

RESEARCH ARTICLE

Open Access



RepTB: a gene ontology based drug repurposing approach for tuberculosis

Anurag Passi^{1,2}, Neeraj Kumar Rajput¹, David J. Wild^{3*} and Anshu Bhardwaj^{1,2*}

Abstract

Tuberculosis (TB) is the world's leading infectious killer with 1.8 million deaths in 2015 as reported by WHO. It is therefore imperative that alternate routes of identification of novel anti-TB compounds are explored given the time and costs involved in new drug discovery process. Towards this, we have developed *RepTB*. This is a unique drug repurposing approach for TB that uses molecular function correlations among known drug-target pairs to predict novel drug-target interactions. In this study, we have created a Gene Ontology based network containing 26,404 edges, 6630 drug and 4083 target nodes. The network, enriched with molecular function ontology, was analyzed using Network Based Inference (NBI). The association scores computed from NBI are used to identify novel drug-target interactions. These interactions are further evaluated based on a combined evidence approach for identification of potential drug repurposing candidates. In this approach, targets which have no known variation in clinical isolates, no human homologs, and are essential for *Mtb*'s survival and or virulence are prioritized. We analyzed predicted DTIs to identify target pairs whose predicted drugs may have synergistic bactericidal effect. From the list of predicted DTIs from RepTB, four TB targets, namely, FolP1 (Dihydropteroate synthase), Tmk (Thymidylate kinase), Dut (Deoxyuridine 5'-triphosphate nucleotidohydrolase) and MenB (1,4-dihydroxy-2-naphthoyl-CoA synthase) may be selected for further validation. In addition, we observed that in some cases there is significant chemical structure similarity between predicted and reported drugs of prioritized targets, lending credence to our approach. We also report new chemical space for prioritized targets that may be tested further. We believe that with increasing drug-target interaction dataset RepTB will be able to offer better predictive value and is amenable for identification of drug-repurposing candidates for other disease indications too.

Keywords: Tuberculosis, RepTB, Repurposing, Network based inference, Polypharmacology, MDR/XDR, Interaction

Background

Emergence of Multi-Drug Resistant (MDR) and Extensively-Drug Resistant (XDR) Tuberculosis is impeding the progress of combating the epidemic of Tuberculosis [1]. According to WHO, in 2015 1.8 million people died of TB and almost 480,000 diagnosed with MDR-TB globally [2]. It takes up to 2 years to cure MDR/XDR patients and more than 50% patients do not respond to the existing treatment regimens [3]. Additionally, the existing

drugs in the TB regimen are toxic (Fluoroquinolones & Aminoglycosides show hepatotoxicity and renal toxicity, respectively) that deters compliance and leads to poor-treatment outcomes [4, 5]. Given that only 10% of the compounds go through from Phase I to final FDA approval [6] and high attrition rates of lead molecules passing from preclinical development to Phase I clinical studies [7], alternative strategies are needed. Drug repurposing is an attractive strategy to identify novel treatment options given that it may reduce R&D timelines by 3–5 years, have less development costs and the improved quality of success [8, 9]. The concept of drug repurposing is not new and many new drug indications have been identified serendipitously. The classic cases include Sildenafil that was initially approved for angina but repurposed for erectile dysfunction and Canakinumab which

*Correspondence: davidjwild@indiana.edu; anshu@imtech.res.in

² Academy of Scientific and Innovative Research, Council of Scientific and Industrial Research, Training and Development Complex, CSIR Campus, CSIR Road, Taramani, Chennai, Tamil Nadu 600113, India

³ School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN 47405, USA

Full list of author information is available at the end of the article

was initially tested for Rheumatoid arthritis but was repurposed to treat cryopyrin-associated periodic syndrome (CAPS) [10].

