

RESEARCH ARTICLE

Uncovering New Pathogen–Host Protein–Protein Interactions by Pairwise Structure Similarity

Tao Cui, Weihui Li, Lei Liu, Qiaoyun Huang, Zheng-Guo He*

National Key Laboratory of Microbiology, Huazhong Agricultural University, Wuhan, 430070, China

* he.zhengguo@hotmail.com



OPEN ACCESS

Citation: Cui T, Li W, Liu L, Huang Q, He Z-G (2016) Uncovering New Pathogen–Host Protein–Protein Interactions by Pairwise Structure Similarity. PLoS ONE 11(1): e0147612. doi:10.1371/journal.pone.0147612

Editor: Byung-Jun Yoon, Texas A&M University, UNITED STATES

Received: August 31, 2015

Accepted: January 6, 2016

Published: January 22, 2016

Copyright: © 2016 Cui et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: TC is funded by the National Natural Science Foundation of China (31300079) and the China Postdoctoral Science Foundation (2013M540586). ZGH is funded by the National Natural Science Foundation of China (81471996 and 31421064) and the Chang Jiang Scholars Program. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Pathogens usually evade and manipulate host-immune pathways through pathogen–host protein–protein interactions (PPIs) to avoid being killed by the host immune system. Therefore, uncovering pathogen–host PPIs is critical for determining the mechanisms underlying pathogen infection and survival. In this study, we developed a computational method, which we named pairwise structure similarity (PSS)-PPI, to predict pathogen–host PPIs. First, a high-quality and non-redundant structure–structure interaction (SSI) template library was constructed by exhaustively exploring heteromeric protein complex structures in the PDB database. New interactions were then predicted by searching for PSS with complex structures in the SSI template library. A quantitative score named the PSS score, which integrated structure similarity and residue–residue contact-coverage information, was used to describe the overall similarity of each predicted interaction with the corresponding SSI template. Notably, PSS-PPI yielded experimentally confirmed pathogen–host PPIs of human immunodeficiency virus type 1 (HIV-1) with performance close to that of *in vitro* high-throughput screening approaches. Finally, a pathogen–host PPI network of human pathogen *Mycobacterium tuberculosis*, the causative agent of tuberculosis, was constructed using PSS-PPI and refined using filtration steps based on cellular localization information. Analysis of the resulting network indicated that secreted proteins of the STPK, ESX-1, and PE/PPE family in *M. tuberculosis* targeted human proteins involved in immune response and phagocytosis. *M. tuberculosis* also targeted host factors known to regulate HIV replication. Taken together, our findings provide insights into the survival mechanisms of *M. tuberculosis* in human hosts, as well as co-infection of tuberculosis and HIV. With the rapid pace of three-dimensional protein structure discovery, the SSI template library we constructed and the PSS-PPI method we devised can be used to uncover new pathogen–host PPIs in the future.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Upon infection by pathogens, various host immune response pathways such as toll-like receptors signaling, NF- κ B signaling, phagocytes, and cell-apoptosis pathways are activated. These pathways then work collectively to recognize, take up, and ultimately kill invading pathogens. However, pathogens have evolved diverse strategies for survival and replication under hostile host environments [1–3]. Pathogens usually counteract host immune defense and even acquire host nutrition through physical protein–protein interactions (PPIs). For example, the *Legionella* effector RavZ inhibits host autophagy through irreversible Atg8 deconjugation [4]. The VirA protein of *Shigella flexneri* targets host Rab1 for inactivation and then contributes to *Shigella* escape from autophagy [5]. *Neisseria* TbpA binds iron-containing transferrin and extracts iron directly from serum transferrin in a human host [6]. Therefore, discovery and analysis of pathogen–host PPIs can enable the determination of pathogen survival mechanisms in hosts.

Pathogen–host PPIs have been screened for several viruses [7–10] and bacteria [11–14] because they are critical for our understanding of pathogen-survival mechanisms. Information on such interactions are available through several organism-specific or comprehensive-public databases, including the Human Immunodeficiency Virus (HIV)-1 Human-Interaction Database [15], VirusMentha [16], and HPIDB [17]. Experimental approaches mostly used to map pathogen–host PPIs include affinity purification coupled with mass spectrometry [7] and yeast two-hybrid assay [8–11]. However, reliable experiment-based methods are time-consuming, expensive, and applicable only in limited species. Computational methods may play important roles in paving the way for experimental pathogen–host PPI verifications by highlighting high potential interactions and limiting the experimental scope, which can help to reduce expenses and accelerate the pace of discovery [18]. Computational methods for pathogen–host PPIs include the 'interologs' method [19, 20], domain-domain interaction method [20, 21], structure similarity-based method [22, 23] and machine learning-based method [24].

Tuberculosis (TB) causes two million deaths annually worldwide, and approximately one-third of the world's population is asymptotically infected with *Mycobacterium tuberculosis*, the main causative agent of this disease. TB infection also activates HIV replication and exacerbates HIV infection [25]. Successful intracellular survival of *M. tuberculosis* in macrophages involves modulation of several host-cell processes, including innate immune response and phagosome maturation [26–28]. However, the molecular mechanisms underlying such processes are unclear, which has been a stumbling block for development of efficient therapeutics.

In the present study, we developed a computational method, which we call pairwise structure similarity-PPI (PSS-PPI) to predict pathogen–host PPIs. PSS-PPI is a structure similarity based method with improvement compared to previous methods [22, 23], including more credible templates and a score function that measures similarity of both global structure and local interaction interface. PSS-PPI uses highly credible complex structures from PDB database as templates to predict new PPIs and integrates structure similarity with residue contact information to score the credibility of each new predicted interaction. Results showed that PSS-PPI successfully recovered experimentally confirmed pathogen–host PPIs of human immunodeficiency virus type 1 (HIV-1) with performance close to that of *in vitro* high-throughput screening. We then constructed a pathogen–human PPI network of human pathogen *M. tuberculosis* using PSS-PPI. Network-analysis results indicated that the serine/threonine protein kinase (STPK) family, the ESAT-6 secretion system (ESX) family, and the PE/PPE family proteins of *M. tuberculosis* interact with human proteins involved in immune response and phagocytosis pathways. Comparison with the HIV-human PPI network indicated that *M. tuberculosis* also targets human proteins involved in HIV infection. These findings can serve as a basis for

understanding the interaction between *M. tuberculosis* and human host at a molecular level and can also provide insights into the molecular mechanisms of co-infection by TB and HIV.

Materials and Methods

Construction of SSI template library

The SSI template library was constructed using experimentally resolved, high-quality, three-dimensional complex structures in the PDB database [29]. At the last data update of the current study, the PDB database contained a total of ~100,000 structures. By excluding structures of non-protein polymer (DNA/RNA), as well as monomeric and homomeric proteins, ~17,000 heteromeric protein complex structures were selected. Structures that were not resolved by X-ray diffraction or NMR-based methods were removed. In addition, structures resolved by X-ray diffraction with >4.0 Å resolution were removed. Finally, ~16,000 structures were collected and downloaded from the PDB database. Then, SSIs were identified from these complex structures by calculating inter-chain residue-residue contacts. Each complex structure was first split into monomer chains. Residue-residue contacts between monomer chains were then calculated by measuring atom-atom contacts using 6.05 Å as the cutoff distance [30]. A total of ~40,000 SSIs with >50 residue-residue contact number were identified from ~16,000 complex structures. Given the redundancy of the PDB database, our current collected SSIs were also redundant. Each SSI corresponded to a PPI, and corresponding SSIs of a PPI were usually largely identical with minor differences. At present, structure-similarity alignment is more computationally intensive than sequence-similarity alignment. Therefore, it is necessary to remove redundancy in our SSI template library. In the current study, we used a “max RRCN” strategy to construct a non-redundant SSI template library. That is, if two or more SSIs were complex structures corresponding to the same PPI, only the SSI with the **maximal residue-residue contacts number** (max RRCN) was selected as representative SSI. Finally, we obtained a high-confidence and non-redundant SSI template library containing 3,375 SSIs out of 6,267 structures.

Protein sequences and structures

Protein sequences of the *M. tuberculosis* proteome (3982 proteins) and human proteome (20,272 proteins) were obtained from the UniProt database [31]. The three-dimensional structures of proteins were identified by sequence-similarity alignment using BLAST+ [32]. First, a local BLAST database was constructed using protein sequences of structures in the PDB database [29]. Then, BLAST search for query proteins was performed against the database. Matching structures were required to have $>90\%$ sequence identity and >80 amino-acid residues. To minimize redundancy, the matched structures for each protein were grouped according to the corresponding protein segments. The structure with the highest sequence identity in each group was selected as the representative structure. If multiple structures had the same sequence identity, the structure with the best resolution was selected. A total of 443 structures for 423 *M. tuberculosis* proteins and 5,859 structures for 4,843 human proteins were obtained. The selected three-dimensional structures were downloaded from the PDB database [29].

Structure-structure alignment

Structure-structure alignment and similarity scoring was performed using the structural superimposition program TM-align [33].

PSS-PPIs and PSS score

We defined sA and sB as the structures of query proteins pA and pB , respectively, and tA and tB as the structures of a known SSI complex from the SSI template library. Then, a potential interaction between pA and pB was predicted if sA and sB shared structural similarity with tA and tB , respectively.

