreased in *esx-1* mutants of *M. marinum* and *M. tuberculosis*, increasing WhiB6 by overexpression resulted in a further decrease in ESX-1 substrate transcription in ESX-1-deficient *M. marinum*. This clearly indicates that WhiB6 is involved in ESX-1-associated gene regulation, as was also suggested previously (40) Indeed, there is accumulating evidence that WhiB proteins function as transcription factors, which may play a role in survival within the host (reviewed in (45)). Recently, also other groups have presented evidence to support a role of WhiB6 in regulating the transcription of *esx-1* genes (38-40).

A remarkable finding in this study was that overproduction of PE35/PPE68_1 resulted in a large increase in EsxA secretion. Previously, deletion of *M. tuberculosis* PE35 was already shown to abolish *esxA* transcription and secretion of its gene product (33). Now, we find that EsxA and PE35 secretion are linked, as an increase in PE35 secretion results in a concomitant increase in EsxA secretion. The fact that the C-terminus of PPE68_1 is required for this effect indicates that this is a specific process, which is supported by the fact that cell-surface localization of another ESX-1 substrate, EspE, is unaffected by overproduction of PE35/PPE68_1. Possibly, the PPE68 proteins serve as a chaperone to escort EsxA outside the bacterium, or it may be part of the secretion apparatus making secretion of specific substrates more efficient.

During *M. marinum* infection of human macrophages, we found that transcription of many *pe_pgrs* and *ppe* family genes were strongly down-regulated when ESX-1 function was abrogated. As much as 50% of all genes with decreased transcript levels in the *esx-1* mutant

strain belongs to one of these gene families (Table S4). It has to be noted that in the wild type strain transcription of *pe_pgrs* and *ppe* genes was decreased during infection, in comparison to growth in 7H9 medium (Table S5). As part of an adaptive response to the macrophage environment, expression of these cell wall localized proteins may be tuned down in order to evade immune recognition or to reduce cell permeability (46). The fact that in absence of a functional ESX-1 secretion system these genes are even further down-regulated, again suggests that there is a functional link or shared transcriptional pathways between ESX-1 and (some of the) PE_PGRS and PPE proteins, which are generally ESX-5 substrates (41).

Taken together, our results show that transcription of the *espA* locus plays an important role in ESX-1 mediated processes during the first hours of infection. Furthermore, we established a functional link between PE35 and EsxA secretion and lastly, we found that WhiB6 may play a regulatory role in transcription of ESX-1 substrates and associated genes.

Materials and Methods

Bacterial strains and growth conditions

The *esx-1* mutants of the *M. marinum* E11 wild type strain used in this study contain transposon insertions in $eccB_1$, $eccCa_1$, $eccCb_1$, $eccD_1$ and $eccE_1$ (47). For *M. tuberculosis*, the attenuated double deletion strains mc^26020 and mc^26030 of H37Rv were used, with deletions of *lysA* and *panCD* or *RD1* and *panCD*, respectively (48, 49). Bacterial strains were grown shaking at 30˚C (*M. marinum)* or 37°C (*M. tuberculosis*) in Middlebrook 7H9 culture medium, supplemented with 10% ADC (Albumin-Dextrose-Catalase, BD Biosciences) and

0.05% Tween-80. Culture medium containing the auxotrophic *M. tuberculosis* deletion strains was supplemented with 50 µg/ml panthothenic acid and, for $mc^2$6020, 100 µg/ml L-lysine.

Infection of human macrophages

THP-1 monocytes were cultured at 37˚C and 5% CO2 in RPMI-1640 with Glutamax-1 (Gibco) supplemented with 10% FBS, 100 µg ml^{-1} streptomycin and 100 U ml^{-1} penicillin. Cells were seeded at a density of 3 x 10^7 cells per T175 flask and differentiated into macrophages by 48 hours of incubation with 25 ng/ml PMA (Sigma-Aldrich). 1,8 x 10^8 THP-1 cells were infected with *M. marinum* at an MOI of 20 for 2 hours, after which cells were washed with PBS to remove extracellular bacteria. After 4 additional hours of infection at 33°C, THP-1 cells were lysed with 1% Triton X-100. After a low speed centrifugation step to remove cellular debris, mycobacteria were pelleted after which RNA was extracted as described in the following section.

Genomic sequence

We sequenced the *M.marinum* E11 strain with the PacBio RSII single-molecule real-time (SMRT) sequencing technology (50). The raw reads were assembled into two pieces (the core and the plasmid) with HGAP assembler (51) using the default parameters. The sequence was improved with ICORN2 (52) with three iterations, correcting 20 single base pair errors and 61 insertions and deletions. To transfer the annotation from the current reference, we used RATT (53) with the PacBio parameter. Gene models around gaps were manually improved on the new sequence. The updated genome annotation was resubmitted under the same accession numbers (HG917972 for *M. marinum E11* main chromosome genome and HG917973 for *M. marinum E11* plasmid pRAW, complete sequence).

444 ***RNA extraction and qRT-PCR***

445 *M. marinum* and *M. tuberculosis* cultures were pelleted and bead-beated in 1 ml Trizol
446 (Invitrogen) with 0.1 mm Zirconia/Silica beads (Biospec Products). After centrifugation,
447 supernatants were extracted with chloroform and RNA was precipitated with isopropanol.
448 RNA pellets were washed with 80% ethanol and dissolved in RNAse-free water.
449 Contaminating DNA was removed by incubation with DNAse I (Fermentas). For RT-PCR,
450 cDNA was generated using a SuperScript VILO cDNA synthesis kit (Invitrogen). An
451 equivalent of 5 ng of RNA was used in the quantitative PCR reactions. Q-RT-PCR was
452 performed using SYBR GreenER (Invitrogen) and the LightCycler 480 (Roche). Transcript
453 levels were normalized to the housekeeping gene *sigA* (54), using ΔΔCt analysis. All primer
454 sequences used for q-RT-PCR are listed in Table S8.
455
456 ***RNA preparation for Illumina Sequencing***
457 Total RNA was extracted with Trizol (Invitrogen) and then purified on RNeasy spin columns
458 (Qiagen) according to the manufacturer's instructions. The RNA integrity (RNA Integrity
459 Score ≥ 6.8) and quantity was determined on the Agilent 2100 Bioanalyzer (Agilent; Palo
460 Alto, CA, USA). As ribosomal RNA comprises the vast majority of the extracted RNA
461 population, depletion of these molecules through RiboMinus-based rRNA depletion was used.
462 For this mRNA enrichment, the Invitrogen's RiboMinus™ Prokaryotic kit was used according
463 to manufacturer's instructions. Briefly, 2 µg of total RNA samples was hybridized with
464 prokaryotic rRNA sequence-specific 5′-biotin labeled oligonucleotide probes to selectively
465 deplete large rRNA molecules from total RNA. Then, these rRNA-hybridized, biotinylated
466 probes were removed from the sample with streptavidin-coated magnetic beads. The resulting
467 RNA sample was concentrated using the RiboMinus™ concentrate module according to the
468 manufacturer's protocol. The final RiboMinus™ RNA sample was subjected to thermal

mRNA fragmentation using Elute, Prime, and fragment Mix from the Illumina TruSeq™ RNA sample preparation kit v2 (Low-Throughput protocol). The fragmented mRNA samples were subjected to cDNA synthesis using the Illumina TruSeq™ RNA sample preparation kit (Low-Throughput protocol) according to manufacturer's protocol. Briefly, cDNA was synthesized from enriched and fragmented RNA using SuperScript III Reverse Transcriptase (Invitrogen) and SRA RT primer (Illumina). The cDNA was further converted into double stranded DNA using the reagents supplied in the kit, and the resulting dsDNA was used for library preparation. To this end, cDNA fragments were end-repaired and phosphorylated, followed by adenylation of 3′ends and adapter ligation. Twelve cycles of PCR amplification were then performed, and the library was finally purified with AMPure beads (Beckman Coulter) as per the manufacturer's instructions. A small aliquot (1 µl) was analyzed on Invitrogen Qubit and Agilent Bioanalyzer. The bar-coded cDNA libraries were pooled together in equal concentrations in one pool before sequencing on Illumina HiSeq2000 using the TruSeq SR Cluster Generation Kit v3 and TruSeq SBS Kit v3. Data were processed with the Illumina Pipeline Software v1.82.