Palomino et al. [11] discussed the viability of drug repurposing methods in treatment of infectious diseases like Tuberculosis. Many of the drugs like Fluoroquinolones (Gatifloxacin and Moxifloxacin); rifamycins (Rifampentine and Clofazimine); oxazolinones (Linezolid and Sutezolid); and beta lactams (Meropenem and Clavulanate) have been repurposed against MDR/XDR TB [12]. Linezolid and Metronidazole have completed the Phase II while Gatifloxacin have completed phase III clinical trials [13]. More recently, cephalosporins in combination with Rifampicin and other anti-TB drugs like Bedaquiline and Delamanid have also shown promising synergistic activity [14]. Besides the anti-infectives, non-anti-infective agents have also been repurposed against TB. These include: Entacapone and Tolcapone, that act as adjunct to treatment of Parkinson's disease [15]; Thioridazine and chlorpromazine are drugs used in treatment of psychoses; and NSAIDs such as Diclofenac, Ibuprofen, and Carprofen also have anti-TB activity [16].

Efforts towards identification of new TB drugs followed either phenotypic screens or target based screens [17]. The first approach, relies on development of whole-cell screening assays and availability of library of chemical compounds against replicating and non-growing *Mycobacterium tuberculosis* (Mtb). The target based screening approach involves purification of target protein, further compound screening assays against the target and in vivo target validation. As opposed to phenotypic screening, the target-based drug discovery has been less successful with few examples like identification of drug-like inhibitors of EthR and malate synthase. In order to systematically explore the drug target space in TB, several computational methods have also been developed which have identified potential drug repurposing candidates for TB. Brindha et al. [18, 19] identified drug candidates against Mtb MurE and PknB using docking based virtual screening method. In addition, methods based on structural proteome of Mtb by Kinnings et al. [20] and identification of polypharmacological drugs by Ramakrishnan et al. [21] have also reported repurposing candidates against Mtb proteins. However, these methods of drug repurposing are limited by the availability of the 3D structure of the target and/or ligand. Consequently, network pharmacology based methods have been adopted to predict drug target interactions (DTIs), for example, pharmaco-chemical-genomics; Random Walk; gene expression and network based analysis, etc. [22–24].

More recently, Cheng et al. [25, 26] used Network Based Inference method to predict drug-target interactions but has not been applied to identify repurposing candidates for TB.