We defined the PSS-score, which integrated structure similarity and residue-residue contact coverage information to quantify the overall match of the query structures to the corresponding SSI template. PSS-score was calculated as follows:

$$\text{PSS - score} = \text{SIMA} \times \text{SIMB} \times \text{COV} \quad (1)$$

where SIMA is the structure-similarity score between query structure sA with template structure tA . In the current study, structure-structure alignments were performed using TM-align, which used a TM-score (ranging from 0.0 to 1.0) to quantify similarity between the two structures. For each pair of structure-structure alignment, the TM-align output of two TM-score values were normalized by the length (residue number) of the two aligned structures. We defined SIMA as the average values of two TM-scores from the alignment of sA and tA . In the same way, SIMB was defined as the average values of two TM-scores from the alignment of sB and tB . COV is the coverage of residue-residue contacts calculated by

$$\text{COV} = \frac{\text{RRCN}(sA, sB)}{\text{RRCN}(tA, tB)} \quad (2)$$

where $\text{RRCN}(tA, tB)$ is the residue-residue contact number of the SSI template. $\text{RRCN}(tA, tB)$ was calculated by measuring atom-atom contacts. A residue-residue contact was defined if the distance of any heavy atom pair from two residues was shorter than the cutoff value of 6.05 Å [30]. $\text{RRCN}(sA, sB)$ is the residue-residue contact number of the predicted interaction model. Notably, $\text{RRCN}(sA, sB)$ was mapped from residue-residue contacts in SSI template according to residue-alignment information generated by structure alignment.

Pathogen-human protein-protein interaction data

HIV-1 and human proteins known to regulate HIV-1 replication were obtained from the HIV-1 Human-Interaction Database [15]. Comprehensive pathogen-host protein interaction data sets were obtained from the VirusMentha Database [16].

Visualization of protein interactions

Protein interaction data was visualized using Cytoscape software [34]. Each protein was represented as a node. Two proteins were linked by an edge if they interacted with each other.

Results

Principle of PSS-PPI and PSS score

Structurally similar monomer proteins can share similar interactions even in the absence of significant sequence similarity. For example, the GCSF:GCSF-receptor complex structure and vIL-6:IL6ST complex structure share structurally similar monomer and interaction conformations (Fig 1A). Structure and sequence alignment indicated that GCSF with vIL-6 and GCSF-receptor with IL6ST shared 77% and 80% structure similarity, respectively, whereas their sequence identities were only 16% and 27%, respectively. This observation indicated the possibility that known complex structures can be used to discover potential new protein interactions using structure similarity as a bridge. In the current example, interaction between vIL-6 and

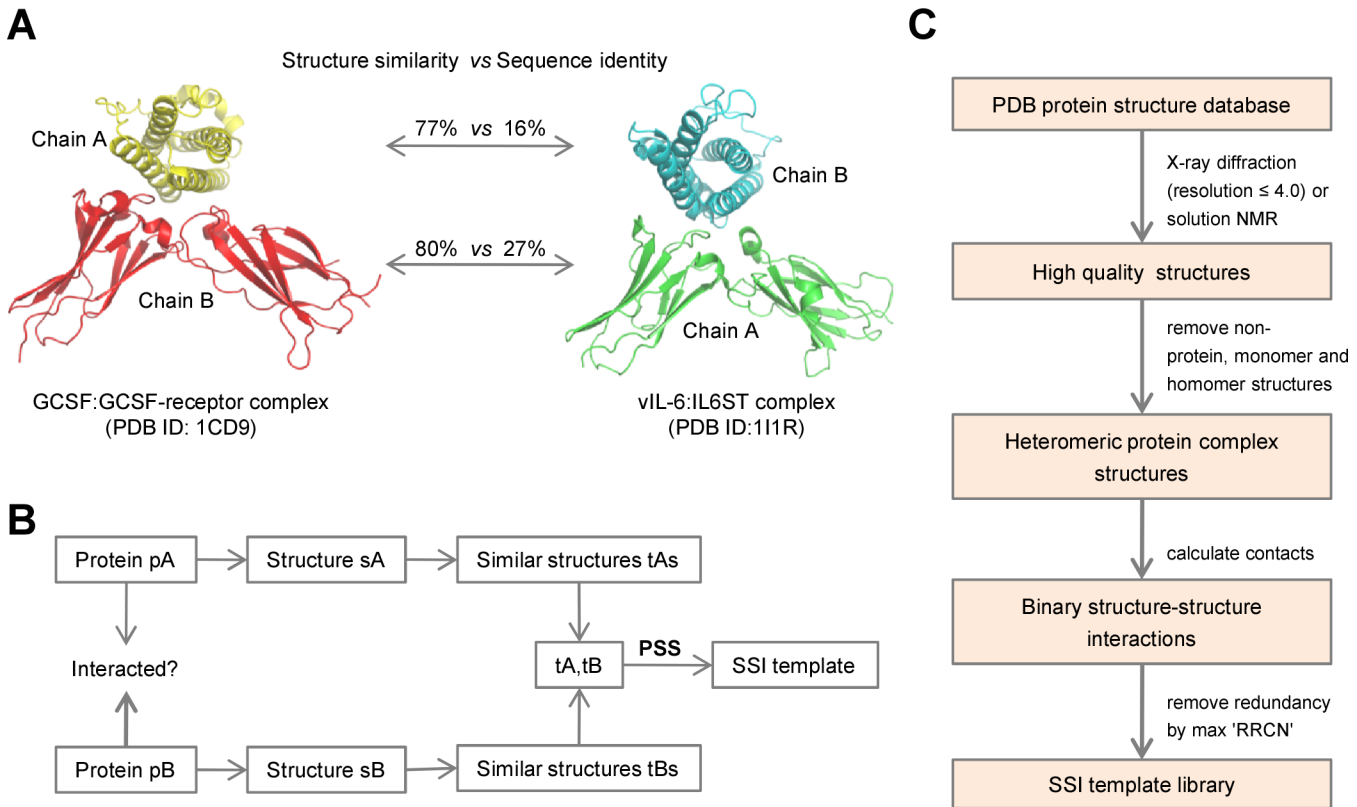


Fig 1. PSS-PPI and SSI template library. (A) An example of PSS. 1CD9 is the complex structure of GCSF (chain A) and GCSF-receptor (chain B). 111R is the complex structure of human herpesvirus 8 protein vIL-6 (chain B) and human interleukin-6 receptor beta subunit IL6ST (chain A). 111R and 1CD9 are structurally similar but have very low sequence similarity. GCSF means granulocyte colony-stimulating factor. (B) Schematic of PSS-PPI. (C) Flow chart showing the approach used for constructing the SSI template library.

doi:10.1371/journal.pone.0147612.g001

IL6ST can be inferred based on pairwise structure similarity (PSS) of vIL-6 and IL6ST monomers to monomers in the GCSF:GCSF-receptor complex.

Based on this observation, we proposed a computational method, which we named PSS-PPI, to predict pathogen–host PPIs. PSS-PPI is based on the idea that given two proteins pA and pB with structures sA and sB that shared structural similarity with tA and tB, respectively, from a known tA:tB complex, then pA and pB are also likely to interact with each other (Fig 1B).

The GCSF:GCSF-receptor complex mentioned in the above example was used as a structure–structure interaction (SSI) template. In order to apply PSS-PPI, a large number of such SSI templates were required. Accordingly, we comprehensively explored the experimentally resolved three-dimensional heteromer complex structures in the PDB database [29]. We obtained ~22,000 highly credible binary SSIs. After removing redundant SSIs, the final SSI template library contained more than 3,000 non-redundant SSIs (Fig 1C, Methods).

Interactions with a high degree of similarity between query structures and template structures can be reasonably predicted to be bonafide interactions. Therefore, it is important to establish a method for quantifying the overall similarity of query structures to structures in the SSI template. Scores from structure alignment may provide the necessary information but may be insufficient and may lead to false positives when a similar region is not located in the region that actually mediates the interaction. Therefore, we introduced residue–residue contact coverage to quantify the overlap between structural region sharing similarities with regions that actually mediate interaction. Finally, an integrated score, which we named the PSS score, was

