SCIENTIFIC REPORTS

OPEN

System-wide coordinates of higher order functions in host-pathogen environment upon *Mycobacterium tuberculosis* infection

P. V. Parvati Sai Arun[1], Sravan Kumar Miryala[2], Aarti Rana[3], Sreenivasulu Kurukuti[4], Yusuf Akhter[5] & Sailu Yellaboina[2]

Molecular signatures and their interactions behind the successful establishment of infection of *Mycobacterium tuberculosis* (*Mtb*) inside macrophage are largely unknown. In this work, we present an inter-system scale atlas of the gene expression signatures, their interactions and higher order gene functions of macrophage-*Mtb* environment at the time of infection. We have carried out large-scale meta-analysis of previously published gene expression microarray studies and have identified a ranked list of differentially expressed genes and their higher order functions in intracellular *Mtb* as well as the infected macrophage. Comparative analysis of gene expression signatures of intracellular *Mtb* with the *in vitro* dormant *Mtb* at different hypoxic and oxidative stress conditions led to the identification of the large number of *Mtb* functional groups, namely operons, regulons and pathways that were common and unique to the intracellular environment and dormancy state. Some of the functions that are specific to intracellular *Mtb* are cholesterol degradation and biosynthesis of immunomodulatory phenolic compounds. The molecular signatures we have identified to be involved in adaptation to different stress conditions in macrophage environment may be critical for designing therapeutic interventions against tuberculosis. And, our approach may be broadly applicable for investigating other host-pathogen interactions.

Mycobacterium tuberculosis (*Mtb*) is the causative factor for tuberculosis (TB). Besides HIV, *Mtb* is the leading cause of death worldwide[1–3]. After successful infection into its host, Mtb survives as an intracellular pathogen, localized inside the macrophages and dendritic cells[4,5]. The phenomenon of infection of a healthy macrophage by *Mtb* can be divided into three inter-related stages such as triggering inflammatory response from macrophage upon infection; activation of cell mediated immunity and followed by formation of granulomas, and finally the reactivation of *Mtb* due to various environmental and genetic factors[6]. *Mtb* is known to survive for many years under the adverse intracellular environment of macrophages by down regulating the metabolic pathways and switching to non-replicative state[7,8]. It was observed that to survive in the adverse conditions, *Mtb* not only alters its gene expression profiles but also it alters the gene expression profiles of the host[9].

Previously, many independent microarray studies were conducted to identify differentially expressed genes in *Mtb* as well as macrophages during the stable infection[10–19]. In addition to these studies, there are also few gene expression microarray studies conducted to identify adaptive mechanisms of *Mtb* to adverse conditions in macrophages by using *in-vitro* models such as nutrient starvation, hypoxia, and oxidative stress[20,21]. Furthermore, Meta-analysis of gene expression microarray data from independent and related datasets was also performed in several studies by which the random errors created in experiments were nullified and also the right estimate of differential expression was obtained[22,23]. Moreover, by performing the meta-analysis, it was also found that

[1]C R Rao AIMSCS, University of Hyderabad campus, Hyderabad, Telangana, India. [2]IOB-YU Centre for Systems Biology and Molecular Medicine, Yenepoya Research Centre Yenepoya University, Mangalore, Karnataka, India. [3]Centre for Computational Biology and Bioinformatics, School of Life Sciences, Central University of Himachal Pradesh, Dharamshala, India. [4]Department of Animal Biology, School of Life Sciences, University of Hyderabad, Hyderabad, India. [5]Department of Biotechnology, Babasaheb Bhimrao Ambedkar University, Vidya Vihar, Raebareli Road, Lucknow, Uttar Pradesh, 226025, India. P. V. Parvati Sai Arun and Sravan Kumar Miryala contributed equally to this work. Correspondence and requests for materials should be addressed to S.Y. (email: bio.sailu@gmail.com)

there is an increase in the reliability and generalization of results[23–25]. In this work, we performed simultaneous meta-analysis of gene expression microarray data from infected host cell (macrophage) as well as pathogen (*Mtb*) to identify the ranked list of gene expression signatures induced in host-pathogen environment.

Furthermore, functional over-representation analysis (FOA) of different gene groups namely complexes, pathways, regulons and gene ontology terms among the gene expression signatures was also performed which led us to identify the system-wide functions in Mtb-macrophage environment. Inclusion of operons for FOA of Mtb gene expression signatures identified many unknown pathways/complexes and adaptive mechanisms with in the macrophage environment.

Furthermore, functional over-representation analysis (FOA) of different gene groups namely complexes, pathways, regulons and gene ontology terms among the gene expression signatures was also performed which led us to identify the system-wide functions in *Mtb*-macrophage environment. Inclusion of operons for FOA of *Mtb* gene expression signatures identified many unknown pathways/complexes and adaptive mechanisms within the macrophage environment.

Materials and Methods

Preparation of gene expression microarray datasets.   Gene expression microarray data sets were downloaded from two major public repository databases Gene Expression Omnibus (GEO) (http://www.ncbi. nlm.nih.gov/geo/) and ArrayExpress (http://www.ebi.ac.uk/microarray-as/ae/). To remove the redundancy in the gene expression data, we have included the datasets from ArrayExpress database, if they are not found in GEO database. To facilitate the data comparison across the different platforms, gene expression microarray features were mapped to Entrez Gene IDS. Mapping information from gene names, refseq IDs, Ensembl IDs, Clone IDS, GenBank IDS to Entrez Gene IDS were obtained from various platform annotation files and also from various databases such as NCBI, BioMart, NIA array[26–28].

Differential expression analysis.   In order to identify differentially expressed genes during infection, we considered the samples during the final stage of the infection. Differential expression analysis of gene expression microarray data was carried out using an R package called 'RankProd'[29]. The package uses a modified 'rank product' method[29,30] to identify up or down-regulated genes in one condition against another. We have used RankProd because, it uses Rank Product method which is a non-parametric statistical method that detects genes which are consistently found among the most strongly up-regulated genes in with replicate or without replicate experiments. It is based on the assumption that under the null hypothesis the order of all items is random. Probability of finding a specific item among the top of items in a list is k/n, where k is the rank of the item and n is the total number of items. Multiplying these probabilities gives the rank product. Smaller the rank product value, smaller the probability that the observed placement of the item at the top of the lists is due to chance.

Meta-analysis of differential expression.   Meta-analysis of gene expression microarrays was carried out using our previously described method[31]. We transformed the initial p-values of each gene obtained from the differential expression analysis of gene expression microarray datasets into the Z-scores, using the inverse cumulative distribution function (CDF)[32]. Then we took the weighted sum of the Z-scores and divided by the square root of sum of the weights to generate the average Z-score. Here the weights are equivalent to number of replicates in the gene expression microarray experiments. Finally the average Z-score was converted to p-value by using CDF[33].

Inference of interactions of between *Mtb* and Human.   Experimentally verified Host-Pathogen interaction data from different bacterial species were retrieved from PHISTO and HPIDB databases[34,35]. PHISTO is a Pathogen Host Interaction Search Tool which contains a total of 23,633 pathogen host PPI data including 9318 interactions between bacteria and human. HPIDB is a Host Pathogen Interaction Database that currently contains 23.485 interactions between 66 host and 541 pathogen species. In addition, we curated a total of 53 experimentally verified host-pathogen protein-protein interaction information from 12 pathogens including *Mtb* from literature. The genome sequences of *Mtb* and other bacteria, which have experimentally verified interactions, were downloaded from NCBI ftp site (ftp://ftp.ncbi.nlm.nih.gov/genomes/). Then, we identified orthologs of *Mtb* in ~1500 sequenced genomes of bacterial species, which have experimentally verified host-pathogen interactions. Finally, the experimentally verified interactions were transferred to *Mtb* on the basis of orthology relationship. Totally we identified 2808 host-pathogen interactions between two sets of proteins 564 from Human and 1725 from *Mtb*.

Preparation of functional gene groups.   Annotated functions of human, i.e., pathways, motifs, gene ontology (GO) and immunological signatures were collected from Molecular Signature Database (MSigDB)[36,37]. Since the MSigDB does not maintain any functional gene groups for *Mtb*, we have curated the various functional gene groups such as complexes, pathways and operons etc., for *Mtb* from the published literature[38].