RNA-Seq analysis

The Illumina reads were mapped with smalt (default parameter) against the new PacBio reference. From the read count, obtain with bedtools ((55) parameter multicov) -D to include duplicates and -q 5 to exclude repetitive mapping reads), we perform a differential expression analysis with DESeq (56), default parameter.

Plasmid construction

The *E.coli* mycobacterial shuttle vector pSMT3 was used for construction of all plasmids. To overexpress PE35-PPE68_1 (MMARE11_01740- MMARE11_01750), we used a previously

494 described plasmid (13). For construction of the plasmid containing *espG$_1$*, this gene was
495 amplified from the *M. marinum* E11 genome by PCR using primers containing NheI and
496 EcoRV restriction sites and a 3' HA-epitope. The resulting PCR product and empty pSMT3
497 were digested with NheI and EcoRV followed by ligation of *espG$_1$* into the vector by T4
498 ligase (Fermentas). For the construction of the plasmid containing *whib6*, this gene was
499 amplified from the *M. marinum* E11 genome by PCR using primers containing NheI and
500 BamHI restriction sites. For the other construct, *espI* was amplified from the *M. marinum* E11
501 genome by PCR using primers containing NheI and BglII restriction site. PCR product was
502 digested with NheI and BamHI. Empty pSMT3 was digested with NheI and BamHI after
503 which PCR products were ligated in the vector. All plasmids were introduced in the *M.*
504 *marinum* wild-type E11 and its isogenic *eccCb$_1$* mutant strain by electroporation. All primer
505 sequences are listed in Table S8.
506
507 ***Analysis of protein expression and secretion***
508 *M. marinum* cultures were grown to mid-logarithmic phase in 7H9 culture medium
509 supplemented with 0.2% glycerol and 0.2% dextrose. Bacteria were pelleted, washed in PBS
510 and incubated in 0.5% Genapol X-080 (Sigma-Aldrich) for 30 minutes to extract cell wall
511 proteins. Genapol X-080-treated *M. marinum* cells were disrupted by sonication. Secreted
512 proteins were precipitated from the culture supernatant by 10% trichloroacetic acid (TCA,
513 Sigma-Aldrich). Proteins were separated by molecular weight on 15% SDS-PAGE gels and
514 subsequently transferred to nitrocellulose membranes (Amersham Hybond ECL, GE
515 Healthcare Life Sciences). Immunostaining was performed with mouse monoclonal
516 antibodies directed against the HA-epitope (HA.11, Covance), EsxA (Hyb76-8), or rabbit
517 polyclonal sera recognizing EspE (57).
518

519 ***LC-MS analysis***

520 Peptide preparation from *M. marinum* E11 strain and their isogenic *esx-1*-mutant was
521 performed as previously described (58). Approximately 100 μg protein digests of each sample
522 were labeled with 4plex iTRAQ reagents (Applied Biosystems). The combined iTRAQ-
523 labeled samples were fractionated using strong cation exchange chromatography. The eluted
524 fractions were dried and desalted using a Sep-Pak C-18 SPE cartridge (Waters, Milford, MA,
525 USA). LC-MS analysis as well as MS data processing was carried out following our
526 published procedure (59). Briefly, each fraction was analyzed three times using LTQ-Orbitrap
527 Velos (Thermo Scientific). The MS spectra were recorded in the Orbitrap whereas MS2
528 spectra were recorded in the c-TRAP for HCD fragmentation and in the LTQ for the CID
529 fragmentation. Both HCD and CID spectra were extracted separately using Proteome
530 Discoverer software and processed by in-house script before Mascot search against *M.*
531 *marinum* E11 strain. The mascot results (.dat file) were processed by Scaffold software for
532 validation of protein identification and quantitative assessment. For protein identification,
533 Scaffold local false positive rates (FDR) were controlled below 1% for both protein and
534 peptide identifications (0.91% and 0.9% for peptide and protein respectively for this dataset).
535 Protein quantitation was processed using Scaffold Q+ which was based on i-Tracker
536 algorithm (60). The iTRAQ quantitation using HCD is highly accurate and a greater than 2-
537 fold change were considered as significantly differential expression in this study.

538

539 **Accession codes.**

540 Sequencing reads have been submitted to the EMBL-EBI European Nucleotide Archive
541 (ENA) Sequence Read Archive (SRA) under the study accession PRJEB8560.

542

543 **Acknowledgements**

We thank Astrid van der Sar and Esther Stoop for providing the *M. marinum E11* ESX-1 mutants. Work in AP's laboratory is supported by KAUST faculty baseline fund (BAS/1/1020-01-01). Authors thank members of the Bioscience Core Lab (BCL) in KAUST for sequencing the RNA-seq libraries on the Illumina Hiseq platform and for running protein samples for the quantitative proteomics workflow with the LTQ-Orbitrap Velos (Thermo Scientific).

References

1. Lewis KN, Liao R, Guinn KM, Hickey MJ, Smith S, Behr MA, et al. Deletion of RD1 from Mycobacterium tuberculosis mimics bacille Calmette-Guerin attenuation. The Journal of infectious diseases. 2003;187(1):117-23.
2. Pym AS, Brodin P, Brosch R, Huerre M, Cole ST. Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti. Mol Microbiol. 2002;46(3):709-17.
3. Stamm LM, Morisaki JH, Gao LY, Jeng RL, McDonald KL, Roth R, et al. Mycobacterium marinum escapes from phagosomes and is propelled by actin-based motility. Journal of Experimental Medicine. 2003;198(9):1361-8.
4. van der Wel N, Hava D, Houben D, Fluitsma D, van Zon M, Pierson J, et al. M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. Cell. 2007;129(7):1287-98.
5. de Jonge MI, Pehau-Arnaudet G, Fretz MM, Romain F, Bottai D, Brodin P, et al. ESAT-6 from Mycobacterium tuberculosis dissociates from its putative chaperone CFP-10 under acidic conditions and exhibits membrane-lysing activity. Journal of bacteriology. 2007;189(16):6028-34.
6. Houben D, Demangel C, van Ingen J, Perez J, Baldeon L, Abdallah AM, et al. ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria. Cellular microbiology. 2012;14(8):1287-98.
7. De Leon J, Jiang G, Ma Y, Rubin E, Fortune S, Sun J. Mycobacterium tuberculosis ESAT-6 exhibits a unique membrane-interacting activity that is not found in its ortholog from non-pathogenic Mycobacterium smegmatis. The Journal of biological chemistry. 2012;287(53):44184-91.
8. Coros A, Callahan B, Battaglioli E, Derbyshire KM. The specialized secretory apparatus ESX-1 is essential for DNA transfer in Mycobacterium smegmatis. Mol Microbiol. 2008;69(4):794-808.
9. Gao LY, Guo S, McLaughlin B, Morisaki H, Engel JN, Brown EJ. A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. Mol Microbiol. 2004;53(6):1677-93.
10. Fortune SM, Jaeger A, Sarracino DA, Chase MR, Sassetti CM, Sherman DR, et al. Mutually dependent secretion of proteins required for mycobacterial virulence. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(30):10676-81.