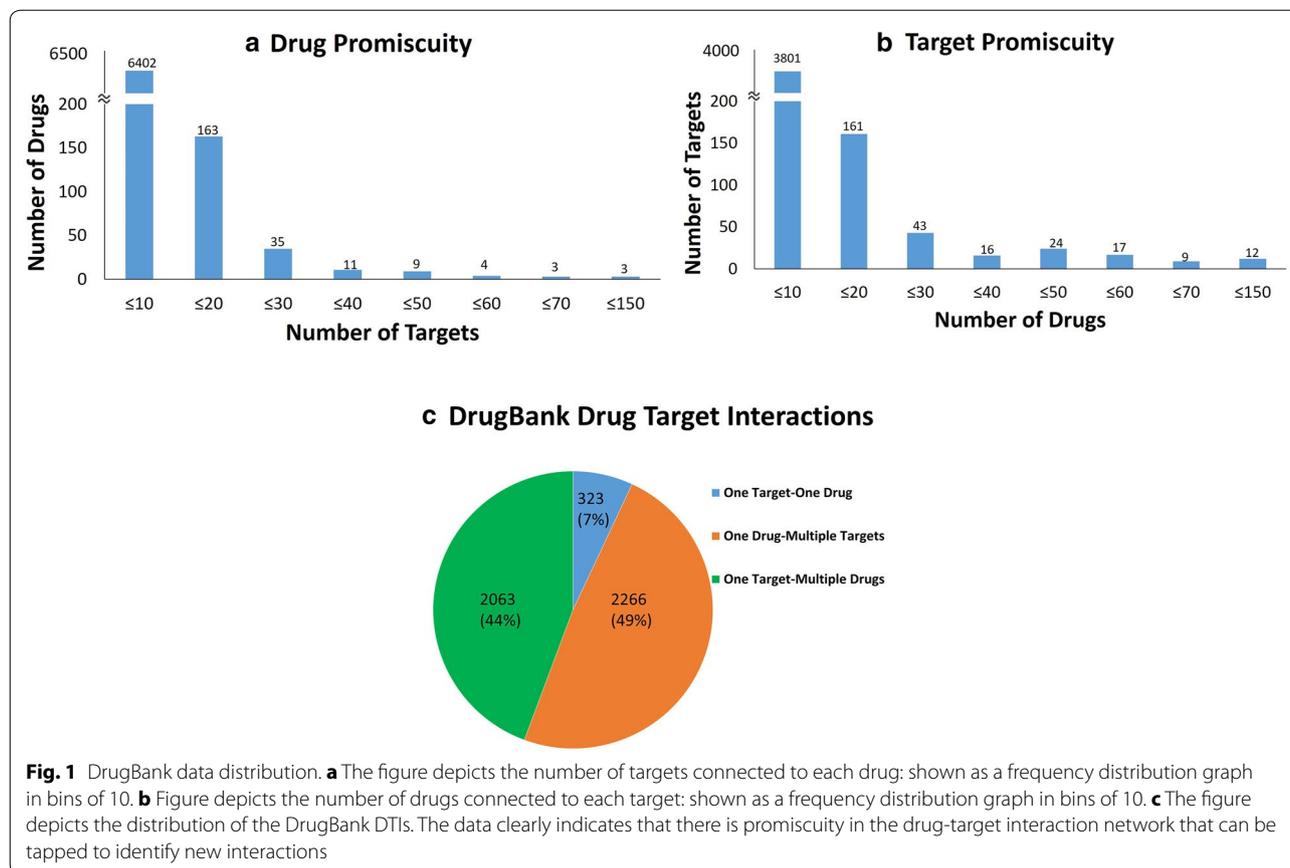
In this study, we have developed a drug-repurposing platform, *RepTB*, using network pharmacology approach to identify potential repurposed candidates for TB. RepTB utilizes Gene Ontology (GO) based drug-target interaction (DTIs) network to compute association scores for identification of new DTIs. Once the DTIs are predicted, a combined evidence based approach is applied to shortlist the potential candidates. Furthermore, to assess the chemical diversity, a near neighbor analysis was performed using the chemical structure similarity index between the known and predicted chemical space. In the end, we propose synergistic DTIs that may be evaluated further as potential starting points in TB drug discovery pipeline.

Results

1. Promiscuity in drug-target interactions from DrugBank

A drug is promiscuous when it acts on multiple targets and exhibits distinct pharmacological effects [27]. To find out the level of promiscuity in the known drugs, we plotted the frequency of drugs against targets (Fig. 1). As seen in Fig. 1a, of the 6630 drugs from DrugBank [28–31], 6402 drugs bind to fewer than 10 targets of which 4364 drugs interact with only one target. The remaining 228 drugs interact with more than 11 targets, some of which are known to interact with wide range of targets (DB02379: 90 targets and DB00157: 144 targets). Likewise, of the total 4083 targets reported in DrugBank, 2020 targets bind to single drug (Fig. 1b). The rest, bind to more than one drug, for example, CDK2 protein (Uniprot ID: P24941) interacts with 114 drugs and Prothrombin (Uniprot ID: P00734) interacts with 101 drugs. In Fig. 1c, the DrugBank drug-target interaction network, 49% drugs and 44% targets show promiscuity. Of the total interactions, only 7% exhibit one drug-one target interaction (singletons).

As discussed above, the promiscuity in the drug-target interaction space may be harnessed to predict novel interactions for drug repurposing. For the same, a unique pipeline, RepTB is established which begins with creation of the Gene Ontology based network followed by a combined evidence based strategy to predict drug repurposing candidates. The pipeline is depicted in Fig. 2. The components of this platform are now discussed.