Functional over representation analysis of gene groups.   Functional over-representation analysis (FOA) was carried out using our previously described method[39]. For each gene in a gene group, we have converted p-value of differential expression into Z-score using, inverse normal Cumulative Distribution Function (CDF). The FOA score of each gene group is the sum of the Z-scores divided by square root of number of genes. In order to identify the overrepresented target genes of various transcription regulators (regulon), we have taken sum of the absolute value of target genes Z-scores, then the total was divided by square root of number of genes. Finally the q-value was converted to P-value by using CDF. To calibrate the gene group (containing n genes) FOA score against the background distribution, we randomly sampled the n number of Z-scores and calculated the random FOA score. We repeated this process over 100,000 times and mean and standard deviation of the sampling distribution was used for correction of original score.
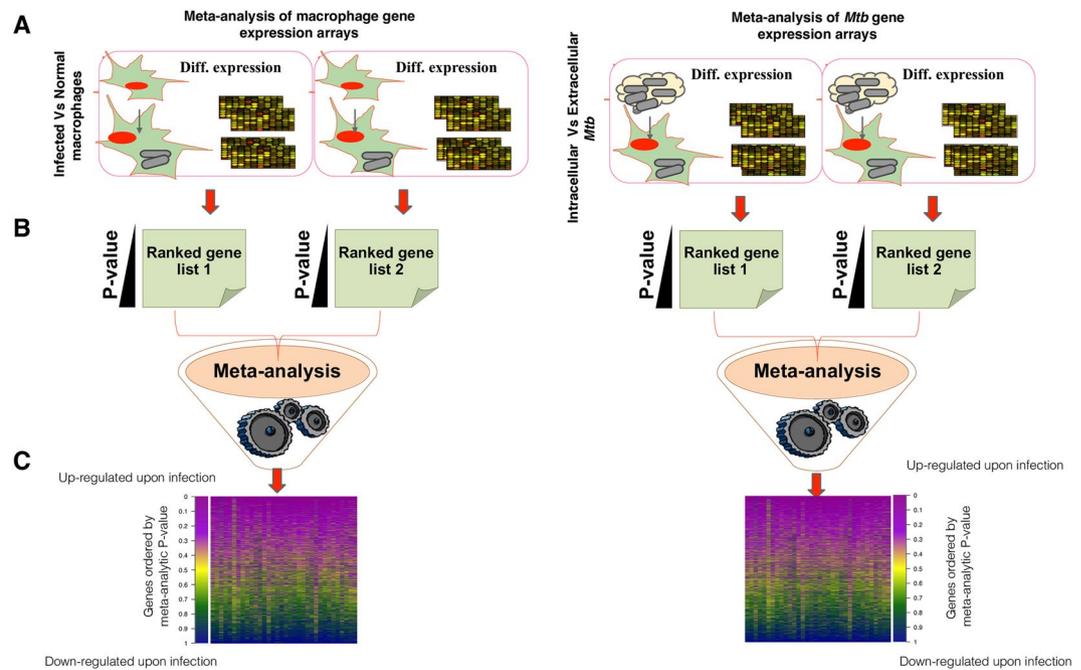
Figure 1. Meta-analytic approach to identify the gene expression signatures induced in host-pathogen environment upon infection by *Mtb*. (**A**) Initially the gene expression microarray datasets of human macrophages and *Mtb* were collected from Gene Expression Omnibus (GEO) (**B**). We have used R package to generate the ranked list of differentially expressed genes by comparing the samples of infected macrophages versus normal and *in vitro Mtb* vs. intracellular. (**C**) The ranked list of differentially expressed genes were combined using previously developed meta-analytic method to generate a consensus ranked list of the differentially expressed genes in macrophages as well as in *Mtb*.

Results and Discussion

Since both the host and the pathogen reciprocally influence the gene expression profiles of each other, it is pertinent to study the gene expression profiling of the pathogen as well as the host. This would provide insight into the strategies employed by the pathogen to survive in host environment and the host response to contest against the virulence of the pathogen[16]. Therefore, we performed the meta-analysis of gene expression changes in both *Mtb* and human macrophages during stable infection. In addition, we carried out Functional over-representation analysis (FOA) to identify higher order functions of induced or suppressed gene expression signatures both in *Mtb* as well as in the infected macrophages by considering functional groups of genes such as operons, regulons, pathways, complexes and biological process. Inclusion of diverse functional gene groups for the analysis ensured functional annotation for most of the genes. We extensively discuss the relevance of higher order functions of the genes in relation to the mechanisms adopted by the pathogen for its adaptation as well as the host defence.

Meta-analysis of differentially expressed gene signatures during infection. We have developed a systematic frame work to identify the genes induced in host-pathogen environment (Fig. 1). Initially, we have identified two sets of differentially expressed genes, one from the *Mtb* and the other from the macrophages during the course of infection. Differentially expressed genes of *Mtb* were identified by comparing the gene expression data of *Mtb* which resides inside macrophage after infection (call it intracellular *Mtb*) with the *Mtb* before infection (*in vitro Mtb* cultures) using RankProd[29]. Similarly, we have also identified differentially expressed genes of macrophages by comparing gene expression data of the infected macrophages with the uninfected macrophages (Supplementary data S1). The differential gene expression scores were given ranks and then combined to generate a single consensus ranked list, using a meta-analytical method described previously[31]. The genes ranked atop the consensus ranked list are more severely and consistently up-regulated during infection across experiments, and those that are ranked at the bottom were down-regulated during infection.

Host-pathogen interactions and essentiality of top ranked genes. We identified orthologs of *Mtb* in ~1500 sequenced genomes of bacterial species, which have experimentally verified host-pathogen interactions in HPIDB and PHISTO databases as well as curated from literature[34,35]. The known interactions were then transferred to *Mtb* on the basis of orthology relationships. Totally, we identified 2808 host-pathogen interactions between two sets of proteins 564 from human and 1725 from *Mtb* (Supplementary data S2). Meta-analytic Z-score distribution analysis of meta-analytic ranked list from macrophage and *Mtb* shows that the most of the genes from both sets have above average Z-scores indicating that most of the top ranked genes encode for proteins involved in host-pathogen interactions (Fig. 2A,B). In addition, the studies conducted using genome wide siRNA screen also have large number of genes involved in the regulation of pathogen load and found to have above
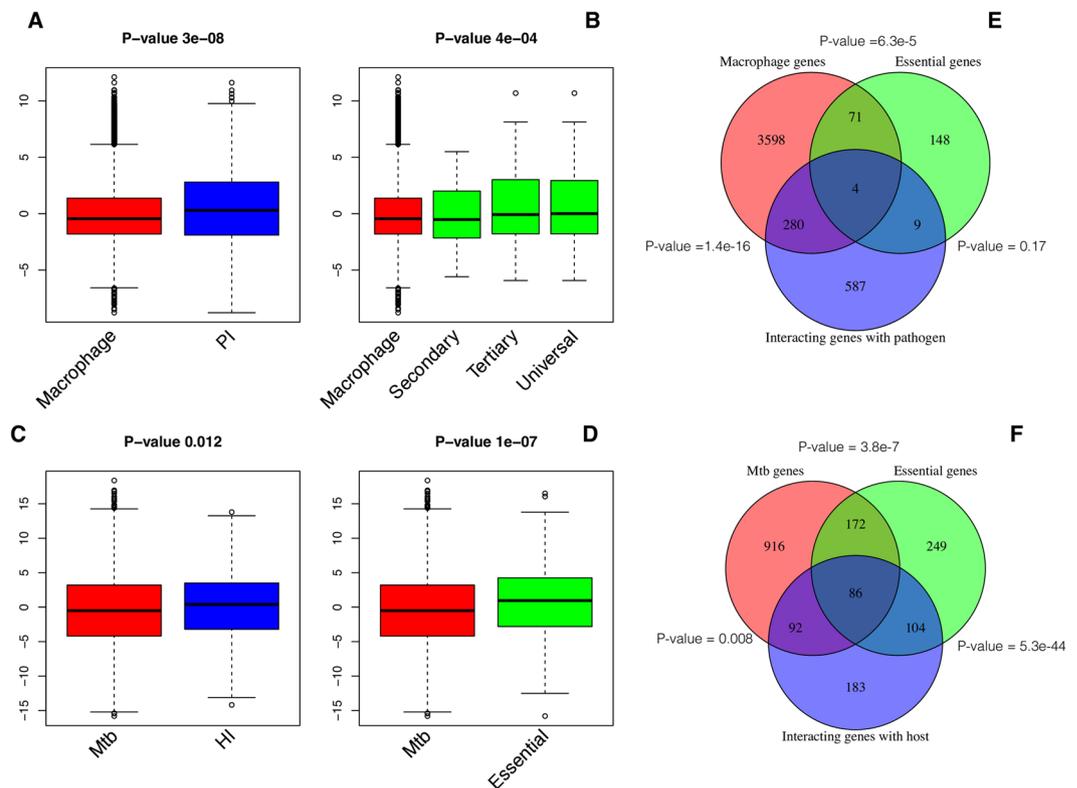
Figure 2. Box plots showing the comparison of differential expression Z-Scores for various gene lists with overall z-scores of the genes induced in the host-pathogen environment. The positive Z-scores indicates up-regulation, whereas negative scores indicate down-regulation; (**A**) Z-scores of total macrophage genes that are differentially expressed upon infection and macrophage genes predicted to be pathogen interacting (PI). (**B**) Z-scores of total macrophage genes that are differentially expressed upon infection and macrophage genes that were reported to be essential to regulate pathogen load using different levels of siRNA screen i.e., secondary, tertiary and universal. (**C**) Z-scores of total *Mtb* genes that are differentially expressed in infected macrophages and *Mtb* genes predicted to be host interacting (HI). (**D**) Z-scores of total *Mtb* genes that are differentially expressed in infected macrophages and *Mtb* genes reported to be essential for survival in different physiological environments. P-value calculated using Wilcoxon rank sum test shows that the distributions of each pair of comparisons are significantly different. (**E**) Venn-diagram with P-value of overlapping statistics between the genes that are expressed with in the macrophages, pathogen interacting and essential (tertiary screen) and (**F**) Venn-diagram with P-value of overlapping statistics between the *Mtb* genes that are expressed with in the macrophages, host interacting and essential.

average Z-scores in meta-analytic ranked list of macrophage genes. (Fig. 2C)[18]. Moreover as the quality of RNAi hits (3 sets of genes) also have above average Z-score in the meta-analytic ranked list, indicating that top ranked genes in the meta-analytic list are highly important for defence against *Mtb*. Similarly, some of the genes reported to be important for survival of *Mtb* in different physiological environments were also have above average Z-scores (Fig. 2D)[32,33,40]. Overlapping statistics using hyper geometric distribution also shows that the top ranked genes from both macrophages as well as *Mtb* were significantly overlap with their respective set of genes that encode for proteins involved in predicted host-pathogen interactions (Fig. 2E,F). There is a significant overlap between a large number of *Mtb* genes that are reported to be essential and the genes, whose products are predicted to interact with the host. The common candidates between the three set *Mtb* genes i.e., essential, host interacting and expressed with in macrophages could be potential targets for therapeutics.