11. MacGurn JA, Raghavan S, Stanley SA, Cox JS. A non-RD1 gene cluster is required for Snm secretion in Mycobacterium tuberculosis. Mol Microbiol. 2005;57(6):1653-63.
12. Pallen MJ. The ESAT-6/WXG100 superfamily -- and a new Gram-positive secretion system? Trends Microbiol. 2002;10(5):209-12.
13. Daleke MH, Ummels R, Bawono P, Heringa J, Vandenbroucke-Grauls CM, Luirink J, et al. General secretion signal for the mycobacterial type VII secretion pathway. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(28):11342-7.
14. McLaughlin B, Chon JS, MacGurn JA, Carlsson F, Cheng TL, Cox JS, et al. A mycobacterium ESX-1-secreted virulence factor with unique requirements for export. Plos Pathog. 2007;3(8):e105.
15. Tan T, Lee WL, Alexander DC, Grinstein S, Liu J. The ESAT-6/CFP-10 secretion system of Mycobacterium marinum modulates phagosome maturation. Cellular microbiology. 2006;8(9):1417-29.
16. Asensio JA, Arbues A, Perez E, Gicquel B, Martin C. Live tuberculosis vaccines based on phoP mutants: a step towards clinical trials. Expert opinion on biological therapy. 2008;8(2):201-11.
17. Walters SB, Dubnau E, Kolesnikova I, Laval F, Daffe M, Smith I. The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. Mol Microbiol. 2006;60(2):312-30.
18. Abramovitch RB, Rohde KH, Hsu FF, Russell DG. aprABC: a Mycobacterium tuberculosis complex-specific locus that modulates pH-driven adaptation to the macrophage phagosome. Mol Microbiol. 2011;80(3):678-94.
19. Gordon BRG, Li YF, Wang LR, Sintsova A, van Bakel H, Tian SH, et al. Lsr2 is a nucleoid-associated protein that targets AT-rich sequences and virulence genes in Mycobacterium tuberculosis (vol 107, pg 5154, 2010). Proceedings of the National Academy of Sciences of the United States of America. 2010;107(43):18741-.
20. Pang XH, Samten B, Cao GX, Wang XS, Tvinnereim AR, Chen XL, et al. MprAB Regulates the espA Operon in Mycobacterium tuberculosis and Modulates ESX-1 Function and Host Cytokine Response. Journal of bacteriology. 2013;195(1):66-75.
21. Raghavan S, Manzanillo P, Chan K, Dovey C, Cox JS. Secreted transcription factor controls Mycobacterium tuberculosis virulence. Nature. 2008;454(7205):717-21.
22. Rickman L, Scott C, Hunt DM, Hutchinson T, Menendez MC, Whalan R, et al. A member of the cAMP receptor protein family of transcription regulators in Mycobacterium tuberculosis is required for virulence in mice and controls transcription of the rpfA gene coding for a resuscitation promoting factor. Mol Microbiol. 2005;56(5):1274-86.
23. Waldbauer JR, Rodrigue S, Coleman ML, Chisholm SW. Transcriptome and proteome dynamics of a light-dark synchronized bacterial cell cycle. Plos One. 2012;7(8):e43432.
24. Joshi SA, Ball DA, Sun MG, Carlsson F, Watkins BY, Aggarwal N, et al. EccA1, a component of the Mycobacterium marinum ESX-1 protein virulence factor secretion pathway, regulates mycolic acid lipid synthesis. Chem Biol. 2012;19(3):372-80.
25. Weston A, Stern RJ, Lee RE, Nassau PM, Monsey D, Martin SL, et al. Biosynthetic origin of mycobacterial cell wall galactofuranosyl residues. Tubercle and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 1997;78(2):123-31.
26. Pastoret S, Fraipont C, den Blaauwen T, Wolf B, Aarsman ME, Piette A, et al. Functional analysis of the cell division protein FtsW of Escherichia coli. Journal of bacteriology. 2004;186(24):8370-9.
27. Schmidt KL, Peterson ND, Kustusch RJ, Wissel MC, Graham B, Phillips GJ, et al. A predicted ABC transporter, FtsEX, is needed for cell division in Escherichia coli. Journal of bacteriology. 2004;186(3):785-93.

636 28. McKinney JD, Honer zu Bentrup K, Munoz-Elias EJ, Miczak A, Chen B, Chan WT, et
637 al. Persistence of Mycobacterium tuberculosis in macrophages and mice requires the
638 glyoxylate shunt enzyme isocitrate lyase. Nature. 2000;406(6797):735-8.
639 29. Munoz-Elias EJ, McKinney JD. Mycobacterium tuberculosis isocitrate lyases 1 and 2
640 are jointly required for in vivo growth and virulence. Nature medicine. 2005;11(6):638-44.
641 30. Liu KY, Yu JZ, Russell DG. pckA-deficient Mycobacterium bovis BCG shows
642 attenuated virulence in mice and in macrophages. Microbiol-Sgm. 2003;149:1829-35.
643 31. Marrero J, Rhee KY, Schnappinger D, Pethe K, Ehrt S. Gluconeogenic carbon flow of
644 tricarboxylic acid cycle intermediates is critical for Mycobacterium tuberculosis to establish
645 and maintain infection. Proceedings of the National Academy of Sciences of the United States
646 of America. 2010;107(21):9819-24.
647 32. Kendall SL, Withers M, Soffair CN, Moreland NJ, Gurcha S, Sidders B, et al. A
648 highly conserved transcriptional repressor controls a large regulon involved in lipid
649 degradation in Mycobacterium smegmatis and Mycobacterium tuberculosis. Mol Microbiol.
650 2007;65(3):684-99.
651 33. Brodin P, Majlessi L, Marsollier L, de Jonge MI, Bottai D, Demangel C, et al.
652 Dissection of ESAT-6 system 1 of Mycobacterium tuberculosis and impact on
653 immunogenicity and virulence. Infection and immunity. 2006;74(1):88-98.
654 34. Daleke MH, van der Woude AD, Parret AHA, Ummels R, de Groot AM, Watson D, et
655 al. Specific Chaperones for the Type VII Protein Secretion Pathway. Journal of Biological
656 Chemistry. 2012;287(38):31939-47.
657 35. Sani M, Houben EN, Geurtsen J, Pierson J, de Punder K, van Zon M, et al. Direct
658 visualization by cryo-EM of the mycobacterial capsular layer: a labile structure containing
659 ESX-1-secreted proteins. Plos Pathog. 2010;6(3):e1000794.
660 36. Zheng F, Long Q, Xie J. The function and regulatory network of WhiB and WhiB-like
661 protein from comparative genomics and systems biology perspectives. Cell Biochem Biophys.
662 2012;63(2):103-8.
663 37. Das C, Ghosh TS, Mande SS. Computational analysis of the ESX-1 region of
664 Mycobacterium tuberculosis: insights into the mechanism of type VII secretion system. Plos
665 One. 2011;6(11):e27980.
666 38. Bosserman RE, Nguyen TT, Sanchez KG, Chirakos AE, Ferrell MJ, Thompson CR, et
667 al. WhiB6 regulation of ESX-1 gene expression is controlled by a negative feedback loop in
668 Mycobacterium marinum. Proceedings of the National Academy of Sciences of the United
669 States of America. 2017;114(50):E10772-E81.
670 39. Chen Z, Hu Y, Cumming BM, Lu P, Feng L, Deng J, et al. Mycobacterial WhiB6
671 Differentially Regulates ESX-1 and the Dos Regulon to Modulate Granuloma Formation and
672 Virulence in Zebrafish. Cell Rep. 2016;16(9):2512-24.
673 40. Solans L, Aguilo N, Samper S, Pawlik A, Frigui W, Martin C, et al. A Specific
674 Polymorphism in Mycobacterium tuberculosis H37Rv Causes Differential ESAT-6
675 Expression and Identifies WhiB6 as a Novel ESX-1 Component. Infection and immunity.
676 2014;82(8):3446-56.
677 41. Abdallah AM, Verboom T, Weerdenburg EM, van Pittius NCG, Mahasha PW,
678 Jimenez C, et al. PPE and PE_PGRS proteins of Mycobacterium marinum are transported via
679 the type VII secretion system ESX-5. Mol Microbiol. 2009;73(3):329-40.
680 42. Ates LS, Dippenaar A, Ummels R, Piersma SR, van der Woude AD, van der Kuij K,
681 et al. Mutations in ppe38 block PE_PGRS secretion and increase virulence of Mycobacterium
682 tuberculosis. Nat Microbiol. 2018;3(2):181-8.
683 43. Cao G, Howard ST, Zhang P, Wang X, Chen XL, Samten B, et al. EspR, a regulator of
684 the ESX-1 secretion system in Mycobacterium tuberculosis, is directly regulated by the two-
685 component systems MprAB and PhoPR. Microbiology. 2015;161(Pt 3):477-89.