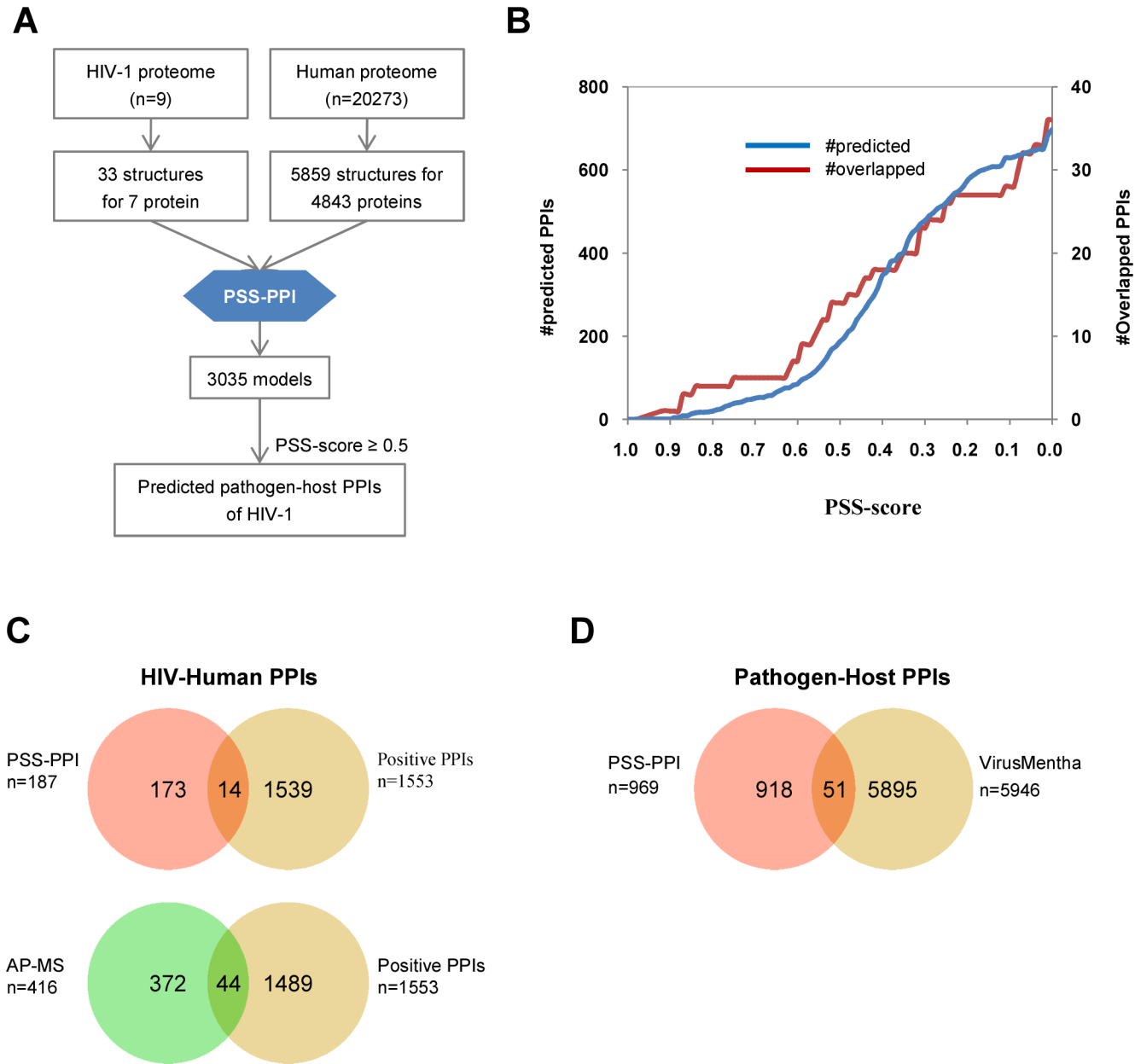


Fig 2. Performance of PSS-PPI. (A) Flow chart showing construction of the pathogen–human PPI network of HIV-1 using PSS-PPI. (B) Predicted PPI number and overlapped positive PPI number vs. PSS score. (C) Venn diagram of overlap between HIV-human PPIs from PSS-PPI (PSS-score cutoff = 0.5), HIV-human protein-interaction database and high-throughput screening.

doi:10.1371/journal.pone.0147612.g002

calculated by integrating structure-similarity information and residue–residue contact-coverage information to assign a global similarity metric for the newly predicted interaction with the corresponding SSI template (Methods).

Assessment of PSS-PPI

To examine the capability of PSS-PPI, we applied this method to predict the pathogen–host PPIs of HIV-1 (Fig 2A). The predicted PPIs were compared with the positive dataset that included 1,538 HIV-human PPIs obtained from the HIV-1 Human-Interaction Database [15].

The overlapped PPI number increased with decreased PSS score (Fig 2B). The ratio between overlapped PPIs with predicted PPIs increased with PSS score (S1 Fig), suggesting that higher PSS score implied better prediction. This finding indicated the capability of the PSS-PPI method to discover the pathogen–host PPIs of HIV-1. We further compared the performance of PSS-PPI with high-throughput experimental screening [7]. At a PSS-score cutoff of 0.5, 14 out of the 187 PPIs predicted by PSS-PPI overlapped with positive PPIs. By contrast, 44 out of 416 PPIs generated from high-throughput experiment overlapped with positive PPIs (Fig 2C). Notably, the ratio of overlapped PPIs to predicted PPIs from PSS-PPI ($14/187 = 0.075$) was close to the ratio of overlapped PPIs to PPIs generated from high-throughput screening ($44/416 = 0.106$). Three common PPIs were further identified between PSS-PPI and the high-throughput experimental screen, which suggested that PPIs obtained from the two approaches were largely complementary.

To further determine the applicability of the PSS-PPI method for other pathogens, we applied PSS-PPI to predict pathogen–host PPIs in the scope of the VirusMentha Database. This database collects literature-curated, credible pathogen–host PPIs that are mostly derived from small-scale targeted studies [16]. Using 0.5 as a PSS-score cutoff, 51 out of 969 new predicted PPIs overlapped with positive PPIs in the VirusMentha Database (Fig 2D). These results demonstrate the capability of the PSS-PPI method to discover pathogen–host PPIs.

Predicting pathogen–host protein interactions of *M. tuberculosis*

We used PSS-PPI to predict pathogen–host PPIs of the human pathogen *M. tuberculosis* (Fig 3A). First, protein sequences of *M. tuberculosis* and the human proteome were downloaded from the UniProt Database. The representative structures for each protein were assigned by sequence-similarity alignment using BLAST search. A total of 443 structures for 423 *M. tuberculosis* proteins and 5,859 structures for 4,843 human proteins were obtained. Using TM-align, all-against-all structure alignment was performed between all 6,302 query protein structures (from *M. tuberculosis* and human) and 8,776 structures in the SSI template library. Structures with TM score ≥ 0.5 with each query structure were deemed similar [33]. Then, 2,595,537 structure–structure pairs between *M. tuberculosis* and human were used to query the SSI template library to identify PSS, which led to the identification of 411,020 matching models. The PSS score for each model was calculated and assigned to the corresponding PPI. If more than one model corresponded to a PPI, the maximal score was used. To reduce false-positive interactions, we used two filtration steps to further refine our predicted *M. tuberculosis*–human PPIs. The first filter removed PPIs whose structure similarity between *M. tuberculosis* protein and human protein were greater than 0.6. This filter was based on our analysis of structure similarity between pathogen and host protein of known pathogen–host PPIs. More specifically, we calculated the structure similarity between pathogen protein and host protein of each known pathogen–host PPI from the VirusMentha database. Then, we analyzed the distribution of these structure similarity scores. The result indicated that structure similarity between pathogen and host protein of almost all known pathogen–host PPIs ($1538/1541 = 99.81\%$) were less than 0.6 (S2A Fig). This suggested that potential PPIs whose pathogen protein and host protein have structural similarity score greater than 0.6 were likely to be false-positive PPIs. However, more than one-third of the predicted PPIs had structure similarity score between *M. tuberculosis* protein and human protein greater than 0.6 in our predicted *M. tuberculosis*–host PPIs (S2B Fig). Therefore, PPIs with structure similarity score between *M. tuberculosis* protein and human protein greater than 0.6 were removed to reduce false-positive predictions.

The second filter removed PPIs according to protein-localization information. Unlike viruses, bacteria have compartmentalized cell structure. This structure should be considered