2. Gene ontology based enriched drug-target interaction network

The DrugBank DT network comprised of 15,824 edges. An additional set of 12,853 edges were added to the network based on molecular function Gene Ontology (GO) based mapping of the targets. On combining the DrugBank DTIs and GO mapping DTIs, the *RepTB* network contained 26,404 edges with 6630 drug and 4083 target nodes. The combined DTI edge list is available as Additional file 1.

3. Identification of new drug-target interactions using NBI

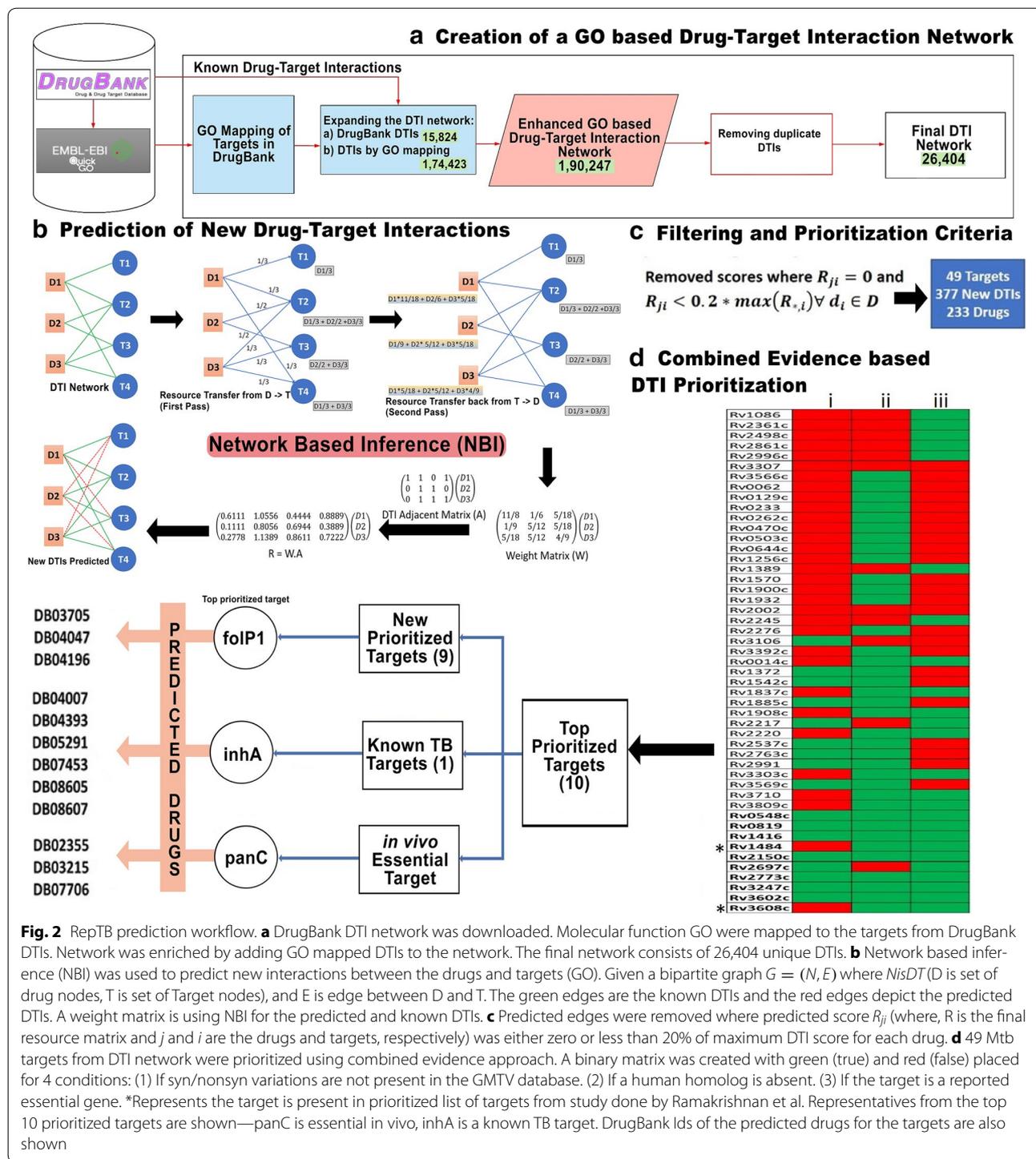
For the entire network comprising of 6630 drugs and 4083 targets, NBI association scores were computed for approximately 27 million (27,070,290) interactions. Of these, 26,404 are known and 27,043,886 are new associations (unconnected drug-target pairs). In order to create a prioritization list, we applied a metric where a target was only prioritized for a drug if its association score was more than 20% of the maximum score in the sorted list with unconnected targets. This led to identification