Higher order functions of *Mtb* genes induced in host-pathogen environment. Previously there were many gene expression microarray experiments conducted to identify the adaptive mechanisms of *Mtb* towards the *in-vitro* models of nutrient starvation, hypoxia, and oxidative stress in macrophage environment. There were two types of *in-vitro* models studied under hypoxic conditions, one is constant oxygen (0.2%) supply model and another one is Wayne growth (WG) model, in which the oxygen supply is gradually depleted[41]. In order to identify the nature of environment created by macrophage towards *Mtb* infection, we compared the meta-analytic ranked list of gene expression signatures induced inside macrophages with a large number of previously published gene expression signatures generated using *in-vitro* models of diverse environmental conditions at various time points (GSE16146, GSE8786, and GSE9331)[20,21,42]. The gene expression signatures induced under conditions of constant O_2 (0.2%) for 4hrs, Wayne growth for 8 days, 5 mM H_2O_2 for 40 mins and 0.5 mM DETA
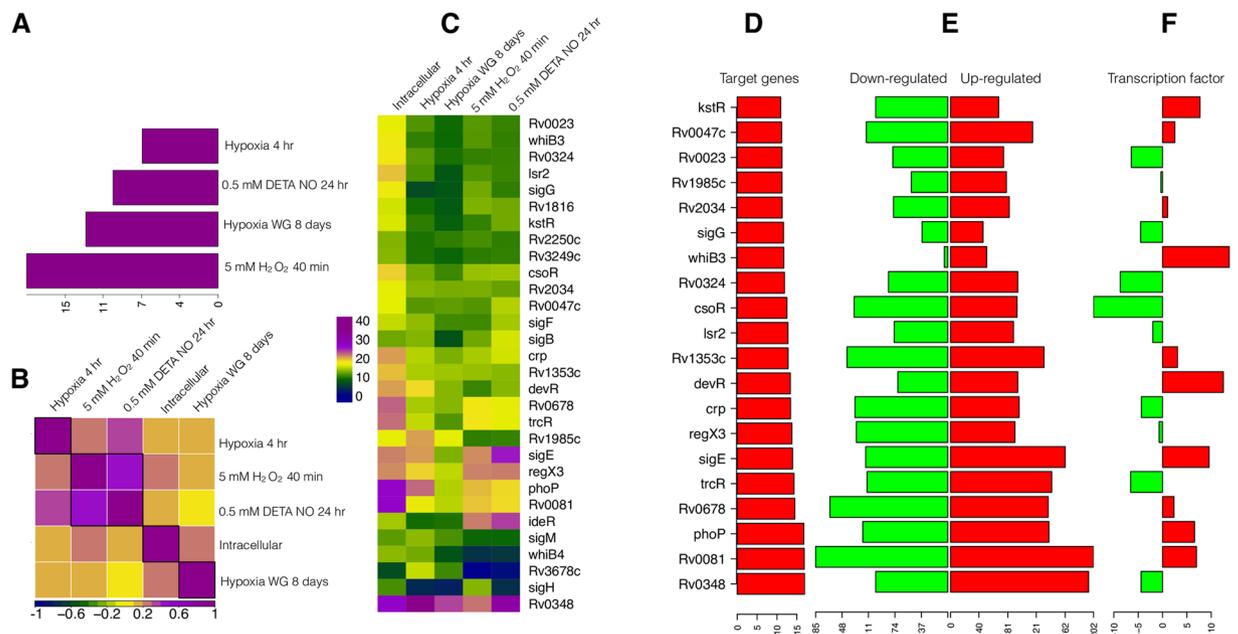
Figure 3. Comparison of meta-analytic ranked list of *Mtb* genes induced in intracellular (macrophage) environment with the genes induced *in-vitro* conditions. (**A**) Bar plot showing the distribution of quantile normalized P-value of Spearman's rank correlation (threshold value for significance is 2.3) between meta-analytic ranked list with genes induced under different hypoxia and oxidative stress conditions. (**B**) Correlation plot showing the Spearman's rank correlation (all the correlations have significant P-value) between meta-analytic ranked list with genes induced under different hypoxia and oxidative stress conditions. (**C**) Heat map showing the over/under representation analysis scores for highly differentially expressed target genes of *Mtb* transcription regulators (score >2 is significant) under different conditions. (**D**) Bar plot showing the over/under representation analysis scores for highly differentially expressed target genes of various *Mtb* transcription regulators in intracellular conditions. (**E**) Divergent bar plot showing the number of down-regulated (green bars) and up-regulated genes (red bars) for each of the transcription regulator. (**F**) Bar plot showing the expression level of each of the transcription regulator.

(Diethylenetriamine) NO for 24hrs best correlates with the meta-analytic ranked list (intracellular), indicating the likely environmental conditions *Mtb* exposed in macrophage environment (Fig. 3A). Gene expression signatures related to hypoxia model of WG shows stronger correlation than constant O2 (0.2%) supply, model. 5 mM H2O2 shows the strongest correlation. In addition, gene expression signatures induced in all the five different conditions show significant correlation with each other (P-value > 1.5e-08). Particularly, the conditions of hypoxia, H_2O_2 and NO show very strong correlation (Fig. 3B).

In order to identify the higher order functions of *Mtb* genes expressed across different conditions and in intracellular environment, we carried out FOA and functional interpretation of regulons, operons and pathways as following[39].

Regulons. We curated *Mtb* regulons by collecting known target genes of all transcription factors of *Mtb* from previously published literature[43,44]. We have performed FOA of target genes of various transcription regulators (regulons) in the ranked list of differentially expressed genes across the different conditions to get the list of the transcription regulators, which are critical for adaptation of *Mtb* to macrophage environment. Our analysis shows that the target genes of the persistence regulator, Rv0348 (MosR) were consistently differentially regulated across the different conditions (Fig. 3C) (Supplementary data S3). In addition, the target genes of Rv0081, PhoP, Rv0678, TrcR, SigE, RegX3, Crp, DevR, Rv1353c, Lsr2, CsoR, Rv0324, WhiB3, SigG, Rv2034, Rv1985c, Rv0023, Rv0047c, kstR were also differentially regulated in intracellular environment (Fig. 3D). Although target genes of the RegX3 and WhiB3 were significantly altered in intracellular conditions (Fig. 3E), the respective transcription regulators were not significantly differentially expressed (Fig. 3F). In contrast to the MosR expression, most of its target genes were up-regulated, indicating that it represses the majority of the genes involved in adaptation to the macrophage environment. MosR and its associated genes in an operon *Rv0347-Rv0348-Rv0349* were induced during the late stage of chronic tuberculosis in mice[45]. PhoP is a global regulator of *Mtb* and it plays an important role in pathogenesis and synthesis of cell wall lipids[46].