686 44. Hunt DM, Sweeney NP, Mori L, Whalan RH, Comas I, Norman L, et al. Long-range
687 transcriptional control of an operon necessary for virulence-critical ESX-1 secretion in
688 Mycobacterium tuberculosis. Journal of bacteriology. 2012;194(9):2307-20.
689 45. den Hengst CD, Buttner MJ. Redox control in actinobacteria. Bba-Gen Subjects.
690 2008;1780(11):1201-16.
691 46. Ates LS, van der Woude AD, Bestebroer J, van Stempvoort G, Musters RJ, Garcia-
692 Vallejo JJ, et al. The ESX-5 System of Pathogenic Mycobacteria Is Involved In Capsule
693 Integrity and Virulence through Its Substrate PPE10. Plos Pathog. 2016;12(6):e1005696.
694 47. van der Sar AM, Abdallah AM, Sparrius M, Reinders E, Vandenbroucke-Grauls
695 CMJE, Bitter W. Mycobacterium marinum strains can be divided into two distinct types based
696 on genetic diversity and virulence. Infection and immunity. 2004;72(11):6306-12.
697 48. Sambandamurthy VK, Derrick SC, Hsu T, Chen B, Larsen MH, Jalapathy KV, et al.
698 Mycobacterium tuberculosis DeltaRD1 DeltapanCD: a safe and limited replicating mutant
699 strain that protects immunocompetent and immunocompromised mice against experimental
700 tuberculosis. Vaccine. 2006;24(37-39):6309-20.
701 49. Sambandamurthy VK, Derrick SC, Jalapathy KV, Chen B, Russell RG, Morris SL, et
702 al. Long-term protection against tuberculosis following vaccination with a severely attenuated
703 double lysine and pantothenate auxotroph of Mycobacterium tuberculosis. Infection and
704 immunity. 2005;73(2):1196-203.
705 50. Levene MJ, Korlach J, Turner SW, Foquet M, Craighead HG, Webb WW. Zero-mode
706 waveguides for single-molecule analysis at high concentrations. Science.
707 2003;299(5607):682-6.
708 51. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid,
709 finished microbial genome assemblies from long-read SMRT sequencing data. Nature
710 methods. 2013;10(6):563-+.
711 52. Otto TD, Sanders M, Berriman M, Newbold C. Iterative Correction of Reference
712 Nucleotides (iCORN) using second generation sequencing technology. Bioinformatics.
713 2010;26(14):1704-7.
714 53. Otto TD, Dillon GP, Degrave WS, Berriman M. RATT: Rapid Annotation Transfer
715 Tool. Nucleic Acids Res. 2011;39(9).
716 54. Manganelli R, Dubnau E, Tyagi S, Kramer FR, Smith I. Differential expression of 10
717 sigma factor genes in Mycobacterium tuberculosis. Mol Microbiol. 1999;31(2):715-24.
718 55. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic
719 features. Bioinformatics. 2010;26(6):841-2.
720 56. Anders S, Huber W. Differential expression analysis for sequence count data. Genome
721 biology. 2010;11(10).
722 57. Carlsson F, Joshi SA, Rangell L, Brown EJ. Polar localization of virulence-related
723 Esx-1 secretion in mycobacteria. Plos Pathog. 2009;5(1):e1000285.
724 58. Abdallah AM, Hill-Cawthorne GA, Otto TD, Coll F, Guerra-Assuncao JA, Gao G, et
725 al. Genomic expression catalogue of a global collection of BCG vaccine strains show
726 evidence for highly diverged metabolic and cell-wall adaptations. Sci Rep. 2015;5:15443.
727 59. Zhu J, Zhang HM, Guo TN, Li WY, Li HY, Zhu Y, et al. Quantitative proteomics
728 reveals differential biological processes in healthy neonatal cord neutrophils and adult
729 neutrophils. Proteomics. 2014;14(13-14):1688-97.
730 60. Shadforth IP, Dunkley TPJ, Lilley KS, Bessant C. i-Tracker: For quantitative
731 proteomics using iTRAQ (TM). BMC genomics. 2005;6.
732 61. Mohn WW, van der Geize R, Stewart GR, Okamoto S, Liu J, Dijkhuizen L, et al. The
733 actinobacterial mce4 locus encodes a steroid transporter. The Journal of biological chemistry.
734 2008;283(51):35368-74.
735

736
737 **Figure legends**

738

739 Figure 1: Global features of the *M. marinum* & *M. tuberculosis* esx-1 mutant transcriptome
740 and proteome
741 Volcano plots obtained from RNA-Seq analysis of wild-type *M. marinum* E11 strain vs. the
742 *eccCb$_1$* transposon mutant (A) and *M. tuberculosis* mc^{2}6020 vs. the ESX-1 mutant strain (B).
743 Red dots indicate statistical significance ($q < 0.05$) and black dots indicate a lack of statistical
744 significance. Selected genes that are most down or up-regulated in the *ESX-1* mutant strains
745 are highlighted. (C) Venn diagram of the number of differentially expressed transcripts and
746 proteins quantified using RNA-seq and quantitative proteomics respectively. Scatterplots of
747 the relationship between differentially expressed genes of *M. marinum eccCb$_1$* transposon
748 mutant compared to its isogenic wild-type strain E11, quantified in both data sets and
749 classified by (D) lipid metabolism, (E) regulatory proteins and (F) cell wall and cell processes
750 categories. Scatterplots and bar chart display the rectilinear equation and coefficient of
751 determination (R^2).

752

753 Figure 2: Most differentially expressed genes of *M. marinum* and *M. tuberculosis*, when
754 grown in culture medium, grouped into broad functional categories. Within each group, genes
755 are ranked in ascending order by *p*-value. (Red) Top-100 annotated *M. marinum* E11 genes
756 that show the most differentially expressed in *M. marinum eccCb$_1$* transposon mutant
757 compared to its isogenic wild-type strain E11 during growth in 7H9 culture medium. Bar
758 chart of log2 fold change for individual genes (RNA, blue; protein, red; locus tags, outer).
759 (Green) Top-100 annotated *M. tuberculosis* genes that show the most differentially expressed
760 in auxotrophic *M. tuberculosis* RD1 deletion mutant strain mc^{2}6030 compared to its isogenic
761 control strain mc^{2}6020 during growth in 7H9 culture medium. Bar chart of log2 fold change