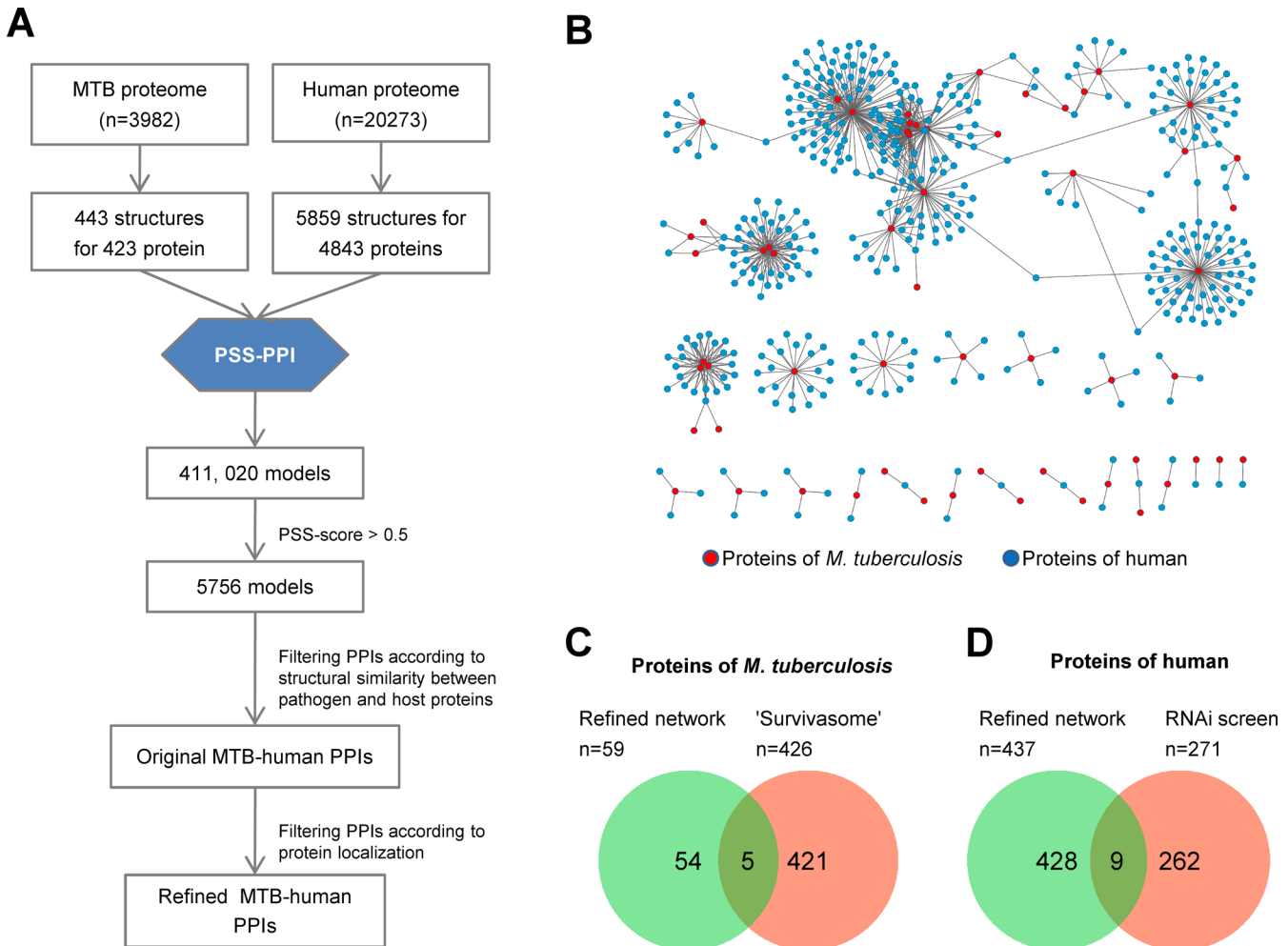


Fig 3. Pathogen–host PPI network of *M. tuberculosis*. (A) Flow chart showing the construction of pathogen–host PPI network of *M. tuberculosis* using PSS-PPI. (B) Graphical representation of predicted pathogen–host PPI network of *M. tuberculosis*. Nodes represent proteins of *M. tuberculosis* (red) and human (blue). Edges represent predicted interactions. (C) Overlap between *M. tuberculosis* proteins in predicted PPI network and “survivasome” of *M. tuberculosis*. (D) Overlap between human proteins in the predicted PPI network and human proteins involved in *M. tuberculosis* infection identified by RNAi screening.

doi:10.1371/journal.pone.0147612.g003

when assessing the feasibility of interactions under *in vivo* conditions in order to decrease false positives. Proteins from the secretome of pathogenic bacteria are more likely to interact with a host. We manually curated culture supernatant proteins and exported proteins of *M. tuberculosis* to build the secretome of *M. tuberculosis* using data from several publications [35–37]. Subsequently, PPIs in which the pathogen protein was not found in our constructed secretome were removed.

After these two filtration steps, we finally obtained 773 predicted PPIs between 59 *M. tuberculosis* proteins and 437 human proteins using 0.5 as the PSS score cutoff (Fig 3B and S1 Table). The PSS score distribution of all PPIs can be found in S3 Fig. The degree of a protein in the PPI network refers to the number of its interacting partners. The degree distribution of proteins in our predicted PPI network approximately followed a power law function (S4 Fig). This suggested that the network was scale-free, which is a general characteristic of biological networks [38].

We manually curated proteins involved in the infection and survival of *M. tuberculosis* in a host to build the “survivasome” of *M. tuberculosis* from several transposon screening studies [39–41]. We found that 5 out of 59 *M. tuberculosis* proteins in our network overlapped with the “survivasome” of *M. tuberculosis* (Fig 3C). In addition, we collected host factors responsible for *M. tuberculosis* infection from genome-wide RNAi screening studies [42]. We obtained 275 proteins that overlapped with 9 out of 437 human proteins in our network (Fig 3D).

To examine the host-cell process targeted by *M. tuberculosis*, we analyzed human proteins in the *M. tuberculosis*–human PPI network in the context of GO terms [43]. We counted the frequency of involved GO biological processes (S2 Table). The top groups also included proteins involved in small GTPase-mediated signal transduction, protein transport, innate immune response, and apoptotic process. A considerable number of proteins were also involved in the regulation of signaling pathways linked to NF- κ B and phagocytosis.

M. tuberculosis targeted host immune-response pathways

To analyze the interaction between *M. tuberculosis* and human anti-microbial pathways, we collected human proteins involved in the NF- κ B signaling pathway according to GO annotation and constructed a subnetwork involved in these processes (Fig 4A). The highest connected node in this local network was the PknB protein of *M. tuberculosis*, which interacted with nine host proteins. The largest group of *M. tuberculosis* proteins came from the ESX family, including five ESX-family proteins (EsxB, EsxJ, EsxK, EsxP, and EsxW). Host proteins included the inhibitor of κ B (I κ B) family proteins I κ B α and BCL-3, ubiquitin-conjugating enzyme E2, non-ATPase regulatory subunit 10 of 26S proteasome, and transforming proteins RhoA and RhoC. To intuitively describe the roles of these target proteins, a schematic of the NF- κ B signaling pathway is shown (Fig 4B). I κ B α is the inhibitor of NF- κ B, whereas ubiquitin-conjugating enzyme and proteasome are essential components of the ubiquitin–proteasome pathway, which is required for degrading I κ B α and activating NF- κ B [44, 45]. These results suggested that *M. tuberculosis* STPK and ESX-family proteins targeted multiple host regulators involved in the activation of NF- κ B pathways.

ESX-family proteins targeted host proteins involved in phagocytosis

M. tuberculosis can escape from phagocytosis-mediated antimicrobial activity and survive in a host [26, 46, 47]. To explore the interaction between *M. tuberculosis* and the host phagocytosis pathway, we collected host proteins involved in phagocytosis according to GO annotation and predicted interactions that targeted these proteins (Fig 5A). This *M. tuberculosis*–human PPI network involved in phagocytosis included 55 interactions between 12 *M. tuberculosis* proteins and 20 human proteins. *M. tuberculosis* proteins included six ESX-family proteins (EsxA/Rv3875, EsxB/Rv3874, EsxJ/Rv1038c, EsxK/Rv1197, EsxW/Rv3620c, and EsxP/Rv2347c), two PE/PPE family proteins (PE25 and PPE41), and one hypothetical protein Rv1794. Clearly, ESX-family proteins accounted for almost half of *M. tuberculosis* proteins that targeted host phagocytosis pathways. Host proteins included multiple small GTPases (RAB5, RAB7, and RAB11), several subunits of phosphatidylinositol 3-kinase PI3K (PIK3R1/P85A, PIK3R2/P85B, and PK3CA), and a component of the ESCRT (endosomal sorting required for transport) complex (CHMP3). To intuitively describe the roles of these target proteins, a schematic of the host phagocytosis-signaling pathway is shown (Fig 5B). Functions of human proteins indicated that *M. tuberculosis* may interfere with phagocytosis pathways at multiple stages, including phagosome formation (Rac1 and Cdc42), phagosome maturation (small GTPases RAB5, RAB7, and RAB11, as well as PI3K and CHMP3), and regulation of NADPH oxidase enzyme activity (Rac1) (Fig 5B).

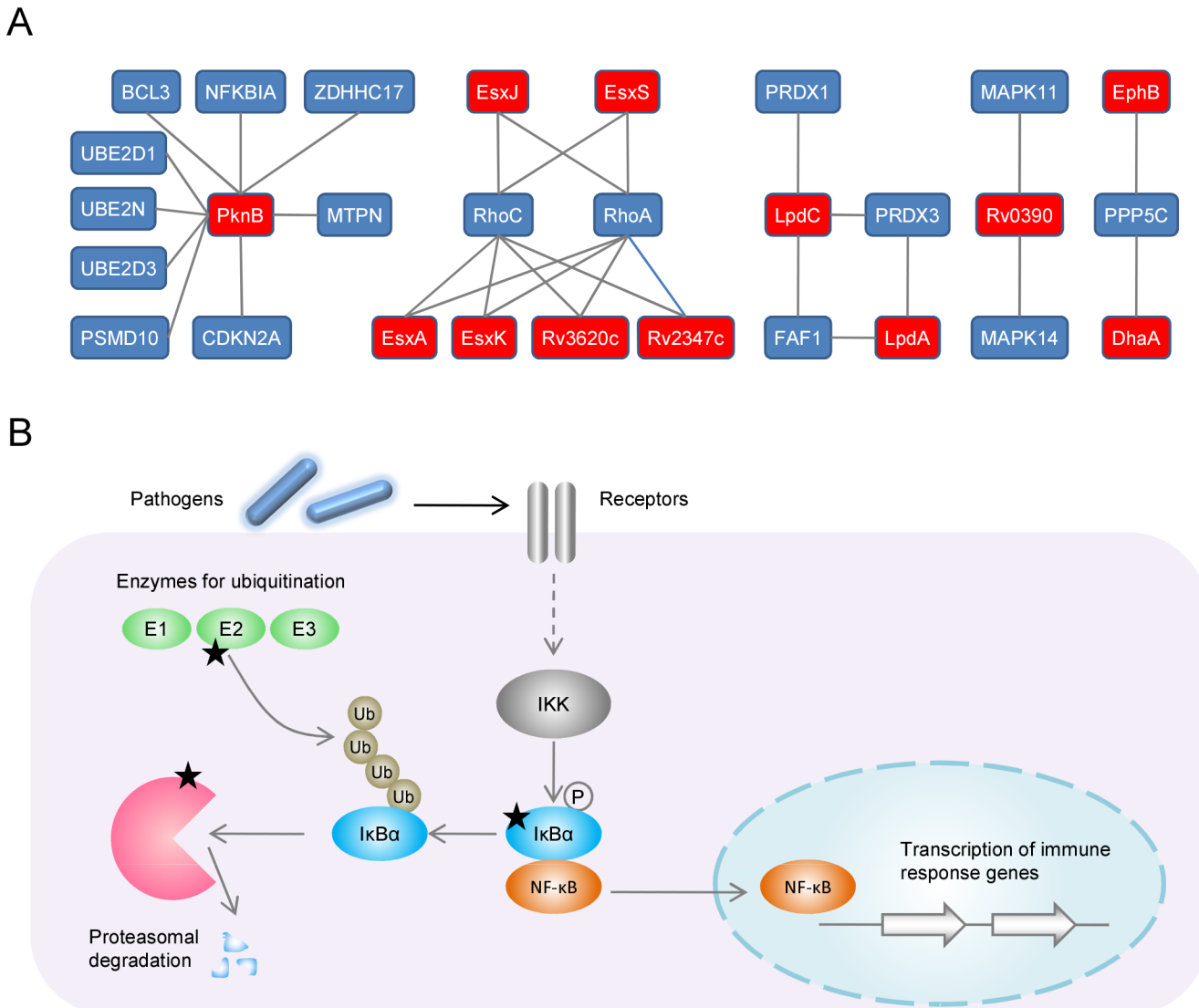


Fig 4. Predicted PPIs between *M. tuberculosis* and human immune response pathways. (A) Local PPIs network between *M. tuberculosis* proteins (red nodes) and human proteins (blue nodes) involved in the NF-κB signaling pathway. (B) Schematic of the NF-κB signaling pathway. Proteins found to be targeted by *M. tuberculosis* are indicated by black stars.