Operons. In most of the bacterial species such as *Mtb*, pathways, complexes and gene ontology terms were not well annotated, which hinders the functional annotation of gene expression signatures. Here, we circumvent the functional annotation problem by inclusion of predicted operons, because genes within the most of operons were known to encode proteins involved in a same pathway or potentially forming a functional protein complex.
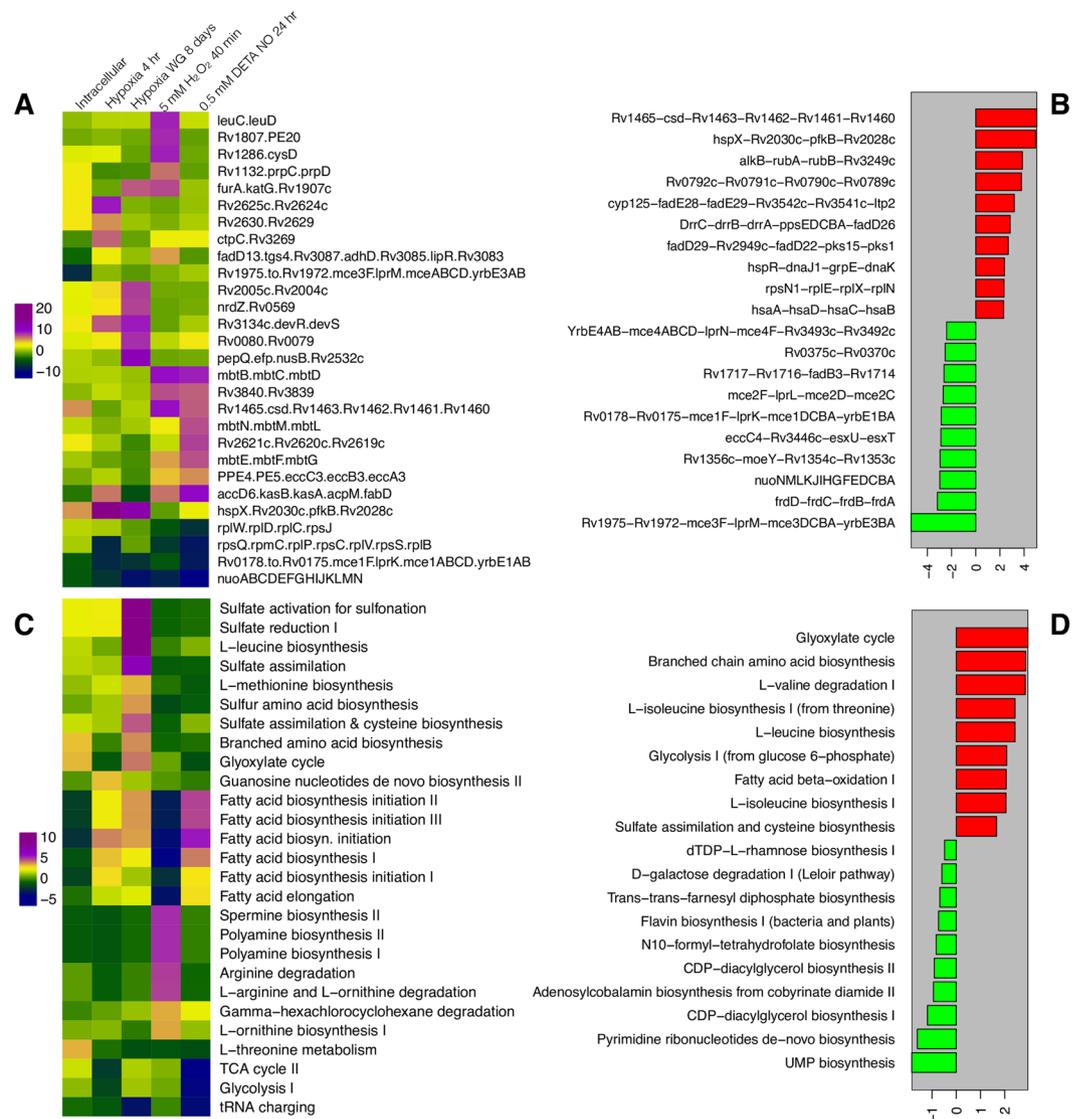
Figure 4. Operons and pathways of *Mtb* with highly differentially expressed genes under different conditions. (**A**) Heat map showing the over/under representation scores of various operons under different stresses and intracellular (macrophage) environment. (**B**) Bar plot showing the operons with highly differentially expressed genes in intracellular environment. (**C**) Heat map showing over/under representation scores of various pathways in different conditions. (**D**) Bar plot showing over/under representation scores of various pathways in intracellular environments.

Over-representation analysis shows that the entire operon encoding for the components of NADH dehydrogenase complex-I were significantly down-regulated across the different conditions of intracellular, hypoxic and oxidative stress (Fig. 4A) (Supplementary data S4), indicating the complete shutdown of oxidative phosphorylation pathway. On the other hand, the operons such as *Rv1465-csd-Rv1463-Rv1462-pps1-Rv1460*, *hspX-Rv2030c-pfkB-Rv2028*, *ceccCa1-eccB1-eccA1-espH-espG1-espF*, *hsaABCD*, *rplNXEN1*, *dnaK-grpE-dnaJ1-hspR*, *pks1-pks15-fad22-fad29*, *fadD26-ppsABCDE-drrABC*, *ltp2-fadE29-fadE28-cyp125* were highly expressed in intracellular conditions (Fig. 4B). The pps1 locus constitutes an operon with seven genes (*Rv1465-Rv1460*), which code for Suf-like proteins, which are part of SUF machinery, an exclusive Fe-S cluster assembly existing in *Mtb*[47]. The Pps1 protein is orthologous to SufB which is essential and conserved component of iron sulphur cluster assembly (Fe-S cluster) complex of prokaryotes. It was reported that as Fe-S clusters serve as cofactors of proteins which are involved in multiple cellular mechanisms[48].

The operon *hspX-Rv2030c-pfkB-Rv2028* was reported to be essential for the *Mtb* during persistent infection in the host and also for its survival during low iron, low pH and hypoxic condition[49–51]. The protein encoded by *hspX* gene of the operon is known to be highly up-regulated during the latent infection and reported to play an important role in anaerobic respiration of *Mtb* during stress conditions. The other genes of the operon are found to be involved in carbohydrate metabolism (*pfkB*) and nucleotide metabolism (*Rv2030c*)[49–51].

The two operons *hsaABCD* and *cyp125-fadE28-fadE29-Rv3542c-Rv3541c-ltp2* were highly conserved across the members of *Mtb* complex[52,53]. They participate in cholesterol degradation pathway and play an important role in intracellular survival of *Mtb* by providing cholesterol as an alternative carbon source[54–59]. Another two operons *pks1-pks15-fad22-fad29* and *fadD26-ppsABCDE-drrABC* were involved in synthesis of phthiodiolone dimycocerosates (DIMs) and immuno-modulatory phenolic glycolipids, which have a role in immuno-modulation[60–62]. Some of the genes (*drrA*, *drrB* and *drrC*) in aforementioned operon were known to encode for a drug efflux pump, which confers resistance to many antibiotics[63]. The genes with in the operon (*alkB-rubA-rubB-Rv3249c*) encode for alpha-ketoglutarate-dependent dioxygenase (*alkB*), which catalyses oxygenation of hydrocarbons[64], could be having role in cholesterol degradation. The genes involved in cholesterol utilization and other genes (*Rv0792c*, *Rv0791c*, *Rv0790c*, *Rv0789c*) with unknown function were highly expressed in intracellular environment in comparison to the other conditions.

Comparative analysis of operons with the genes induced across different conditions has identified several functions that are common and unique to each of the environment. Pathways involved in biosynthesis of iron-sulfur cluster, mycobactins (*mbtB-mbtC-mbtD*) and meromycolic acids (*accD6-kasB-kasA-acpB-fabD*) were highly expressed in presence of NO and H_2O_2 indicating their role in oxidative stress defence. On the other hand the genes with in the operons (*devR-devS-Rv3134c* and *hspX-Rv2030c-pfkB-Rv2028c*) were highly expressed during hypoxic conditions. Furthermore the genes with in the operons (*furA-katG-Rv1907c* and *ahpC-ahpD*) were expressed across intracellular, hypoxic, H_2O_2 induced oxidative stress conditions. In contrast to the previous reports, co-expression of *katG* and *furA* indicates that *furA* may not be a repressor of *katG* under these conditions[65].

Pathways. Over-representation analysis of various *Mtb* pathways among the genes induced during infection shows that the glyoxylate pathway which is known to play an important role in persistence of *Mtb* in macrophages[66] is only active in Wane growth model of hypoxia and intracellular conditions (Fig. 4C) (Supplementary data S4). In addition, pathways involved in the biosynthesis of chorismate and branched chain amino acids such as L-leucine, L-isoleucine (from threonine) and assimilation of sulfate and cysteine biosynthesis were highly expressed during hypoxia and intracellular conditions, but not during oxidative stress. Furthermore, various catabolic pathways such as fatty acid & beta-oxidation I, glycolysis I (from glucose 6-phosphate), and glyoxylate cycles are highly expressed during intracellular conditions (Fig. 4D). The chorismate pathway, also called as shikimate pathway is found in *Mtb* but absent in humans and other mammals[67,68]. This pathway produces chorismate, which is a key intermediate compound for the biosynthesis of amino acids, ubiquinones, siderophores and other important metabolites essential for *Mtb* growth and survival in the host[69–73]. Overall, the pathways involved in the biosynthesis of L-Leucine, chorismate, cysteine, and assimilation of sulfate are linked to the long-time survival of *Mtb* during the pathogenesis.

### Higher order functions of differentially expressed macrophages genes upon *Mtb* infection.

Pathways. In order to identify the macrophage pathways that were induced during infection, we have carried out over-representation analysis of pathway genes from various databases such as BioCarta, KEGG, and Reactome[74–76] (Fig. 5) (Supplementary data S5). The BioCarta pathways namely, IL-10 anti-inflammatory signalling, induction of apoptosis through DR3 and DR4/5 Death Receptors, IL22 soluble receptor signalling, NF-kB, TNFR2 signalling and Acetylation/deacetlyation of RelA were found to be highly expressed (Fig. 5A).