762 for individual genes. The genes Rv3872-Rv3878 are not included as these genes are deleted in
763 the RD1 mutant strain.
764
765 Figure 3. Effect of ESX-1 disruption (*eccCb*$_1$ transposon mutant) on gene transcription during
766 infection (indicated as ínt') and growth in culture medium in *M. marinum* compared to wild-
767 type strain E11 during growth in 7H9 culture medium. (A) Relative transcript expression
768 levels of ESX-1 secretion system-associated genes including the main ESX-1 locus as well as
769 the EspR regulator and accessory factors EspA operon encoded outside the RD-1 region. (B)
770 Gene expression levels, as measured by q-RT-PCR, Gene expression levels were compared to
771 those of the wild-type strain E11 grown in similar conditions. Values represent mean ±
772 standard error of two biological replicates. (C, D) Regulation of cell wall synthesis including
773 genes involved in mycolic acid synthesis (C) and phthiocerol dimycocerosates (PDIM) (D).
774
775 Figure 4. *esx-1* transposon mutants have similar gene transcription profiles. Gene expression
776 levels for *M. marinum eccB*$_1$, *eccCa*$_1$, *eccCb*$_1$, *eccD*$_1$ and *eccE*$_1$ transposon mutants as
777 measured by q-RT-PCR. All strains were grown in 7H9 culture medium and gene expression
778 levels were compared to those of the wild-type strain E11. Values represent mean ± standard
779 error of at least three biological replicates.
780
781 Figure 5. Regulation of ESX-1 secretion system. (A) Down-regulation of *espL* is the result of
782 a regulatory process. A functional copy of *espL* was introduced in *M. marinum* wild type and
783 *eccCb1* mutant strains and espK and espL gene expression levels were measured by qRT-
784 PCR. Gene expression levels were compared to those of the wild type strain E11. Values
785 represent mean ± standard error of two biological replicates. (B) Introduction of
786 PE35/PPE68_1 results in increased EsxA secretion but not in gene regulation. Pellet (p), cell

787 wall extract (cw) and supernatant (s) fractions of *M. marinum* wild type and *eccCb1* mutant
788 strains expressing PE35/PPE68_1, PE35/PPE68 containing a C-terminal deletion of PPE68_1,
789 or PE35/PPE68_1 containing a 15-amino acid C-terminal deletion of PE35, were analyzed for
790 the presence of EspE, EsxA and the introduced PE35 by Western blot. Fractions represent 0.5,
791 1 or 2 OD units of original culture, respectively. In all cases, PE35 contained a C-terminal
792 HA-tag. (C) EspG1, EspI and PE35/PPE68_1 do not regulate transcription of selected esx-1
793 (associated) genes. EspG1, EspI or PE35/PPE68_1 were overexpressed in the *M. marinum*
794 eccCb1 mutant strain and expression levels of *espK, espL, esxA, pe_pgrs1* and *eccD1* were
795 measured by q-RT-PCR. Gene expression levels were compared to those of the wild-type
796 strain E11. Values represent mean ± standard error of at least two biological replicates. (D)
797 WhiB6 is involved in transcriptional regulation of ESX-1 substrates and associated genes. The
798 *whib6* gene was overexpressed in the *M. marinum eccCb1* mutant strain and transcript levels
799 of *espK, espL, esxA, pe_pgrs1* and *eccD1* were measured by q-RT-PCR. Gene expression
800 levels were compared to those of the *eccCb1* mutant strain. Values represent mean ± standard
801 error of two biological replicates.
802
803 Figure 6. Expression profile of the complementary strains (*M. marinum* M^{USA} –
804 Complementary and *M. marinum* M^{VU} –Complementary) and knock-out strains (*M. marinum*
805 M^{USA} -Δ*whiB6*–and *M. marinum* M^{VU} –Δ*whiB6*) compared with the corresponding control
806 strains (*M. marinum* M^{USA} –Empty Vector strain and *M. marinum* M^{VU} –Empty Vector). The
807 whiB6-activated gene set expression heat map is shown in (A), the ESX-1 locus expression is
808 shown in (B) and (C) shows the whiB6-repressed gene set.
809
810 **Supplementary table and figure legends**
811

812 **Table S1.**

813 Complete list of genes that are significantly changed in the *M. marinum eccCb$_1$* transposon
814 mutant compared to its isogenic wild-type strain E11 during growth in 7H9 culture medium.
815 P<0.05.

816

817 **Table S2.**

818 Complete list of proteins of which expression is changed in *M. marinum eccCb$_1$* transposon
819 mutant compared to its isogenic wild-type strain E11 during growth in 7H9 culture medium.
820 Protein with greater than 2-fold change was considered as significantly differentially
821 expression

822

823 **Table S3.**

824 Complete list of genes that are significantly changed in the auxotrophic *M. tuberculosis* RD1
825 deletion mutant strain mc^{2}6030 compared to its isogenic control strain mc^{2}6020 during
826 growth in 7H9 culture medium. P<0.05.

827

828 **Table S4.**

829 Complete list of genes that are significantly changed in the *M. marinum eccCb$_1$* transposon
830 mutant compared to the wild-type strain E11 during infection of human THP-1 macrophages.
831 P<0.05.

832

833 **Table S5.**

834 Complete list of genes that are significantly changed in the *M. marinum* wild-type strain
835 during infection of macrophages compared to growth in 7H9 culture medium. P<0.05.

836

837 **Table S6.**

838 Complete list of genes that are significantly changed in the *M. marinum eccCb$_1$* transposon
839 mutant strain during infection of macrophages compared to the wild-type strain E11 during
840 growth in 7H9 culture medium. P<0.05.

841

842 **Table S7.**

843 Complete list of genes that are significantly changed in the *M. marinum eccCb$_1$* transposon
844 mutant strain during infection of macrophages compared to growth in 7H9 culture medium.
845 P<0.05.

846

847 **Table S8.**

848 Primers used in this study. Restriction sites showed in bold.

849

850 **Figure S1**

851 Euclidean distance matrices of RNA-seq transcriptome data showing clustering of *M.*
852 *marinum* wild-type (E11) and *eccCb$_1$* transposon mutant samples (ESX-1) grown in culture
853 medium (three biological replicates) or during infection of THP-1 cells (indicated as 'int').

854

855 **Figure S2**

856 Principal component cluster analysis (PCA) of biological replicates of proteome data showing
857 clustering of *M. marinum* wild type (E11) and *eccCb$_1$* transposon mutant samples (ESX-1).
858 PCA mapping showed clustering of biological replicates from E11 wild type and *esx-1*
859 mutant.

860

861 **Figure S3**

Correlation between protein and mRNA expression of *M. marinum eccCb$_1$* transposon mutant compared to its isogenic wild-type strain E11 during growth in 7H9 culture medium. (A) Scatterplot of the relationship between differentially expressed genes quantified in both data sets. (B-F) Scatterplots for protein and transcript gene expression classified by functional categories. Scatterplots display the rectilinear equation and coefficient of determination (R^2).

Figure S4

Functional categories classification of genes that are significantly changed in the transcriptome and proteome of the *M. marinum eccCb$_1$* transposon mutant compared to its isogenic wild-type strain E11 during growth in 7H9 culture medium. Genes exhibiting differential expression at the RNA and protein level were grouped according to the Marinolist classification (http://mycobrowser.epfl.ch/marinolist.html).

Figure S5

Most differentially expressed genes of *M. marinum eccCb$_1$* transposon mutant compared to its isogenic wild-type strain E11 during infection of primary macrophages, grouped into broad functional categories. Within each group, genes are ranked in ascending order by *P*-value. (A). Top-100 annotated *M. marinum E11* strain genes that show the most differentially expressed in *M. marinum* wild-type strain E11 during infection of primary macrophages. Bar chart of log2 fold change for individual genes (tags, left). (B). Top-100 annotated *M. marinum E11* strain genes that show the most differentially expressed in *M. marinum eccCb$_1$* transposon mutant compared to its isogenic wild-type strain E11 during infection of primary macrophages (tags, left). Bar chart of log2 fold change for individual genes.

Figure S6:

887 Regulation of genes encoding proteins predicted to be involved in metabolic adaptation,
888 energy metabolism and transcriptional regulatory in *M. marinum eccCb$_1$* transposon mutant
889 grown in 7H9 culture medium as well as wild type and *eccCb$_1$* transposon mutant during
890 infection in human THP-1 macrophages (indicated as 'int') compared to wild-type strain E11
891 during growth in 7H9 culture medium. (A) *Catabolism of Fatty Acid*. Genes were selected
892 based on their annotation and ordered based on expression. (B) Energy generation and NAD^+
893 regeneration. Genes were selected based on their annotation and ordered based on expression.
894 (C) Genes of the *kstR* regulon which are required for uptake and metabolism of cholesterol
895 (32, 61). (D) Transcriptional regulatory. Genes were selected based on their annotation and
896 ordered based on expression.