doi:10.1371/journal.pone.0147612.g004

M. tuberculosis targeted proteins involved in HIV infection

HIV infection increases the risk of latent TB reactivation by 20-fold. TB infections activate HIV replication and exacerbate the progression of HIV infection [48, 49]. Exploring the molecular mechanisms underlying these effects can facilitate control of both pathogens. Accordingly, we compared *M. tuberculosis*-human PPIs with HIV-human PPIs. We constructed a HIV-human PPI network using PPI data from the NCBI HIV-1 Human-Interaction Database [15]. Interestingly, 90 human proteins in the *M. tuberculosis*-human PPI network overlapped with human proteins in the HIV-human PPI network (Fig 6A and 6B). We also found that *M. tuberculosis* targeted 53 human proteins involved in regulating HIV replication (Fig 6C). These results suggest that *M. tuberculosis* is likely to affect the function of proteins targeted directly by HIV, as well as those that regulate HIV replication.

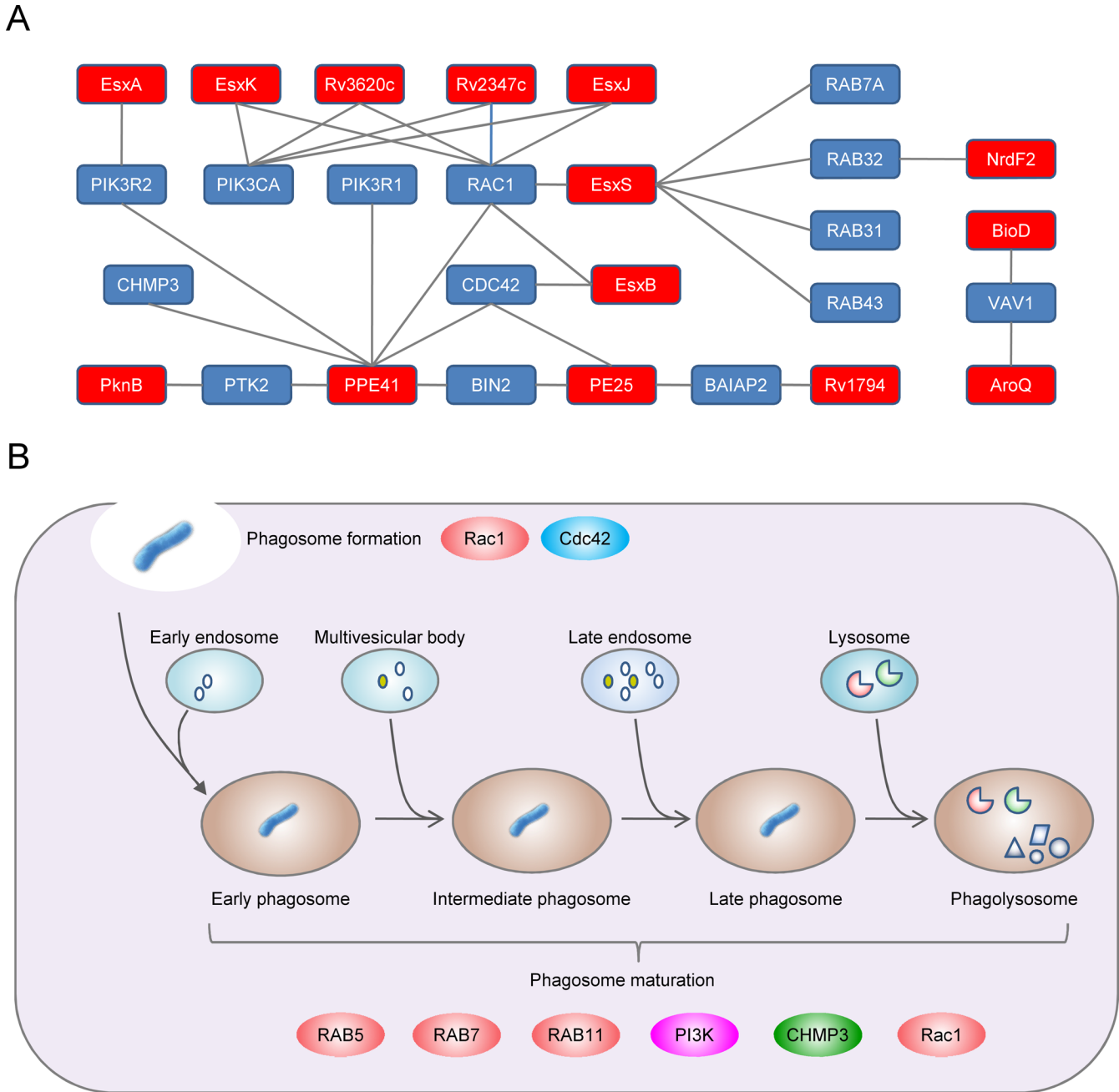


Fig 5. *M. tuberculosis* targets human phagocytosis pathways. (A) Local PPI network between *M. tuberculosis* proteins (red nodes) and human proteins (blue nodes) involved in phagocytosis. (B) Schematic showing phagocytosis pathways. The RHO family GTPases RAC1 and CDC42 play roles in phagosome formation. Phagosomes undergo sequential fusion with early endosomes, late endosomes and lysosomes. The small GTPase RAB5A is involved in the fusion of phagosomes with early endosomes. The small GTPase RAB7A is known to mediate trafficking between phagosomes and late endosomes or lysosomes. The GTPase RAB11A mediates recycling of endosomes to the plasma membrane. PI3K (PIK3R1, PIK3R2 and PIK3CA) regulates metabolism of phosphoinositides, which play essential roles during phagocytosis. CHMP3 is a component of the ESCRT complex. RAC1 is also involved in stimulation of NADPH oxidase activity in macrophages.

doi:10.1371/journal.pone.0147612.g005

Discussion

Discovery of new pathogen–host PPIs is essential for understanding the infection and survival mechanisms of pathogens in a host. In this study, we constructed a comprehensive

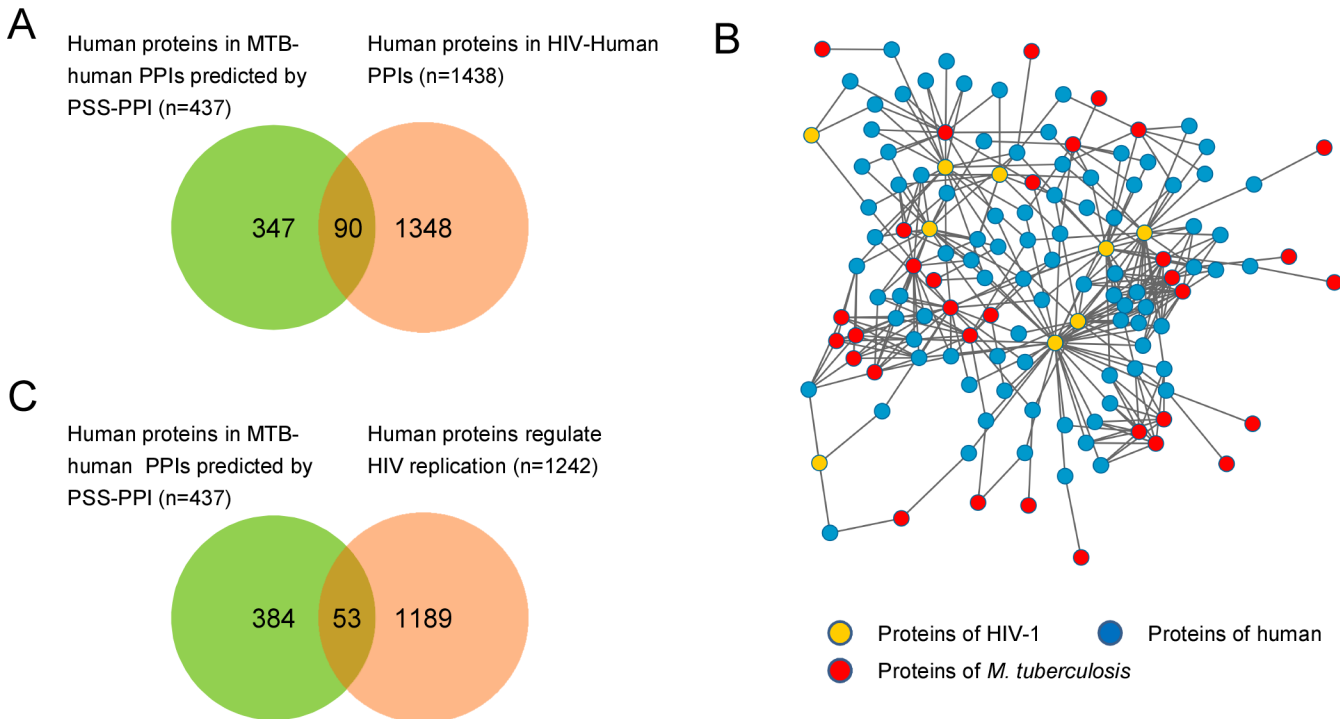


Fig 6. *M. tuberculosis* targets human proteins involved in HIV infection. (A) *M. tuberculosis* was predicted to target human proteins that are also targeted by HIV-1. (B) The *M. tuberculosis*-HIV-human PPI network. Nodes represent proteins of *M. tuberculosis* (red), HIV-1 (yellow), and human (blue). (C) *M. tuberculosis* was predicted to target human proteins that regulate HIV replication.