IL-10 family of cytokines has a total of 6 IL's. In contrast to the rest of the IL-10 family of cytokines, IL-10 acts as an anti-inflammatory cytokine and is also capable of inhibiting the synthesis of pro-inflammatory cytokines such as IL-2, TNF-α, IL-3, IFN-α which indirectly leads to the survival of *Mtb* inside the infected macrophage. Supporting this fact, there are reported evidence from the infection studies of *Mycobacterium avium sub species paratuberclosis* (MAP). It was documented that MAP adopts a different strategy to escape from the host defence. MAP activates the mitogen activated protein kinase pathway, which finally results in up-regulation of IL-10 and repression of other inflammatory cytokines. Further the up-regulated IL-10 also inhibits the apoptosis in macrophages[6,77,78]. According to our findings and the supporting evidence, it may be assumed that the up-regulation of IL-10, an anti-inflammatory pathway of macrophage may aid in the growth and virulence of *Mtb* and may also be considered as one of the immune-modulatory strategies adopted my *Mtb* to escape from host defence mechanisms for its survival.

The pathway 'induction of apoptosis through the engagement of death receptors' has a role in apoptosis of infected macrophages via engagement of variety of death receptors such as Fas, TNFR, DR3, DR4 and DR5[79]. Apoptosis of this kind, where there is engagement of death receptors is the sign of innate immune response of macrophages against *Mtb*[79]. IL22 soluble receptor signalling pathway was reported to inhibit the growth of *Mtb* by enhancing the expression of intracellular signalling molecule, calgranulin A. In addition, it was reported to be involved in tissue modulation, induction of proliferative and anti-apoptotic pathways and production of anti-microbial substances[80].

NF-kB pathway, inflammatory cytokine and TNF-α are responsible for inflammatory response towards *Mtb* infection[81]. TNF-α shows its immune response in variety of ways such as macrophage activation, induction of cytokines, chemokines and induction of apoptosis and also restriction of bacterial growth in infected macrophage[81–85]. TNFR2 signalling pathway is mainly responsible for apoptosis, acetylation/deacetlyation of RelA and is responsible for innate immune response against *Mtb*. Overall the over-representation analysis in Biocarta pathway shows that all the pathways related to recognition of molecular patterns of *Mtb* are responsible for triggering inflammation and apoptosis, are up-regulated.

Over-representation of analysis of KEGG pathways show that the genes involved in cytokine-cytokine receptor interaction, toll-like receptor signalling, NOD-like receptor signalling, RIG-I like signalling, chemokine
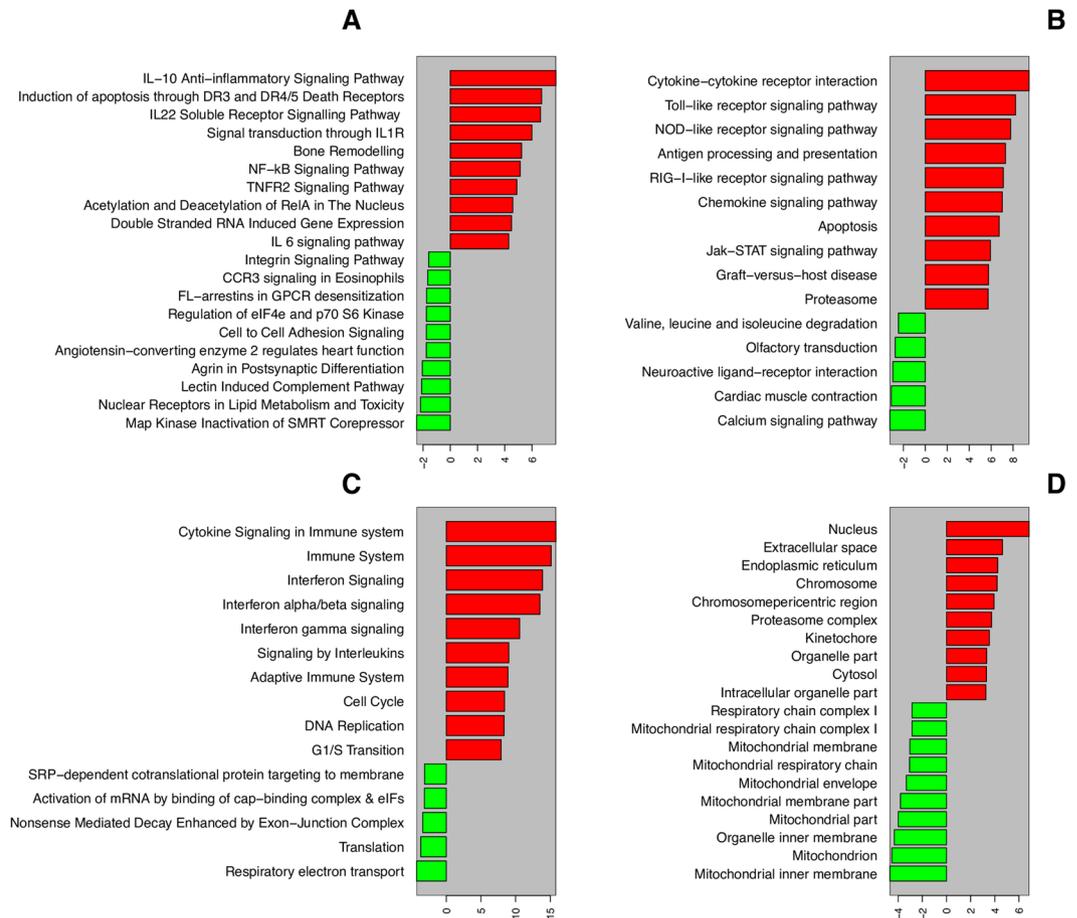
Figure 5. Macrophage pathways and cellular components with highly differentially expressed genes upon infection. (**A**,**B** and **C**) shows over/under represented pathways from Biocarta, KEGG and Reactome, respectively. (**D**) shows the over/under represented cellular components.

signalling and apoptosis signalling were highly up-regulated in the macrophages during *Mtb* infection (Fig. 5B). After infection of *Mtb* in pulmonary alveoli, macrophages express pattern recognition receptors (PRR) which in turn recognize the pathogen associated molecular patterns (PAMPS). The recognition of PAMPS by PRRs leads to the activation of transcription factors, which are involved in providing innate immunity[86]. In the case of myco-bacterial infections, mycobacterial cell wall-associated components such as lipoarabinomannan are recognized by TLR1 and TLR2. Similarly there are several cell wall associated components which may be recognized during the time of mycobacterial infection as reported earlier[87–89]. In NOD-like receptor signalling pathway, the NOD like receptors belonging to NLR family and they may play key roles in innate immunity and inflammation[90]. They recognize the peptidoglycan fragments of the pathogens escaped from endosomal compartments leading to the activation of cytokine production and ultimately resulting in programmed cell death[91,92].

Apart from the NOD like receptors, there are also few other proteins called as cytosolic RNA helicases (RIG-I like signalling pathway) that were also highly up-regulated. These cytosolic RNA helicases (RIG-I-like receptors) are reported to recognize the antigens from viral pathogens and induce the innate immune response against them by triggering the signalling pathways which produce the type I interferon and inflammatory cytokines[93,94]. These have been reported to be involved in the intracellular detection of all pathogens including *Mtb*[95]. Chemokine signalling pathway plays an important role in migration of leukocytes to the site of infection and orchestrates the host defence mechanism[96].

The genes involved in apoptotic pathways were highly up-regulated (Fig. 5A,B). The cells follow two different kinds of pathways for undergoing apoptosis: the death receptor or extrinsic pathway and the mitochondrial or intrinsic pathway, where the final dead cells are cleared through phagocytosis by the macrophages[97–99]. In addition, the pathways in involved in branched chain amino-acid degradation were down-regulated, unlike in intra-cellular *Mtb*, in which the biosynthetic pathways of branched chain amino-acids were up-regulated as mentioned previously.

Analysis of Reactome pathways show that there are two kinds of the pathways that were up-regulated dur-ing infection: pathways pertaining to immune system activation and pathways responsible for cell survival and growth. The former category consists of cytokine signalling related genes, interferon signalling, interleukins signalling and adaptive immunity. The second category consisting of the pathways related to cell cycle, DNA replication and G1/S transition pathways were also observed to be highly up-regulated (Fig. 5C). In the first

category, most of the pathways which were highly regulated were found to be cytokines with diverse functions. As previously described, cytokines have immune response triggering properties against the infections in the host. Interferon's are also cytokines (pleiotropic cytokines) which have immune-regulatory properties 103. In our analysis, we also have found the pathways pertaining to interferon $\alpha/\beta/\gamma$ signalling as highly up-regulated which are also actively involved in triggering immune response. These results confirm that the pathways in the first category are involved in providing innate and adaptive immunity against infection. The pathways which are involved in cell cycle, DNA replication, G1/S transition are well documented to be linked with cell growth and survival.