897 **Figure S7**

898 The enriched Gene ontology (GO) Terms of the Activation (Exclude ESX-1 locus genes) gene
899 set and Repression gene set. The molecular function GO terms are in red color while the
900 biological processes are in blue color.

901

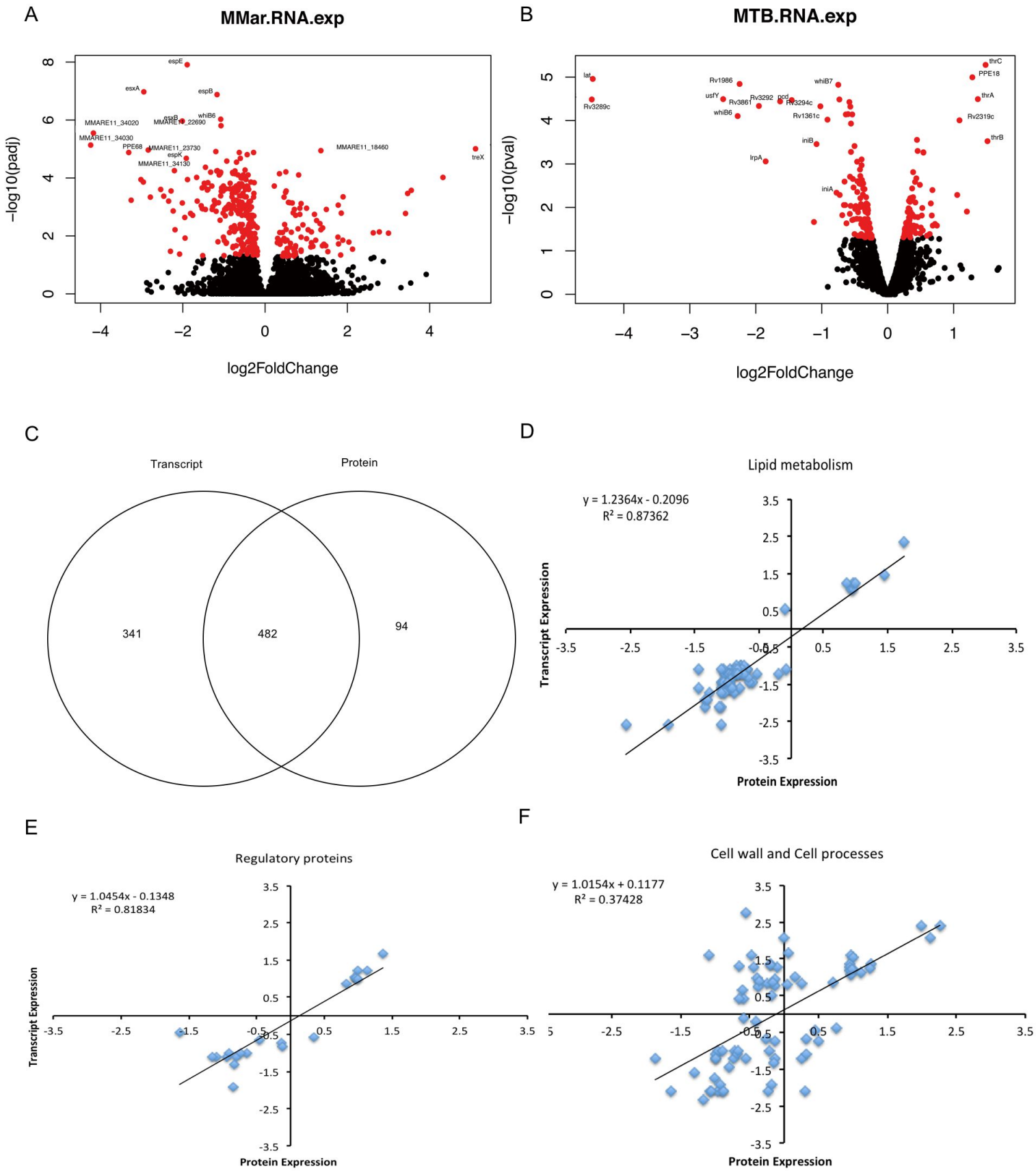

Figure.1

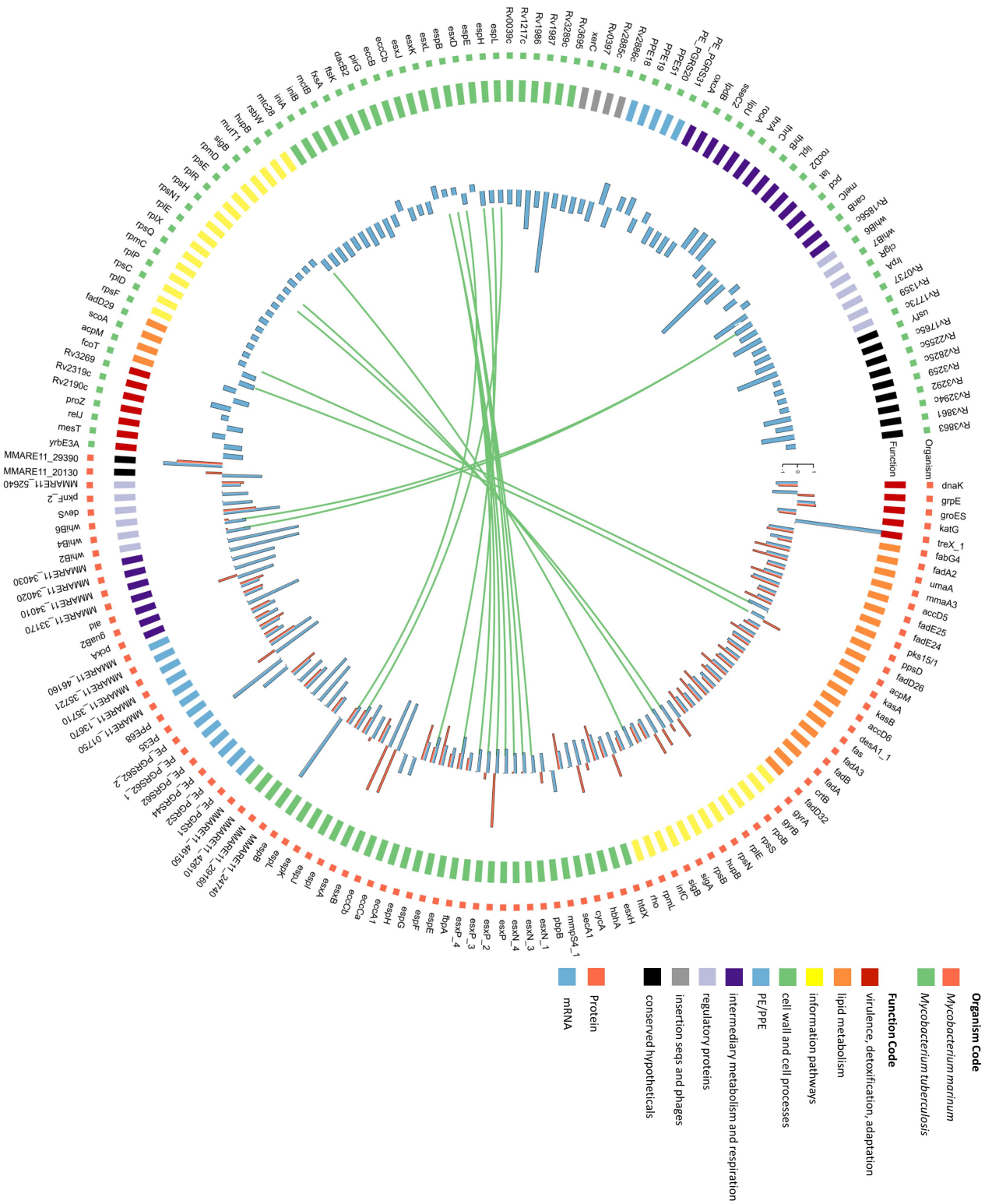

Figure.2