of 25,323 potential DTIs which are prioritized based on their NBI association scores.

4. Target prioritization from combined evidence approach

In the prioritized dataset of 25,323 DTIs, there are 49 *Mtb* targets interacting with 233 drugs (Additional file 2). In order to further prioritize targets from the potential DTIs, a combined evidence approach was implemented with three major criteria—role in drug resistance, absence of a human homolog and essentiality for survival or virulence.

Based on these criteria, of the 49 targets, 10 targets were prioritized. Of these 10, seven targets qualify for all three parameters. The remaining three were included in the prioritized list based on various factors. For example, targets, namely, FolP1 (Rv3608c) and InhA (Rv1484), were also prioritized by another group [21]. *Mtb* dUT-PASE (Dut, Rv2697) is also prioritized despite that it shares a 39% sequence identity with the human dUT-Pase. One might argue that Dut might not be a viable target after all. However, according to the studies by Chan et al. [34] structural differences in the active site between



Human and Mtb Dut enables development of inhibitors specific to Mtb. We do stress that it is imperative to understand the key structural features of various proteins to consider them as viable targets.

5. Repurposed drug candidates for TB

As depicted in Table 1, we predicted 57 potential drug candidates for the 10 prioritized targets. These targets belong to metabolic pathways that are essential for survival of the bacteria such as folate biosynthesis, pyrimidine biosynthesis, cell division, mycothiol biosynthesis, menaquinone biosynthesis, mycolic acid biosynthesis, lysine and riboflavin biosynthesis, and pentothenate biosynthesis. Although many of the predicted drugs are antibacterial, some drugs have reported indications as antiviral, antifungal, anti-cancer and treatment of bipolar disorder. Moreover, some nutraceuticals and obesity drugs are also predicted as potential candidates for TB targets.

In order to understand the chemical diversity of the predicted drugs in context of the known drugs of the prioritized targets, chemical structural similarity analysis was performed. At a dissimilarity coefficient of 0.15, while known and predicted drugs for FolP1, MenB, FtsZ and Tmk were observed to be structurally similar (Fig. 3a–d), no significant structural similarity was observed between the known and predicted drugs for Dut, InhA, RibH and MshD (Fig. 3e–h).