Cellular components.   Over-representation analysis of genes encoding various cellular components shows that the genes that encode for proteins localized to Endoplasmic Reticulum (ER) were observed to be up-regulated during infection. In macrophages, the mycobacterial antigen ESAT-6 evokes the ER stress, which in turn mediates the process of apoptosis[100]. This induction of apoptosis further leads to the killing of *Mtb*. The other highly overrepresented proteins in the cellular localization data set are mainly related to the intracellular proteins associated with kinetochore (Fig. 5D) (Supplementary data S5). In contrast to the ER, most of the mitochondrial components were down-regulated, indicating the shutdown of most of the mitochondrial functions in infected macrophages. *Mtb* virulence is known to correlate with mitochondrial cytochrome c release in infected macrophages[101]. Mitochondria could be an important target for mycobacteria, because of its involvement in the regulation of immune response and cell survival/programmed cell death[102–104]. Furthermore, the *Mtb* could also be exploiting some of the mitochondrial functions considering that mitochondria derived from endosymbiosis of an ancestral bacteria[105]. It has been shown that mycobacterial proteins carry mitochondrial localization signals, which may help its virulence proteins to be transported to the mitochondria and further hijack it's functions[106].

Conclusions

In this report by performing the meta-analysis of publicly available gene expression datasets and analysis of higher order functions and interactions, we have derived comprehensive view about cross talk between *Mtb* and macrophages during the stable infection. We observed a strong correlation of gene expression signatures between the intracellular bacteria with *in-vitro* models of dormancy conditions, which reflects the likely environmental conditions in existing in macrophages at the time of infection. Moreover from this analysis, it is clear that *Mtb* effectively fights against the oxidative, starvation and hypoxic conditions presented by the host as defence against infection. Analysis of different functional signatures associated with different stress conditions provides the valuable information related to the strategies adapted by *Mtb* inside the macrophage for its survival and persistence. For instance such as, the mechanisms adapted by *Mtb* to encounter the oxidative stress induced by intracellular H_2O_2, through catalase-peroxidase and alkyl hydro peroxide reductase, similarly, recruitment of iron-sulfur clusters, mycobactins and meromycolic acids to encounter the oxidative stress induced by NO shows the strategies adapted by *Mtb* to survive in infected macrophage. Further, from our analysis it is observed that, *Mtb* effectively tackles the harsh environmental conditions of macrophage by intracellular environment by up-regulation of cholesterol utilization pathways and bio-synthesis of immuno-modulatory phenolic glycolipids. Further study of these pathways and other genes with unknown function would establish their role and mechanisms of action to counter each of the stress conditions posed by macrophage environment. In addition possible pathways of *Mtb* that target and exploit mitochondria should be identified. This study clearly inferred that, when therapeutic strategies to combat dormant *Mtb* should be designed, the multiple pathways involved in adaptation to different stress conditions and pathways that target mitochondrial functions should be taken into consideration.

References

1. Daniel, T. M. The history of tuberculosis. *Respiratory medicine* **100**, 1862–1870, https://doi.org/10.1016/j.rmed.2006.08.006 (2006).
2. Gideon, H. P. & Flynn, J. L. Latent tuberculosis: what the host "sees. *Immunologic research* **50**, 202–212, https://doi.org/10.1007/s12026-011-8229-7 (2011).
3. Organization, W. H. *Global tuberculosis report 2015*. (World Health Organization, 2015).
4. Bloom, B. R. *Tuberculosis: pathogenesis, protection, and control*. (ASM press, 1994).
5. Snider, D., Raviglione, M., Kochi, A. & Bloom, B. Tuberculosis: pathogenesis, protection and control. *American Society for Microbiology, Washington, DC* (1994).
6. Sasindran, S. J. & Torrelles, J. B. Mycobacterium Tuberculosis Infection and Inflammation: what is Beneficial for the Host and for the Bacterium? *Frontiers in microbiology* **2**, 2, https://doi.org/10.3389/fmicb.2011.00002 (2011).
7. Gengenbacher, M. & Kaufmann, S. H. Mycobacterium tuberculosis: success through dormancy. *FEMS Microbiol Rev* **36**, 514–532, https://doi.org/10.1111/j.1574-6976.2012.00331.x (2012).
8. Gupta, U. D. & Katoch, V. M. Understanding the phenomenon of persistence in mycobacterial infections. *Indian J Lepr* **69**, 385–393 (1997).
9. Rana, A., Ahmed, M., Rub, A. & Akhter, Y. A tug-of-war between the host and the pathogen generates strategic hotspots for the development of novel therapeutic interventions against infectious diseases. *Virulence* **6**, 566–580, https://doi.org/10.1080/21505594.2015.1062211 (2015).
10. Dubnau, E. & Smith, I. Mycobacterium tuberculosis gene expression in macrophages. *Microbes and infection* **5**, 629–637 (2003).
11. Fu, Y. R., Gao, K. S., Ji, R. & Yi, Z. J. Differential transcriptional response in macrophages infected with cell wall deficient versus normal Mycobacterium Tuberculosis. *International journal of biological sciences* **11**, 22–30, https://doi.org/10.7150/ijbs.10217 (2015).
12. Kendall, S. L., Rison, S. C., Movahedzadeh, F., Frita, R. & Stoker, N. G. What do microarrays really tell us about M. tuberculosis? *Trends in microbiology* **12**, 537–544, https://doi.org/10.1016/j.tim.2004.10.005 (2004).
13. Ragno, S. *et al.* Changes in gene expression in macrophages infected with Mycobacterium tuberculosis: a combined transcriptomic and proteomic approach. *Immunology* **104**, 99–108 (2001).
14. Schnappinger, D. *et al.* Transcriptional Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the Phagosomal Environment. *The Journal of experimental medicine* **198**, 693–704, https://doi.org/10.1084/jem.20030846 (2003).
15. Tailleux, L. *et al.* Probing host pathogen cross-talk by transcriptional profiling of both Mycobacterium tuberculosis and infected human dendritic cells and macrophages. *PloS one* **3**, e1403, https://doi.org/10.1371/journal.pone.0001403 (2008).