doi:10.1371/journal.pone.0147612.g006

and non-redundant SSI template library and developed PSS-PPI to uncover new pathogen-host PPIs. Using an HIV-human model system, we demonstrated that PSS-PPI successfully predicted pathogen-host PPIs with performance close to that of *in vitro* experimental high-throughput screening. In addition, PSS-PPI predicted PPIs that were complementary to those predicted by experiment-based screening. Given the scarcity of pathogen-host PPI data for most pathogens, PSS-PPI shows promise for opening new doors to explore the survival mechanism of these pathogens.

All interaction templates known to date have originated mainly from model organisms such as human, yeast, and *Escherichia coli*. Therefore, search for target proteins similar to unique pathogen proteins that are associated with virulence in the interaction template-library have usually failed. This weakness can be overcome by utilizing structural similarity information. PSS-PPI is a structure-based PPI prediction method that uses structural similarity between proteins as a bridge to identify new interactions. Compared with sequence similarity-based methods, this structure similarity-based approach enabled us to explore new interactions of proteins that lacked significant sequence similarity with a known interaction-template library. This advantage is of vital importance in predicting pathogen-host PPIs. Indeed, PSS-PPI enabled the identification of potential interactions of unique ESX and PE/PPE family proteins from *Mycobacteria* in the current study.

PSS-PPI can be classified into template-based PPI prediction methods that transform known interactions to predict new interactions. Therefore, high quality of the interaction-template library (source interactions), including coverage and credibility, is essential for successful application of template-based protein-interaction prediction. In the current study, the SSI template library was constructed with these goals in mind. First, PSS-PPI used templates generated

only from known complex structures and was thus more reliable than templates obtained from high-throughput screening. Second, to improve coverage, all currently available structures in the PDB database were considered for identifying SSI templates. In addition, to control quality, complex structures and SSI templates were strictly filtered according to experimental methods, resolution, structure length, and residue–residue contact information.

PSS-PPI used precise structure–structure alignment to define structure similarity. Compared with sequence alignment, structure alignment is computationally intensive. Prediction of new interactions requires the alignment of query proteins to all proteins in the template library. Therefore, the size of the template library is closely associated with computational cost. In this study, we used a “max RRCN” strategy for the first time to reduce template-library redundancy. This strategy reduced the size of the template library while preserving its comprehensiveness and diversity. This optimized template library then enabled practical application of PSS-PPI.

Taking advantage of precise structure–structure alignment information, PSS-PPI predicted new interactions and provided a potential three-dimensional complex model for each predicted interaction. Each potential complex model was generated by superposing two query structures to the corresponding complex template. Despite being rough, this predicted complex model could provide useful information regarding the interaction interface, which was helpful for us to understand interaction mechanisms at residue level and provide information for structure-based inhibitor design.

Structure is the only required input element of PSS-PPI. Therefore, PSS-PPI can be used to predict the pathogen–host PPIs of many other pathogens for which structures are available. Humans are the most important host in pathogen–host studies. A considerable number of human protein structures are available at present, but availability of structural information for pathogen proteins differs greatly. Therefore, utility of PSS-PPI is limited when structure information is incomplete or absent. This obstacle can be overcome by quickly expanding structural information in the PDB database. In addition, structures from homologous modeling can be considered in the future. Nevertheless, our results indicate that the PSS-PPI method can be used to explore the pathogen–host PPIs of pathogens with available structures.

M. tuberculosis is an obligate intracellular pathogen that can infect and survive in a host. The prediction and analysis of pathogen–host PPIs of *M. tuberculosis* is valuable in exploring the survival mechanisms of *M. tuberculosis* in a host. In the present study, we predicted the pathogen–host PPIs of the human pathogen *M. tuberculosis* using PSS-PPI. This network can serve as a basis for exploring the survival mechanism of *M. tuberculosis* within a host.

We found that multiple host factors involved in NF- κ B signaling pathways are targeted by *M. tuberculosis*. NF- κ B family proteins are transcriptional factors that regulate the expression of immune-response genes. I κ B α protein inactivates the NF- κ B transcription factor by masking the nuclear localization signals of NF- κ B proteins and keeping them sequestered in an inactive state in the cytoplasm. Upon cellular stimulation by immune and pro-inflammatory responses, I κ B kinase (IKK) specifically phosphorylates I κ B α protein. This phosphorylation results in I κ B α ubiquitination and degradation. The dissociation of I κ B α from NF- κ B enables NF- κ B to translocate to the nucleus and activate the transcription of immune-response genes [50]. IKK-mediated phosphorylation of I κ B α proteins represents a convergence point for most signal-transduction pathways leading to NF- κ B activation [51]. In the present study, we identified a potential interaction between Ser/Thr protein kinase PknB and I κ B α . This interaction may affect the phosphorylation of I κ B α and thereby perturb the activation of NF- κ B signaling. Moreover, I κ B α degradation requires the ubiquitin–proteasome pathway. Several subunits of the host ubiquitin-conjugating enzyme E2 and proteasome were also found to be targeted by *M. tuberculosis* (Fig 4A). Interestingly, PknB has been identified by mass spectrometry in a

culture filtrate of *M. tuberculosis* H37Rv [37] and *M. tuberculosis* H37Rv-infected guinea-pig lungs [52]. This supports the validity of the predicted interactions under *in vivo* conditions.

The pathogenicity of *M. tuberculosis* is attributed largely to its ability to survive within macrophages [26, 46, 47]. Interestingly, we found that groups of human proteins involved in phagocytosis were targeted by the ESX-family proteins of *M. tuberculosis*. These interactions highlight the essential roles of ESX-family proteins in the pathogen–host interaction of *M. tuberculosis*, which is highly consistent with the findings of previous genetic-screening studies [53, 54]. Some of the predicted *M. tuberculosis* target proteins such as PI3K and CHMP3 were particularly interesting. PI3K activity is essential for proper phagosomal maturation, and *Mycobacteria* have been shown to use Man-LAN to interfere with PI3K signaling pathways [26, 46]. CHMP3 is a component of the ESCRT complex, and ESCRT has been shown to play roles in restricting the growth of *M. tuberculosis* and to be targeted by another ESX family protein EsxH [54, 55]. Therefore, these new predicted interactions suggest an alternative strategy used by *M. tuberculosis* to interfere with PI3K signaling pathways and phagosomal maturation. In addition, we found that *M. tuberculosis* targeted Rac1, which is involved in activating NADPH oxidase. NADPH oxidase mediates the phagocytic killing of ingested pathogens by producing reactive oxygen species (ROS). Therefore, this interaction may contribute to the observation that *M. tuberculosis* is relatively resistant to the microbicidal effects of ROS [56, 57]. Taken together, our results indicate that these interactions are likely to interfere with phagosomal maturation and ROS signaling and are of interest for further experimental exploration.

TB and HIV co-infection is a major challenge in the global control of TB [25]. As mentioned above, HIV infection increases the risk of latent TB reactivation 20-fold. TB infection activates HIV replication and exacerbates the progression of HIV infection [48, 49]. Interactions of HIV and *M. tuberculosis* with the human immune system have already been explored at the cellular level [48, 49]. However, their interaction at the molecular level is largely unknown. Interestingly, we found that *M. tuberculosis* targets host proteins that regulate HIV replication. We speculate that these interactions may positively regulate the function of host proteins required for HIV replication or negatively regulate the function of host proteins limiting HIV replication. In addition, a group of common host proteins targeted by both *M. tuberculosis* and HIV was identified by comparing the *M. tuberculosis* PPI network constructed here with known HIV-human protein interactions. We speculate that co-targets of *M. tuberculosis* and HIV enhance the interference effect on host cell pathways and facilitate the infection of both pathogens. These findings suggest a concerted attack by *M. tuberculosis* and HIV on the host immune system that may contribute to TB-HIV co-infection.

Conclusions

We constructed an optimized SSI template library and developed a structure-based computational method named PSS-PPI to predict pathogen–host PPIs. We demonstrate that PSS-PPI can effectively discover pathogen–host PPIs with performance close to that of high-throughput screening. Specifically, we constructed a pathogen–host PPI network of the human pathogen *M. tuberculosis* using PSS-PPI. Analysis of the network indicated that *M. tuberculosis* targeted host immune response and phagocytosis pathways. In addition, *M. tuberculosis* targeted host proteins that interact with HIV proteins as well as those that regulate HIV replication. These PPIs provide a resource for exploring the survival mechanisms of *M. tuberculosis* in a host and TB-HIV co-infection. With continuous increase in the availability of three-dimensional protein structures in public databases, the SSI template library and the PSS-PPI method reported here can be used to predict pathogen–host PPIs at a larger scale and for diverse pathogens.