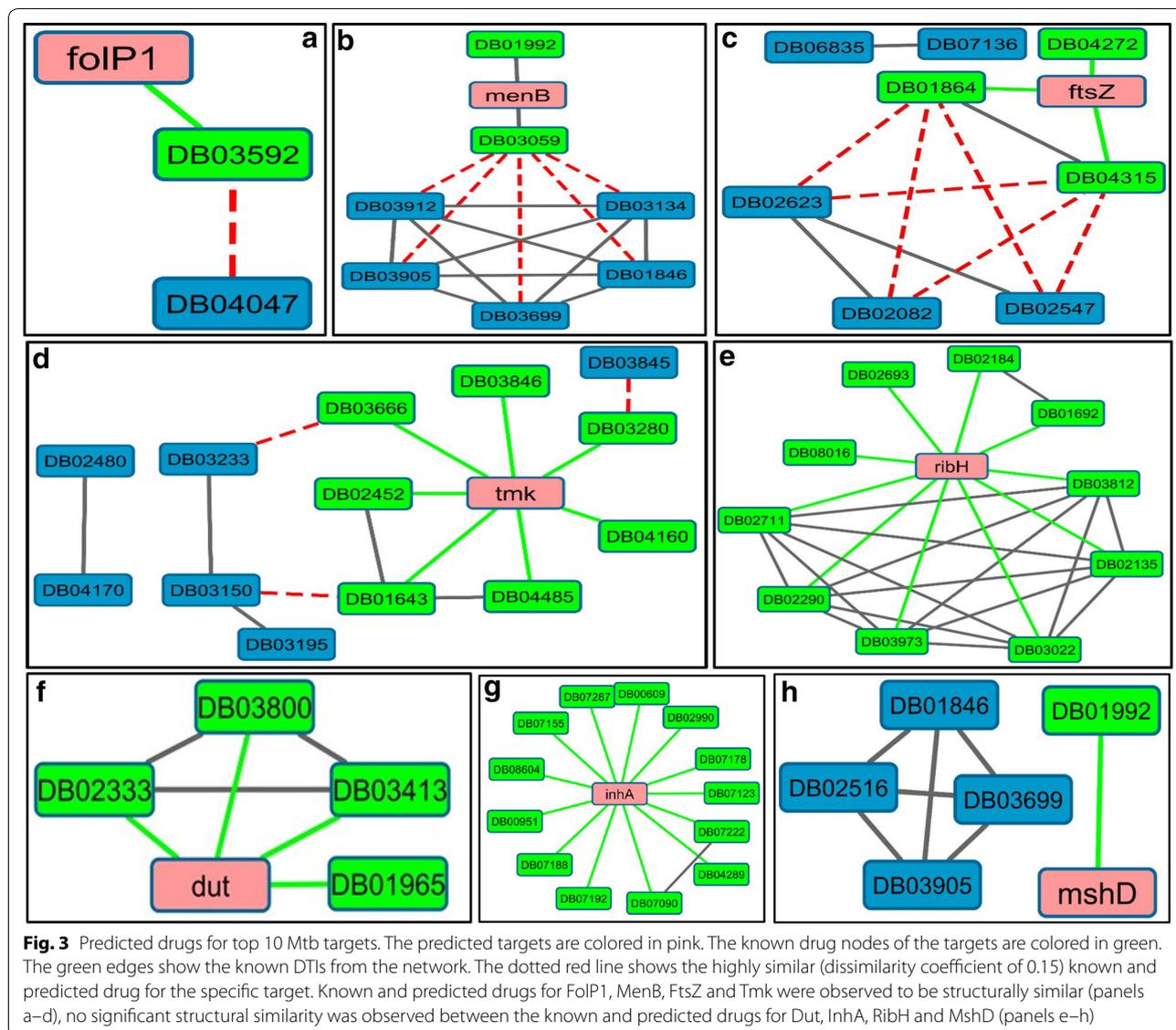
Discussion

RepTB is a systems based platform to identify repurposing candidates for TB. In this work, we have used a data driven approach towards prioritizing TB targets and identifying existing drugs that may act on Mtb. We use the concept of resource allocation between the connected nodes (reported drug-target pairs) to find association scores predicting the degree of association between previously unconnected drugs and targets. We use this concept to identify if a drug not known to be associated with a TB target can be repurposed for TB using NBI scores. Based on the combined evidence based approach, 10 Mtb targets were prioritized. These genes are deemed essential for Mtb and do not have human homolog. Recently, there has been a keen interest in identifying combinations of non-lethal genes whose incomplete inhibition can produce a lethal phenotype [35].