16. Volpe, E. *et al*. Gene expression profiling of human macrophages at late time of infection with Mycobacterium tuberculosis. *Immunology* **118**, 449–460, https://doi.org/10.1111/j.1365-2567.2006.02378.x (2006).
17. Homolka, S., Niemann, S., Russell, D. G. & Rohde, K. H. Functional genetic diversity among Mycobacterium tuberculosis complex clinical isolates: delineation of conserved core and lineage-specific transcriptomes during intracellular survival. *PLoS pathogens* **6**, e1000988, https://doi.org/10.1371/journal.ppat.1000988 (2010).
18. Kumar, D. *et al*. Genome-wide analysis of the host intracellular network that regulates survival of Mycobacterium tuberculosis. *Cell* **140**, 731–743, https://doi.org/10.1016/j.cell.2010.02.012 (2010).
19. Thuong, N. T. *et al*. Identification of tuberculosis susceptibility genes with human macrophage gene expression profiles. *PLoS pathogens* **4**, e1000229, https://doi.org/10.1371/journal.ppat.1000229 (2008).
20. Rustad, T. R., Harrell, M. I., Liao, R. & Sherman, D. R. The enduring hypoxic response of Mycobacterium tuberculosis. *PloS one* **3**, e1502, https://doi.org/10.1371/journal.pone.0001502 (2008).
21. Voskuil, M. I., Visconti, K. C. & Schoolnik, G. K. Mycobacterium tuberculosis gene expression during adaptation to stationary phase and low-oxygen dormancy. *Tuberculosis (Edinb)* **84**, 218–227, https://doi.org/10.1016/j.tube.2004.02.003 (2004).
22. Normand, S. L. Meta-analysis: formulating, evaluating, combining, and reporting. *Statistics in medicine* **18**, 321–359 (1999).
23. Ramasamy, A., Mondry, A., Holmes, C. C. & Altman, D. G. Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS medicine* **5**, e184, https://doi.org/10.1371/journal.pmed.0050184 (2008).
24. Ferguson, L. External validity, generalizability, and knowledge utilization. *Journal of nursing scholarship: an official publication of Sigma Theta Tau International Honor Society of Nursing* **36**, 16–22 (2004).
25. Rhodes, D. R., Barrette, T. R., Rubin, M. A., Ghosh, D. & Chinnaiyan, A. M. Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer research* **62**, 4427–4433 (2002).
26. Guberman, J. M. *et al*. BioMart Central Portal: an open database network for the biological community. *Database: the journal of biological databases and curation* **2011**, bar041, https://doi.org/10.1093/database/bar041 (2011).
27. Haider, S. *et al*. BioMart Central Portal–unified access to biological data. *Nucleic acids research* **37**, W23–27, https://doi.org/10.1093/nar/gkp265 (2009).
28. Sharov, A. A., Dudekula, D. B. & Ko, M. S. A web-based tool for principal component and significance analysis of microarray data. *Bioinformatics* **21**, 2548–2549, https://doi.org/10.1093/bioinformatics/bti343 (2005).
29. Hong, F. *et al*. RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics* **22**, 2825–2827, https://doi.org/10.1093/bioinformatics/btl476 (2006).
30. Breitling, R., Armengaud, P., Amtmann, A. & Herzyk, P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS letters* **573**, 83–92, https://doi.org/10.1016/j.febslet.2004.07.055 (2004).
31. Cinghu, S. *et al*. Integrative framework for identification of key cell identity genes uncovers determinants of ES cell identity and homeostasis. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E1581–1590, https://doi.org/10.1073/pnas.1318598111 (2014).
32. Sassetti, C. M., Boyd, D. H. & Rubin, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular microbiology* **48**, 77–84 (2003).
33. Griffin, J. E. *et al*. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS pathogens* **7**, e1002251 (2011).
34. Kumar, R. & Nanduri, B. HPIDB-a unified resource for host-pathogen interactions. *BMC bioinformatics* **11**, S16 (2010).
35. Tekir, S. D. *et al*. PHISTO: pathogen–host interaction search tool. *Bioinformatics* **29**, 1357–1358 (2013).
36. Liberzon, A. *et al*. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell systems* **1**, 417–425, https://doi.org/10.1016/j.cels.2015.12.004 (2015).
37. Liberzon, A. *et al*. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740, https://doi.org/10.1093/bioinformatics/btr260 (2011).
38. Karp, P. D. *et al*. Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. *Nucleic acids research* **33**, 6083–6089, https://doi.org/10.1093/nar/gki892 (2005).
39. Arun, P. *et al*. Identification and functional analysis of essential, conserved, housekeeping and duplicated genes. *FEBS letters* **590**, 1428–1437 (2016).
40. Zhang, Y. J. *et al*. Global assessment of genomic regions required for growth in Mycobacterium tuberculosis. *PLoS pathogens* **8**, e1002946 (2012).
41. Wayne, L. G. & Hayes, L. G. An *in vitro* model for sequential study of shiftdown of Mycobacterium tuberculosis through two stages of nonreplicating persistence. *Infection and immunity* **64**, 2062–2069 (1996).
42. Voskuil, M. I., Bartek, I. L., Visconti, K. & Schoolnik, G. K. The response of mycobacterium tuberculosis to reactive oxygen and nitrogen species. *Frontiers in microbiology* **2**, 105, https://doi.org/10.3389/fmicb.2011.00105 (2011).
43. Sanz, J. *et al*. The transcriptional regulatory network of Mycobacterium tuberculosis. *PloS one* **6**, e22178, https://doi.org/10.1371/journal.pone.0022178 (2011).
44. Turkarslan, S. *et al*. Network portal: a database for storage, analysis and visualization of biological networks. *Nucleic acids research* **42**, D184–190, https://doi.org/10.1093/nar/gkt1190 (2014).
45. Talaat, A. M. *et al*. Mycobacterial bacilli are metabolically active during chronic tuberculosis in murine lungs: insights from genome-wide transcriptional profiling. *Journal of bacteriology* **189**, 4265–4274, https://doi.org/10.1128/JB.00011-07 (2007).
46. Walters, S. B. *et al*. The Mycobacterium tuberculosis PhoPR two-component system regulates genes essential for virulence and complex lipid biosynthesis. *Molecular microbiology* **60**, 312–330, https://doi.org/10.1111/j.1365-2958.2006.05102.x (2006).
47. Huet, G., Daffe, M. & Saves, I. Identification of the Mycobacterium tuberculosis SUF machinery as the exclusive mycobacterial system of [Fe-S] cluster assembly: evidence for its implication in the pathogen's survival. *Journal of bacteriology* **187**, 6137–6146, https://doi.org/10.1128/JB.187.17.6137-6146.2005 (2005).
48. Roche, B. *et al*. Iron/sulfur proteins biogenesis in prokaryotes: formation, regulation and diversity. *Biochimica et biophysica acta* **1827**, 455–469, https://doi.org/10.1016/j.bbabio.2012.12.010 (2013).
49. Mushtaq, K. *et al*. Rv2031c of Mycobacterium tuberculosis: a master regulator of Rv2028-Rv2031 (HspX) operon. *Frontiers in microbiology* **6**, 351, https://doi.org/10.3389/fmicb.2015.00351 (2015).
50. Sherman, D. R. *et al*. Regulation of the Mycobacterium tuberculosis hypoxic response gene encoding alpha -crystallin. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 7534–7539, https://doi.org/10.1073/pnas.121172498 (2001).
51. Yuan, Y., Crane, D. D. & Barry, C. E. 3rd Stationary phase-associated protein expression in Mycobacterium tuberculosis: function of the mycobacterial alpha-crystallin homolog. *Journal of bacteriology* **178**, 4484–4492 (1996).
52. Capyk, J. K. *et al*. Mycobacterial cytochrome p450 125 (cyp125) catalyzes the terminal hydroxylation of c27 steroids. *The Journal of biological chemistry* **284**, 35534–35542, https://doi.org/10.1074/jbc.M109.072132 (2009).
53. Garcia-Fernandez, E. *et al*. A highly conserved mycobacterial cholesterol catabolic pathway. *Environ Microbiol* **15**, 2342–2359, https://doi.org/10.1111/1462-2920.12108 (2013).
54. Chang, J. C., Harik, N. S., Liao, R. P. & Sherman, D. R. Identification of Mycobacterial genes that alter growth and pathology in macrophages and in mice. *The Journal of infectious diseases* **196**, 788–795, https://doi.org/10.1086/520089 (2007).
55. Rengarajan, J., Bloom, B. R. & Rubin, E. J. Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8327–8332, https://doi.org/10.1073/pnas.0503272102 (2005).

56. Yang, M., Guja, K. E., Thomas, S. T., Garcia-Diaz, M. & Sampson, N. S. A distinct MaoC-like enoyl-CoA hydratase architecture mediates cholesterol catabolism in Mycobacterium tuberculosis. *ACS chemical biology* **9**, 2632–2645, https://doi.org/10.1021/cb500232h (2014).

57. Dresen, C. *et al*. A flavin-dependent monooxygenase from Mycobacterium tuberculosis involved in cholesterol catabolism. *The Journal of biological chemistry* **285**, 22264–22275, https://doi.org/10.1074/jbc.M109.099028 (2010).

58. Ouellet, H., Johnston, J. B. & de Montellano, P. R. Cholesterol catabolism as a therapeutic target in Mycobacterium tuberculosis. *Trends in microbiology* **19**, 530–539, https://doi.org/10.1016/j.tim.2011.07.009 (2011).

59. Van der Geize, R. *et al*. A gene cluster encoding cholesterol catabolism in a soil actinomycete provides insight into Mycobacterium tuberculosis survival in macrophages. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 1947–1952, https://doi.org/10.1073/pnas.0605728104 (2007).

60. Constant, P. *et al*. Role of the pks15/1 gene in the biosynthesis of phenolglycolipids in the Mycobacterium tuberculosis complex. Evidence that all strains synthesize glycosylated p-hydroxybenzoic methyl esters and that strains devoid of phenolglycolipids harbor a frameshift mutation in the pks15/1 gene. *The Journal of biological chemistry* **277**, 38148–38158, https://doi.org/10.1074/jbc.M206538200 (2002).

61. Pang, J. M. *et al*. The polyketide Pks1 contributes to biofilm formation in Mycobacterium tuberculosis. *Journal of bacteriology* **194**, 715–721, https://doi.org/10.1128/JB.06304-11 (2012).

62. Vergnolle, O. *et al*. Biosynthesis of cell envelope-associated phenolic glycolipids in Mycobacterium marinum. *Journal of bacteriology* **197**, 1040–1050, https://doi.org/10.1128/JB.02546-14 (2015).

63. Choudhuri, B. S. *et al*. Overexpression and functional characterization of an ABC (ATP-binding cassette) transporter encoded by the genes drrA and drrB of Mycobacterium tuberculosis. *The Biochemical journal* **367**, 279–285, https://doi.org/10.1042/BJ20020615 (2002).

64. Churchill, S. A., Harper, J. P. & Churchill, P. F. Isolation and characterization of a Mycobacterium species capable of degrading three- and four-ring aromatic and aliphatic hydrocarbons. *Appl Environ Microbiol* **65**, 549–552 (1999).

65. Zahrt, T. C., Song, J., Siple, J. & Deretic, V. Mycobacterial FurA is a negative regulator of catalase-peroxidase gene katG. *Molecular microbiology* **39**, 1174–1185 (2001).

66. McKinney, J. D. *et al*. Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature* **406**, 735–738, https://doi.org/10.1038/35021074 (2000).

67. Bentley, R. The shikimate pathway–a metabolic tree with many branches. *Critical reviews in biochemistry and molecular biology* **25**, 307–384, https://doi.org/10.3109/10409239009090615 (1990).

68. Mdluli, K. & Spigelman, M. Novel targets for tuberculosis drug discovery. *Current opinion in pharmacology* **6**, 459–467, https://doi.org/10.1016/j.coph.2006.06.004 (2006).

69. Prakash, P., Pathak, N. & Hasnain, S. E. pheA (Rv3838c) of Mycobacterium tuberculosis encodes an allosterically regulated monofunctional prephenate dehydratase that requires both catalytic and regulatory domains for optimum activity. *The Journal of biological chemistry* **280**, 20666–20671, https://doi.org/10.1074/jbc.M502107200 (2005).