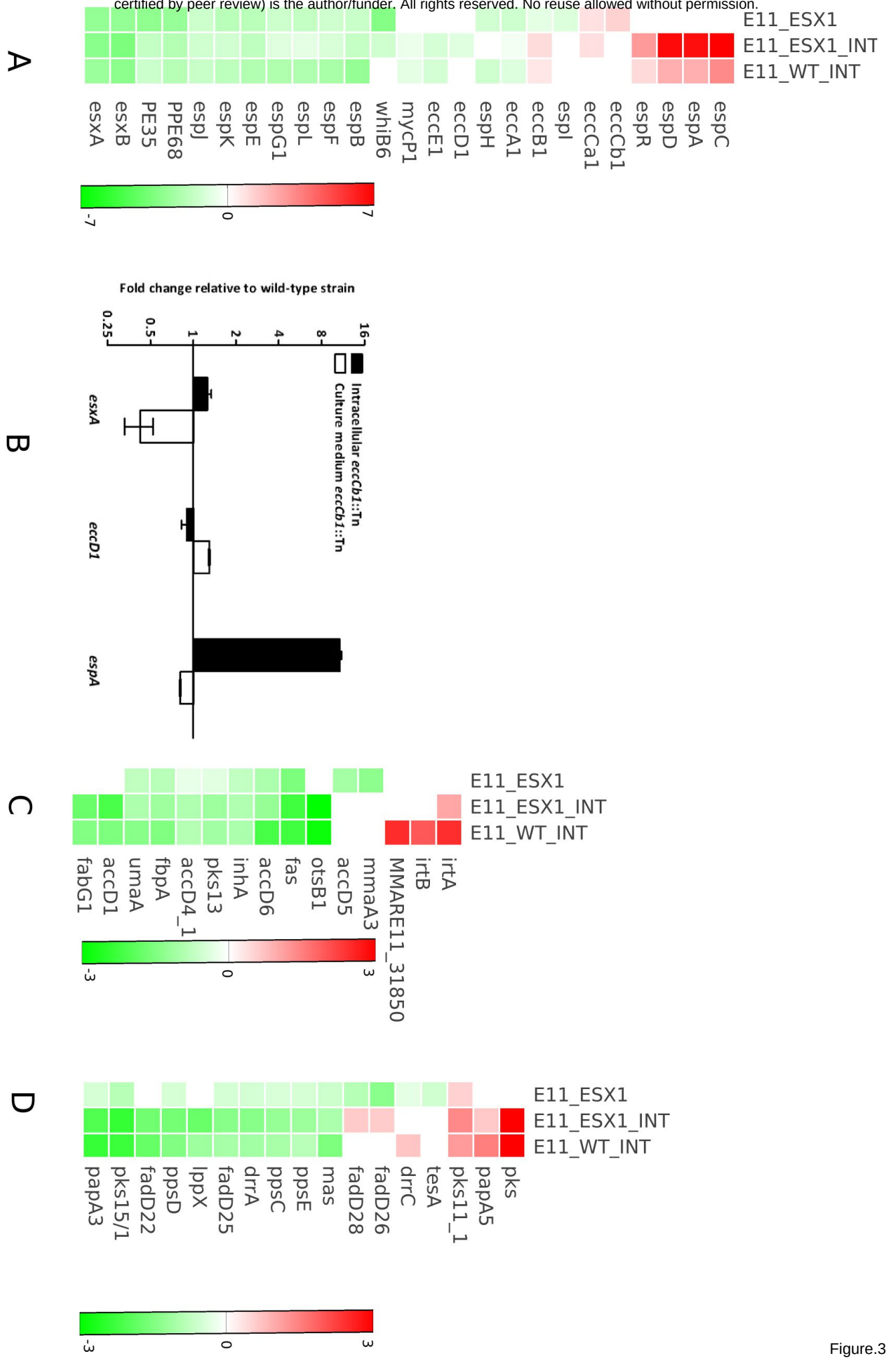

Figure.3

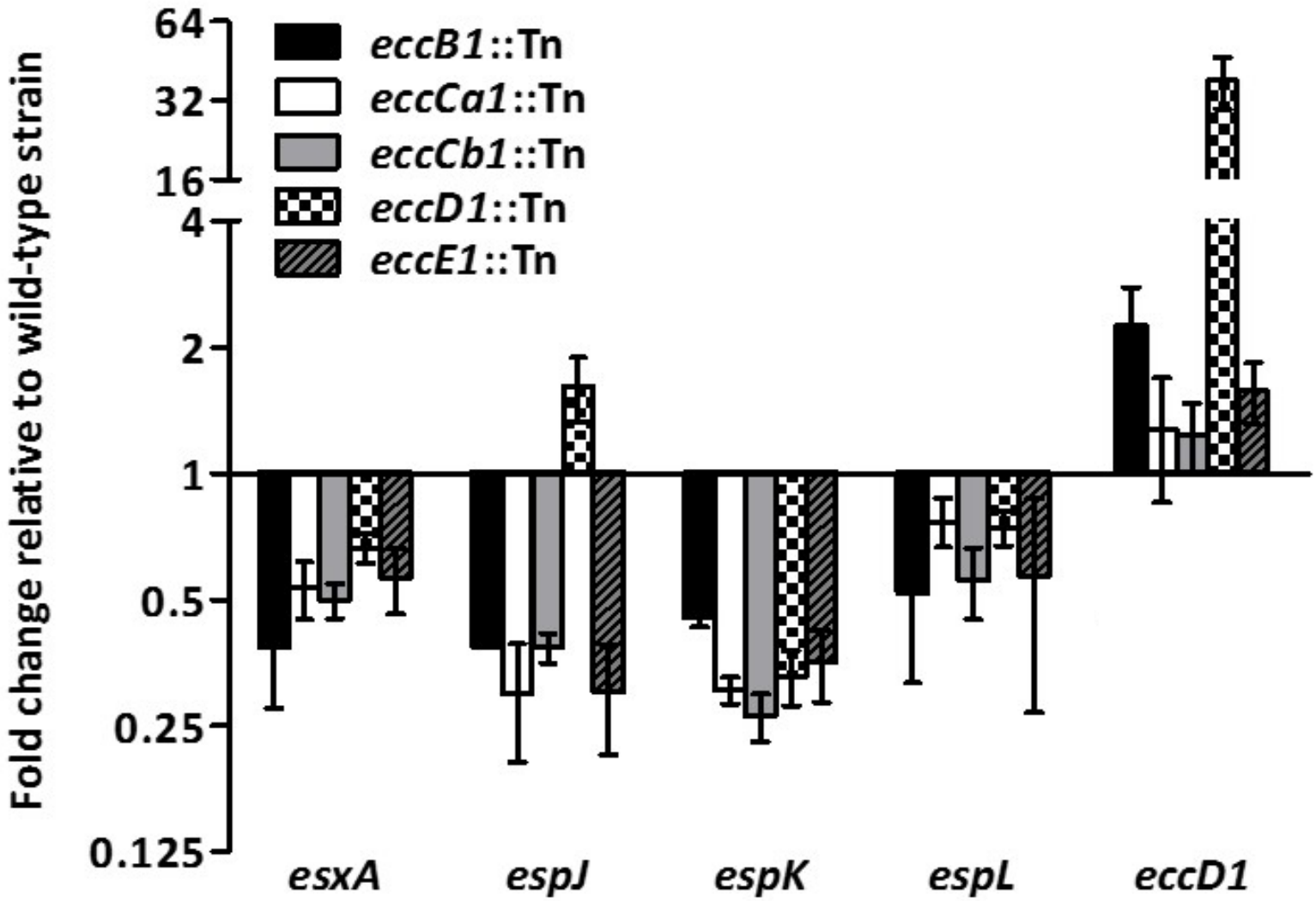

Figure.4

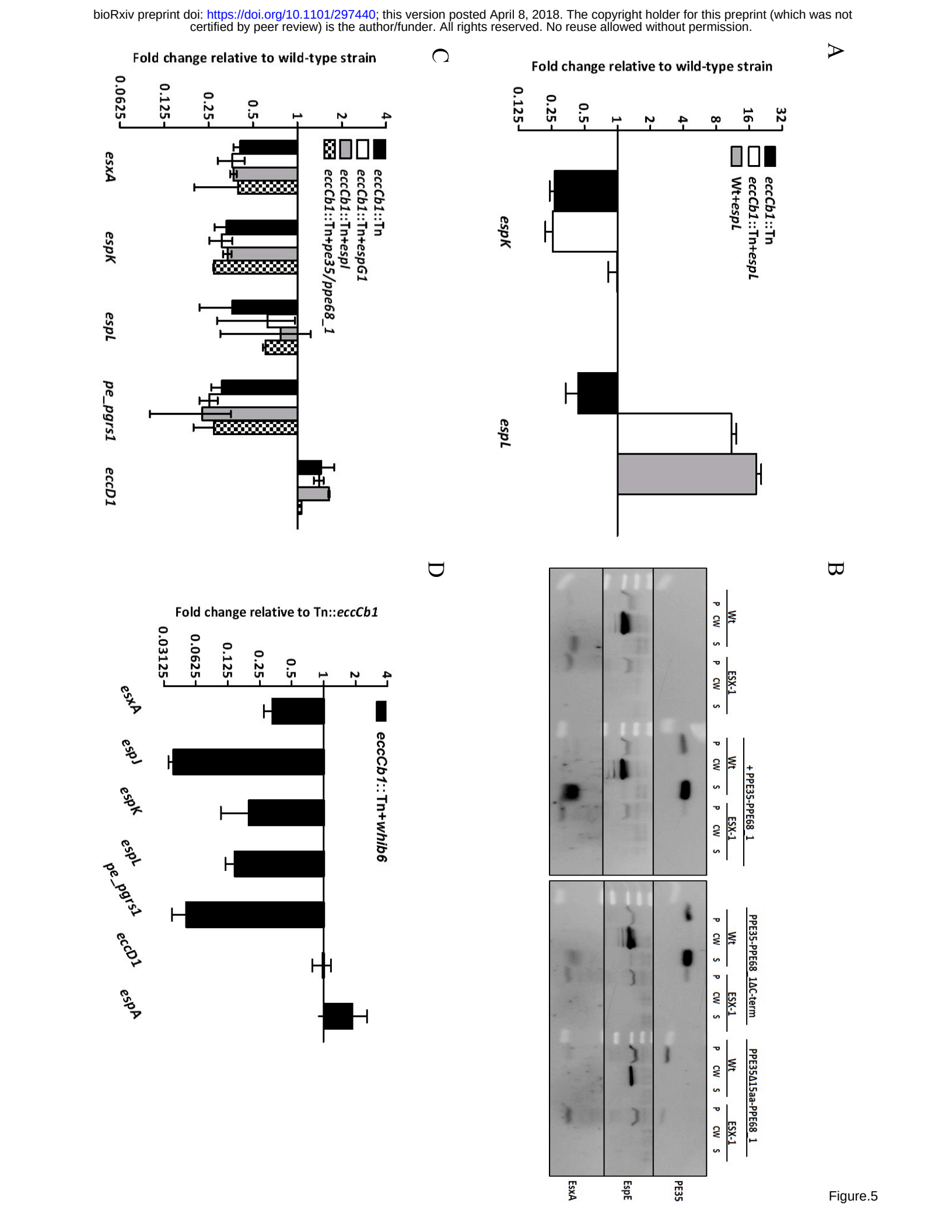

Figure.5

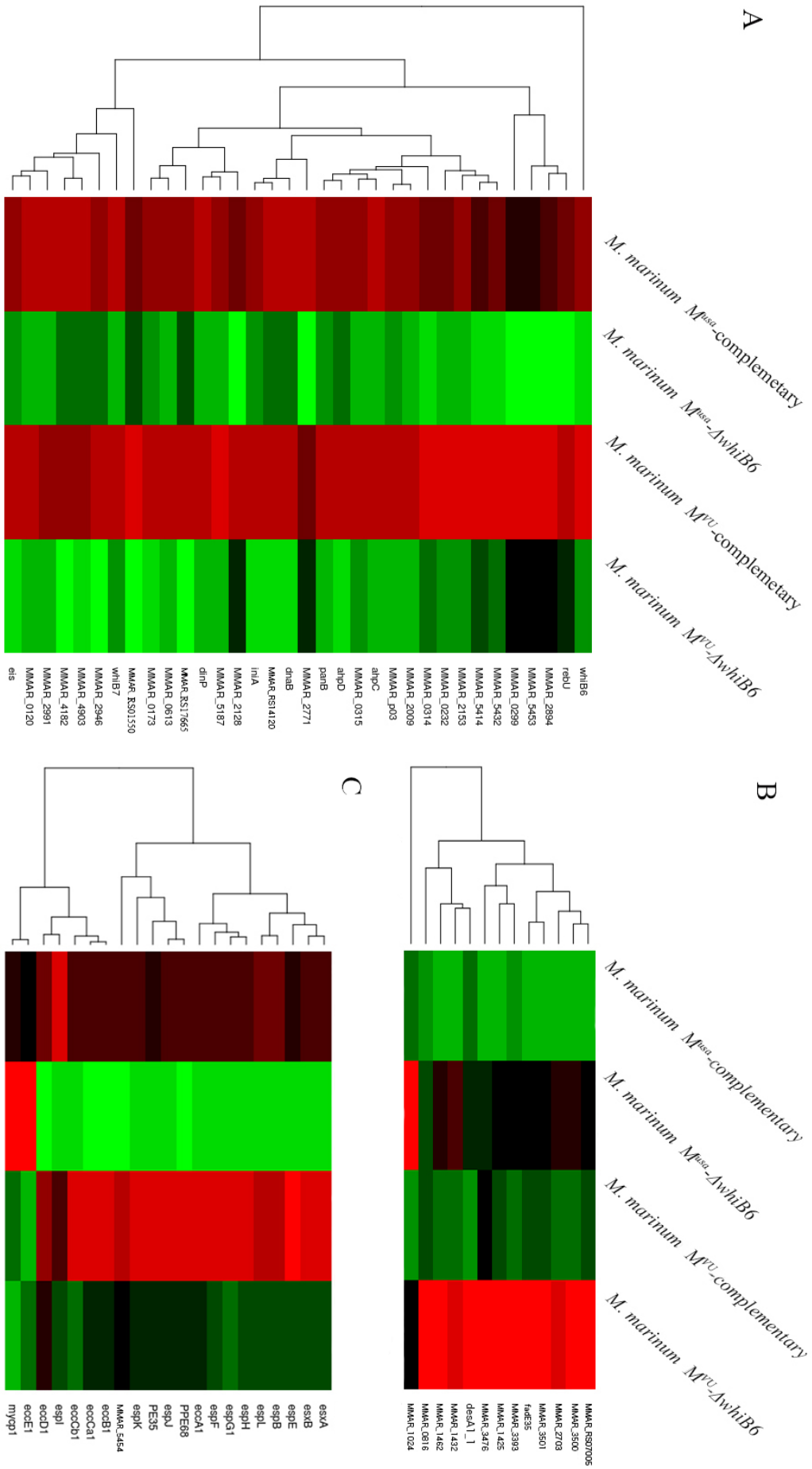

Figure.6