Supporting Information

S1 Fig. Ratio of overlapped PPIs and predicted PPIs. The ratio between overlapped PPIs and predicted PPIs increased along with increase in PSS score, suggesting that higher PSS score implied better prediction.

(TIF)

S2 Fig. Distribution of structure similarity scores between pathogen protein and host protein in pathogen-host PPIs. (A) PPIs of known pathogen-host PPIs from public databases. (B) PPIs predicted using the PSS-PPI method in the current study.

(TIF)

S3 Fig. PSS score distribution of predicted *M. tuberculosis*-human PPIs.

(TIF)

S4 Fig. Degree distribution of the *M. tuberculosis*-human PPI network. The degree distribution approximately followed a power law function ($y = ax^b$, $a = 87.975$, $b = -1.247$, $R^2 = 0.774$). This suggested that the predicted PPI network was a scale-free network.

(TIF)

S1 Table. *M. tuberculosis*-human PPI network predicted by the PSS-PPI method.

(XLS)

S2 Table. GO term count of human proteins in predicted pathogen-host PPIs of *M. tuberculosis*.

(XLS)

Acknowledgments

TC is funded by the National Natural Science Foundation of China (31300079) and the China Postdoctoral Science Foundation (2013M540586). ZGH is funded by the National Natural Science Foundation of China (81471996 and 31421064) and the Chang Jiang Scholars Program.

Author Contributions

Conceived and designed the experiments: TC ZGH. Performed the experiments: TC WL LL. Analyzed the data: TC WL LL QH ZGH. Contributed reagents/materials/analysis tools: TC ZGH. Wrote the paper: TC ZGH.

References

1. Bhavsar AP, Guttman JA, Finlay BB. Manipulation of host-cell pathways by bacterial pathogens. *Nature*. 2007 Oct 18; 449(7164):827–34. PMID: [17943119](#)
2. Flannagan RS, Cosío G, Grinstein S. Antimicrobial mechanisms of phagocytes and bacterial evasion strategies. *Nat Rev Microbiol*. 2009 May; 7(5):355–66. doi: [10.1038/nrmicro2128](#) PMID: [19369951](#)
3. Diacovich L, Gorvel JP. Bacterial manipulation of innate immunity to promote infection. *Nat Rev Microbiol*. 2010 Feb; 8(2):117–28. doi: [10.1038/nrmicro2295](#) PMID: [20075926](#)
4. Choy A, Dancourt J, Mugo B, O'Connor TJ, Isberg RR, Melia TJ, et al. The *Legionella* effector RavZ inhibits host autophagy through irreversible Atg8 deconjugation. *Science*. 2012 Nov 23; 338(6110):1072–6. doi: [10.1126/science.1227026](#) PMID: [23112293](#)
5. Dong N, Zhu Y, Lu Q, Hu L, Zheng Y, Shao F. Structurally distinct bacterial TBC-like GAPs link Arf GTPase to Rab1 inactivation to counteract host defenses. *Cell*. 2012 Aug 31; 150(5):1029–41. doi: [10.1016/j.cell.2012.06.050](#) PMID: [22939626](#)
6. Noinaj N, Easley NC, Oke M, Mizuno N, Gumbart J, Boura E, et al. Structural basis for iron piracy by pathogenic *Neisseria*. *Nature*. 2012 Feb 12; 483(7387):53–8. doi: [10.1038/nature10823](#) PMID: [22327295](#)

7. Jäger S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, Clarke SC, et al. Global landscape of HIV-human protein complexes. *Nature*. 2011 Dec 21; 481(7381):365–70. doi: [10.1038/nature10719](https://doi.org/10.1038/nature10719) PMID: [22190034](https://pubmed.ncbi.nlm.nih.gov/22190034/)
8. Uetz P, Dong YA, Zeretzke C, Atzler C, Baiker A, Berger B, et al. Herpesviral protein networks and their interaction with the human proteome. *Science*. 2006 Jan 13; 311(5758):239–42. PMID: [16339411](https://pubmed.ncbi.nlm.nih.gov/16339411/)
9. Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM, Wu L, et al. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell*. 2009 Dec 24; 139(7):1255–67. doi: [10.1016/j.cell.2009.12.018](https://doi.org/10.1016/j.cell.2009.12.018) PMID: [20064372](https://pubmed.ncbi.nlm.nih.gov/20064372/)
10. Davis ZH, Verschueren E, Jang GM, Kleffman K, Johnson JR, Park J, et al. Global mapping of herpesvirus-host protein complexes reveals a transcription strategy for late genes. *Mol Cell*. 2015 Jan 22; 57(2):349–60. doi: [10.1016/j.molcel.2014.11.026](https://doi.org/10.1016/j.molcel.2014.11.026) Epub 2014 Dec 24. PMID: [25544563](https://pubmed.ncbi.nlm.nih.gov/25544563/)
11. Dyer MD, Neff C, Dufford M, Rivera CG, Shattuck D, Bassaganya-Riera J, et al. The human-bacterial pathogen protein interaction networks of *Bacillus anthracis*, *Francisella tularensis*, and *Yersinia pestis*. *PLoS One*. 2010 Aug 9; 5(8):e12089. doi: [10.1371/journal.pone.0012089](https://doi.org/10.1371/journal.pone.0012089) PMID: [20711500](https://pubmed.ncbi.nlm.nih.gov/20711500/)
12. Yang H, Ke Y, Wang J, Tan Y, Myeni SK, Li D, et al. Insight into bacterial virulence mechanisms against host immune response via the *Yersinia pestis*-human protein-protein interaction network. *Infect Immun*. 2011 Nov; 79(11):4413–24. doi: [10.1128/IAI.05622-11](https://doi.org/10.1128/IAI.05622-11) PMID: [21911467](https://pubmed.ncbi.nlm.nih.gov/21911467/)
13. Memisević V, Zavaljevski N, Pieper R, Rajagopala SV, Kwon K, Townsend K, et al. Novel Burkholderia mallei virulence factors linked to specific host-pathogen protein interactions. *Mol Cell Proteomics*. 2013 Nov; 12(11):3036–51. doi: [10.1074/mcp.M113.029041](https://doi.org/10.1074/mcp.M113.029041) PMID: [23800426](https://pubmed.ncbi.nlm.nih.gov/23800426/)
14. Blasche S, Arens S, Ceol A, Siszler G, Schmidt MA, Häuser R, et al. The EHEC-host interactome reveals novel targets for the translocated intimin receptor. *Sci Rep*. 2014 Dec 18; 4:7531. doi: [10.1038/srep07531](https://doi.org/10.1038/srep07531) PMID: [25519916](https://pubmed.ncbi.nlm.nih.gov/25519916/)
15. Fu W, Sanders-Beer BE, Katz KS, Maglott DR, Pruitt KD, Ptak RG. Human immunodeficiency virus type 1, human protein interaction database at NCBI. *Nucleic Acids Res*. 2009 Jan; 37(Database issue):D417–22. doi: [10.1093/nar/gkn708](https://doi.org/10.1093/nar/gkn708) PMID: [18927109](https://pubmed.ncbi.nlm.nih.gov/18927109/)
16. Calderone A, Licata L, Cesareni G. VirusMentha: a new resource for virus-host protein interactions. *Nucleic Acids Res*. 2015 Jan; 43(Database issue):D588–92. doi: [10.1093/nar/gku830](https://doi.org/10.1093/nar/gku830) PMID: [25217587](https://pubmed.ncbi.nlm.nih.gov/25217587/)
17. Kumar R, Nanduri B. HPIDB—a unified resource for host-pathogen interactions. *BMC Bioinformatics*. 2010 Oct 7; 11 Suppl 6:S16. doi: [10.1186/1471-2105-11-S6-S16](https://doi.org/10.1186/1471-2105-11-S6-S16) PMID: [20946599](https://pubmed.ncbi.nlm.nih.gov/20946599/)
18. Nourani E, Khunjush F, Durmuş S. Computational approaches for prediction of pathogen-host protein-protein interactions. *Front Microbiol*. 2015 Feb 24; 6:94. doi: [10.3389/fmicb.2015.00094](https://doi.org/10.3389/fmicb.2015.00094) PMID: [25759684](https://pubmed.ncbi.nlm.nih.gov/25759684/)
19. Rapanoel HA, Mazandu GK, Mulder NJ. Predicting and analyzing interactions between *Mycobacterium tuberculosis* and its human host. *PLoS One*. 2013 Jul 2; 8(7):e67472. doi: [10.1371/journal.pone.0067472](https://doi.org/10.1371/journal.pone.0067472) PMID: [23844013](https://pubmed.ncbi.nlm.nih.gov/23844013/)
20. Liu X, Huang Y, Liang J, Zhang S, Li Y, Wang J, Shen Y, Xu Z, Zhao Y. Computational prediction of protein interactions related to the invasion of erythrocytes by malarial parasites. *BMC Bioinformatics*. 2014 Nov 30; 15:393. doi: [10.1186/s12859-014-0393-z](https://doi.org/10.1186/s12859-014-0393-z) PMID: [25433733](https://pubmed.ncbi.nlm.nih.gov/25433733/)
21. Dyer MD, Murali TM, Sobral BW. Computational prediction of host-pathogen protein-protein interactions. *Bioinformatics*. 2007 Jul 1; 23(13):i159–66. PMID: [17646292](https://pubmed.ncbi.nlm.nih.gov/17646292/)
22. de Chasse B, Meyniel-Schicklin L, Aublin-Gex A, Navratil V, Chantier T, André P, Lotteu V. Structure homology and interaction redundancy for discovering virus-host protein interactions. *EMBO Rep*. 2013 Oct; 14(10):938–44. doi: [10.1038/embor.2013.130](https://doi.org/10.1038/embor.2013.130) PMID: [24008843](https://pubmed.ncbi.nlm.nih.gov/24008843/)
23. Doolittle JM1, Gomez SM. Mapping protein interactions between Dengue virus and its human and insect hosts. *PLoS Negl Trop Dis*. 2011 Feb 15; 5(2):e954. doi: [10.1371/journal.pntd.0000954](https://doi.org/10.1371/journal.pntd.0000954) PMID: [21358811](https://pubmed.ncbi.nlm.nih.gov/21358811/)
24. Mei S, Zhu H. A novel one-class SVM based negative data sampling method for reconstructing proteome-wide HTLV-human protein interaction networks. *Sci Rep*. 2015 Jan 26; 5:8034. doi: [10.1038/srep08034](https://doi.org/10.1038/srep08034) PMID: [25620466](https://pubmed.ncbi.nlm.nih.gov/25620466/)
25. World Health Organization. Global tuberculosis report 2013. 2013.
26. Koul A, Herget T, Klebl B, Ullrich A. Interplay between *mycobacteria* and host signalling pathways. *Nat Rev Microbiol*. 2004 Mar; 2(3):189–202. PMID: [15083155](https://pubmed.ncbi.nlm.nih.gov/15083155/)
27. Srinivasan L, Ahlbrand S, Briken V. Interaction of *Mycobacterium tuberculosis* with host cell death pathways. *Cold Spring Harb Perspect Med*. 2014 Jun 26; 4(8). pii: a022459. doi: [10.1101/cshperspect.a022459](https://doi.org/10.1101/cshperspect.a022459) PMID: [24968864](https://pubmed.ncbi.nlm.nih.gov/24968864/)
28. Cui T, He ZG. Improved understanding of pathogenesis from protein interactions in *Mycobacterium tuberculosis*. *Expert Rev Proteomics*. 2014 Dec; 11(6):745–55. doi: [10.1586/14789450.2014.971762](https://doi.org/10.1586/14789450.2014.971762) PMID: [25327725](https://pubmed.ncbi.nlm.nih.gov/25327725/)

29. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res.* 2000 Jan 1; 28(1):235–42. PMID: [10592235](#)
30. Robert CH, Janin J. A soft, mean-field potential derived from crystal contacts for predicting protein-protein interactions. *J Mol Biol.* 1998 Nov 13; 283(5):1037–47. PMID: [9799642](#)
31. UniProt Consortium. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* 2012 Jan; 40(Database issue):D71–5. doi: [10.1093/nar/gkr981](#) PMID: [22102590](#)
32. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009 Dec 15; 10:421. doi: [10.1186/1471-2105-10-421](#) PMID: [20003500](#)
33. Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res.* 2005 Apr 22; 33(7):2302–9. PMID: [15849316](#)
34. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, et al. Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc.* 2007; 2(10):2366–82. PMID: [17947979](#)
35. Mattow J, Schaible UE, Schmidt F, Hagens K, Siejak F, Brestrich G, et al. Comparative proteome analysis of culture supernatant proteins from virulent *Mycobacterium tuberculosis* H37Rv and attenuated *M. bovis* BCG Copenhagen. *Electrophoresis.* 2003 Oct; 24(19–20):3405–20. PMID: [14595687](#)
36. Målen H, Berven FS, Fladmark KE, Wiker HG. Comprehensive analysis of exported proteins from *Mycobacterium tuberculosis* H37Rv. *Proteomics.* 2007 May; 7(10):1702–18. PMID: [17443846](#)
37. de Souza GA, Leversen NA, Målen H, Wiker HG. Bacterial proteins with cleaved or uncleaved signal peptides of the general secretory pathway. *J Proteomics.* 2011 Dec 21; 75(2):502–10. doi: [10.1016/j.jprot.2011.08.016](#) PMID: [21920479](#)
38. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet.* 2004 Feb; 5(2):101–13. PMID: [14735121](#)
39. Sassetti CM, Rubin EJ. Genetic requirements for *mycobacterial* survival during infection. *Proc Natl Acad Sci U S A.* 2003 Oct 28; 100(22):12989–94. PMID: [14569030](#)
40. Rengarajan J, Bloom BR, Rubin EJ. Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. *Proc Natl Acad Sci U S A.* 2005 Jun 7; 102(23):8327–32. PMID: [15928073](#)
41. Stewart GR, Patel J, Robertson BD, Rae A, Young DB. *Mycobacterial* mutants with defective control of phagosomal acidification. *PLoS Pathog.* 2005 Nov; 1(3):269–78. PMID: [16322769](#)
42. Kumar D, Nath L, Kamal MA, Varshney A, Jain A, Singh S, et al. Genome-wide analysis of the host intracellular network that regulates survival of *Mycobacterium tuberculosis*. *Cell.* 2010 Mar 5; 140(5):731–43. doi: [10.1016/j.cell.2010.02.012](#) PMID: [20211141](#)
43. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000 May; 25(1):25–9. PMID: [10802651](#)
44. Chen ZJ. Ubiquitin signalling in the NF-kappaB pathway. *Nat Cell Biol.* 2005 Aug; 7(8):758–65. PMID: [16056267](#)
45. Ashida H, Kim M, Sasakawa C. Exploitation of the host ubiquitin system by human bacterial pathogens. *Nat Rev Microbiol.* 2014 Jun; 12(6):399–413. doi: [10.1038/nrmicro3259](#) PMID: [24801936](#)
46. Deretic V, Singh S, Master S, Harris J, Roberts E, Kyei G, et al. *Mycobacterium tuberculosis* inhibition of phagolysosome biogenesis and autophagy as a host defence mechanism. *Cell Microbiol.* 2006 May; 8(5):719–27. PMID: [16611222](#)
47. Russell DG. *Mycobacterium tuberculosis*: here today, and here tomorrow. *Nat Rev Mol Cell Biol.* 2001 Aug; 2(8):569–77. PMID: [11483990](#)
48. Diedrich CR, Flynn JL. HIV-1/*mycobacterium tuberculosis* coinfection immunology: how does HIV-1 exacerbate tuberculosis? *Infect Immun.* 2011 Apr; 79(4):1407–17. doi: [10.1128/IAI.01126-10](#) PMID: [21245275](#)
49. Pawlowski A, Jansson M, Sköld M, Rottenberg ME, Källenius G. Tuberculosis and HIV co-infection. *PLoS Pathog.* 2012 Feb; 8(2):e1002464. doi: [10.1371/journal.ppat.1002464](#) PMID: [22363214](#)
50. Rahman MM, McFadden G. Modulation of NF-κB signalling by microbial pathogens. *Nat Rev Microbiol.* 2011 Apr; 9(4):291–306. doi: [10.1038/nrmicro2539](#) PMID: [21383764](#)
51. Häcker H, Karin M. Regulation and function of IKK and IKK-related kinases. *Sci STKE.* 2006 Oct 17; 2006(357):re13. PMID: [17047224](#)
52. Kruh NA, Troutt J, Izzo A, Prenni J, Dobos KM. Portrait of a pathogen: the *Mycobacterium tuberculosis* proteome in vivo. *PLoS One.* 2010 Nov 11; 5(11):e13938. doi: [10.1371/journal.pone.0013938](#) PMID: [21085642](#)

53. MacGurn JA, Cox JS. A genetic screen for *Mycobacterium tuberculosis* mutants defective for phagosome maturation arrest identifies components of the ESX-1 secretion system. *Infect Immun*. 2007 Jun; 75(6):2668–78. PMID: [17353284](#)
54. Mehra A, Zahra A, Thompson V, Sirisaengtaksin N, Wells A, Porto M, et al. *Mycobacterium tuberculosis* type VII secreted effector EsxH targets host ESCRT to impair trafficking. *PLoS Pathog*. 2013 Oct; 9(10):e1003734. doi: [10.1371/journal.ppat.1003734](#) PMID: [24204276](#)
55. Philips JA, Porto MC, Wang H, Rubin EJ, Perrimon N. ESCRT factors restrict *mycobacterial* growth. *Proc Natl Acad Sci U S A*. 2008 Feb 26; 105(8):3070–5. doi: [10.1073/pnas.0707206105](#) PMID: [18287038](#)
56. Yang CS, Shin DM, Kim KH, Lee ZW, Lee CH, Park SG, Bae YS, Jo EK. NADPH oxidase 2 interaction with TLR2 is required for efficient innate immune responses to *mycobacteria* via cathelicidin expression. *J Immunol*. 2009 Mar 15; 182(6):3696–705. doi: [10.4049/jimmunol.0802217](#) PMID: [19265148](#)
57. Sun J, Singh V, Lau A, Stokes RW, Obregón-Henao A, Orme IM, et al. *Mycobacterium tuberculosis* nucleoside diphosphate kinase inactivates small GTPases leading to evasion of innate immunity. *PLoS Pathog*. 2013; 9(7):e1003499. doi: [10.1371/journal.ppat.1003499](#) PMID: [23874203](#)