70. Yellaboina, S., Ranjan, S., Vindal, V. & Ranjan, A. Comparative analysis of iron regulated genes in mycobacteria. *FEBS letters* **580**, 2567–2576, https://doi.org/10.1016/j.febslet.2006.03.090 (2006).

71. Dosselaere, F. & Vanderleyden, J. A metabolic node in action: chorismate-utilizing enzymes in microorganisms. *Critical reviews in microbiology* **27**, 75–131, https://doi.org/10.1080/20014091096710 (2001).

72. Parish, T. & Stoker, N. G. The common aromatic amino acid biosynthesis pathway is essential in Mycobacterium tuberculosis. *Microbiology* **148**, 3069–3077, https://doi.org/10.1099/00221287-148-10-3069 (2002).

73. Schneider, C. Z., Parish, T., Basso, L. A. & Santos, D. S. The two chorismate mutases from both Mycobacterium tuberculosis and Mycobacterium smegmatis: biochemical analysis and limited regulation of promoter activity by aromatic amino acids. *Journal of bacteriology* **190**, 122–134, https://doi.org/10.1128/JB.01332-07 (2008).

74. Croft, D. *et al*. Reactome: a database of reactions, pathways and biological processes. *Nucleic acids research*, gkq1018 (2010).

75. Kanehisa, M. The KEGG database. *silico simulation of biological processes* **247**, 91–103 (2002).

76. Nishimura, D. BioCarta. *Biotech Software & Internet Report: The Computer Software Journal for Scient* **2**, 117–120 (2001).

77. Hussain, T., Shah, S. Z., Zhao, D., Sreevatsan, S. & Zhou, X. The role of IL-10 in Mycobacterium avium subsp. paratuberculosis infection. *Cell communication and signaling: CCS* **14**, 29, https://doi.org/10.1186/s12964-016-0152-z (2016).

78. Yoo, K. H., Kim, S. K., Chung, J. H. & Chang, S. G. Association of IL10, IL10RA, and IL10RB polymorphisms with benign prostate hyperplasia in Korean population. *J Korean Med Sci* **26**, 659–664, https://doi.org/10.3346/jkms.2011.26.5.659 (2011).

79. Li, J., Chai, Q. Y. & Liu, C. H. The ubiquitin system: a critical regulator of innate immunity and pathogen-host interactions. *Cellular & molecular immunology* **13**, 560–576, https://doi.org/10.1038/cmi.2016.40 (2016).

80. Karpov, L. M. [The role of Na K ATPase in thiamine and lipoic acid interrelations during their absorption in the gastrointestinal tract of mice]. *Fiziologicheskii zhurnal* **35**, 51–57 (1989).

81. Fallahi-Sichani, M., Kirschner, D. E. & Linderman, J. J. NF-kappaB Signaling Dynamics Play a Key Role in Infection Control in Tuberculosis. *Frontiers in physiology* **3**, 170, https://doi.org/10.3389/fphys.2012.00170 (2012).

82. Algood, H. M. *et al*. TNF influences chemokine expression of macrophages *in vitro* and that of CD11b+ cells *in vivo* during Mycobacterium tuberculosis infection. *Journal of immunology* **172**, 6846–6857 (2004).

83. Beg, A. A. & Baltimore, D. An essential role for NF-kappaB in preventing TNF-alpha-induced cell death. *Science* **274**, 782–784 (1996).

84. Gutierrez, M. G. *et al*. NF-kappa B activation controls phagolysosome fusion-mediated killing of mycobacteria by macrophages. *Journal of immunology* **181**, 2651–2663 (2008).

85. Mosser, D. M. & Edwards, J. P. Exploring the full spectrum of macrophage activation. *Nature reviews. Immunology* **8**, 958–969, https://doi.org/10.1038/nri2448 (2008).

86. Sanchez, D. *et al*. Role of TLR2- and TLR4-mediated signaling in Mycobacterium tuberculosis-induced macrophage death. *Cell Immunol* **260**, 128–136, https://doi.org/10.1016/j.cellimm.2009.10.007 (2010).

87. Bulut, Y. *et al*. Mycobacterium tuberculosis heat shock proteins use diverse Toll-like receptor pathways to activate pro-inflammatory signals. *The Journal of biological chemistry* **280**, 20961–20967, https://doi.org/10.1074/jbc.M411379200 (2005).

88. Means, T. K. *et al*. Human toll-like receptors mediate cellular activation by Mycobacterium tuberculosis. *Journal of immunology* **163**, 3920–3927 (1999).

89. Tapping, R. I. & Tobias, P. S. Mycobacterial lipoarabinomannan mediates physical interactions between TLR1 and TLR2 to induce signaling. *J Endotoxin Res* **9**, 264–268, https://doi.org/10.1179/096805103225001477 (2003).

90. Benko, S., Philpott, D. J. & Girardin, S. E. The microbial and danger signals that activate Nod-like receptors. *Cytokine* **43**, 368–373, https://doi.org/10.1016/j.cyto.2008.07.013 (2008).

91. Franchi, L., Warner, N., Viani, K. & Nunez, G. Function of Nod-like receptors in microbial recognition and host defense. *Immunological reviews* **227**, 106–128, https://doi.org/10.1111/j.1600-065X.2008.00734.x (2009).

92. Wang, X. L. *et al*. Bivariate whole genome linkage analyses for total body lean mass and BMD. *Journal of bone and mineral research: the official journal of the American Society for Bone and Mineral Research* **23**, 447–452, https://doi.org/10.1359/jbmr.071033 (2008).

93. Takeuchi, O. & Akira, S. MDA5/RIG-I and virus recognition. *Current opinion in immunology* **20**, 17–22, https://doi.org/10.1016/j.coi.2008.01.002 (2008).

94. Yoneyama, M. *et al.* The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature immunology* **5**, 730–737, https://doi.org/10.1038/ni1087 (2004).

95. Dixit, E. & Kagan, J. C. Intracellular pathogen detection by RIG-I-like receptors. *Adv Immunol* **117**, 99–125, https://doi.org/10.1016/B978-0-12-410524-9.00004-9 (2013).

96. Wong, M. M. & Fish, E. N. Chemokines: attractive mediators of the immune response. *Seminars in immunology* **15**, 5–14 (2003).

97. Elmore, S. Apoptosis: a review of programmed cell death. *Toxicologic pathology* **35**, 495–516, https://doi.org/10.1080/01926230701320337 (2007).

98. Jin, Z. & El-Deiry, W. S. Overview of cell death signaling pathways. *Cancer biology & therapy* **4**, 139–163 (2005).

99. Savitskaya, M. A. & Onishchenko, G. E. Mechanisms of Apoptosis. *Biochemistry. Biokhimiia* **80**, 1393–1405, https://doi.org/10.1134/S0006297915110012 (2015).

100. Choi, H. H. *et al.* Endoplasmic reticulum stress response is involved in Mycobacterium tuberculosis protein ESAT-6-mediated apoptosis. *FEBS letters* **584**, 2445–2454, https://doi.org/10.1016/j.febslet.2010.04.050 (2010).

101. Abarca-Rojano, E., Rosas-Medina, P., Zamudio-Cortez, P., Mondragon-Flores, R. & Sanchez-Garcia, F. J. Mycobacterium tuberculosis virulence correlates with mitochondrial cytochrome c release in infected macrophages. *Scand J Immunol* **58**, 419–427 (2003).

102. Lobet, E., Letesson, J. J. & Arnould, T. Mitochondria: a target for bacteria. *Biochem Pharmacol* **94**, 173–185, https://doi.org/10.1016/j.bcp.2015.02.007 (2015).

103. Weinberg, S. E., Sena, L. A. & Chandel, N. S. Mitochondria in the regulation of innate and adaptive immunity. *Immunity* **42**, 406–417, https://doi.org/10.1016/j.immuni.2015.02.002 (2015).

104. Mahajan, G. & Mande, S. C. Using structural knowledge in the protein data bank to inform the search for potential host-microbe protein interactions in sequence space: application to Mycobacterium tuberculosis. *BMC Bioinformatics* **18**, 201, https://doi.org/10.1186/s12859-017-1550-y (2017).

105. Nass, S. & Nass, M. M. Intramitochondrial Fibers with DNA Characteristics. Ii. Enzymatic and Other Hydrolytic Treatments. *J Cell Biol* **19**, 613–629 (1963).

106. Rana, A., Kumar, D., Rub, A. & Akhter, Y. Proteome-scale identification and characterization of mitochondria targeting proteins of Mycobacterium avium subspecies paratuberculosis: Potential virulence factors modulating host mitochondrial function. *Mitochondrion* **23**, 42–54, https://doi.org/10.1016/j.mito.2015.05.005 (2015).

Acknowledgements

Author Contributions

S.Y. designed the study; P.V.P.S.A., S.K.M., A.R. and S.Y. collected the data; S.Y. carried out the computation and generated Figures; P.V.P.S.A., Y.A., S.K.M., S.K. and S.Y. analysed the data; P.V.P.S.A., Y.A. and S.Y. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-22884-8.

Competing Interests: The authors declare no competing interests.

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