Rv3608c (FolP1, encoded as DHPS) is part of the folate biosynthesis in which the reduced folate species participate in one-carbon metabolism that produce biochemically important biomolecules such as purines, thymidine, methionine, serine, and N-formylmethionyl-tRNA. The dTMP formed as a by-product of folate biosynthesis

feeds into the DNA synthesis. DHPS converts 2-Amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine- P_2 to 7,8-Dihydro-pterolate which in turn converts to 7,8-Dihydrofolate (DHF) that gets converted into folate by *dfrA*. Mtb is unable to acquire folate from environment and is dependent on de-novo synthesis [36]. Dapsone and other sulfonamides have been known to inhibit folP1 by exhibiting bacteriostatic effect [37, 38]. Sulfa drugs target the folP encoded protein DHPS. Sulfa drugs are structural analogs pABA and act as competitive antagonists of DHPS. In addition, these drugs can be alternate substrates of DHPS and form sulfa-pterolates that cannot be further converted to folate that leads to a dead end in the pathway. Recently, mutations in the sulfa drug binding pockets of DHPS have caused resistance towards these drugs that has led to identification of alternate binding site in DHPS [39, 40]. The pterin binding pocket within the DHPS has a high degree of conservation and no sulfa drug resistance has been reported [41]. Hence, there have been studies designing compounds for inhibiting DHPS by binding to pterin binding pocket [42]. Three drugs were predicted for FolP1. DB04196, DB03705, and DB04047 are known to target *Bacillus anthracis* FolP protein and DB04047 is also known to target *B. anthracis* FolP and *E. coli* FolK proteins. The molecular function GO terms of *folP* and *folP1* genes are same (GO:0004156) even though the term assigned to *B. anthracis folP* gene is by IEA (Inferred from Electronic Annotation). FolK differs from FolP1 by transferring phosphorus-c moiety rather than alkyl or aryl group. These drugs can be repurposed against Mtb. The three predicted drugs for FolP1 are antibacterial drugs DB03705 (6-Methylamino-5-Nitrosocytosine), DB04196 (Pteric Acid), and DB04047 ([Pterin-6-Yl Methanyl]-Phosphonophosphate). DB03705 and DB04196 interact with single targets in the DTI network while DB04047 interacts with only 2 targets in the network indicating low promiscuity in the DTI network suggesting a high specificity for the target. It has been reported that *folP2* gene blocks sulfamethoxazole (SMX), which inhibits FolP1. This causes resistance against the drug. It was shown that combining FolP2 and FolP1 inhibitors can overcome the resistance in Mtb [38]. Similarly, in our study, the prioritized target FolP1 is similar to the known targets of its predicted drugs and those drugs do not act via binding to the sulfa binding sites. Hence, show no resistance.

Rv3247c (Tmk) and Rv2697c (Dut) are involved in Pyrimidine biosynthesis. Thymidylate kinase (Tmk) of Mtb is essential for DNA synthesis by converting dTMP to dTDP. This enzyme lies in the junction of de novo and salvage pathways for biosynthesis of dTTP [43]. It also utilizes dUMP as substrate however with low affinity than dTMP [44]. Out of 11 predicted drugs for Tmk, 7



are gram-negative antibacterial drugs namely, DB01799 (4-Hydroxy-3-Methyl Butyl Diphosphate), DB02480 ((S)-4-bromo-3-hydroxy-3-methylbutyl diphosphate), DB03165 (2-Dimethylamino-Ethyl-Diphosphate), DB03233 (Phosphoric Acid Mono-[3-Amino-5-(5-Methyl-2,4-Dioxo-3,4-Dihydro-2h-Pyrimidin-1-Yl)-Tetrahydro-Furan-2-Ylmethyl] Ester), DB04170 (4-bromo-3-hydroxy-3-methyl butyl diphosphate), DB02594 (2'-Deoxycytidine), DB03723 (2'-Deoxy-Thymidine-Beta-L-Rhamnose). One is an antiviral drug DB03150 (2',3'-Dideoxythymidine-5'-Monophosphate), one is a bipolar disorder drug DB02745 (Uridine) and two drugs, DB03195 (Phosphoric Acid Mono-[3-Fluoro-5-(5-Methyl-2,4-Dioxo-3,4-Dihydro-2h-Pyrimidin-1-Yl)-Tetrahydro-Furan-2-Ylmethyl] Ester), and DB03845

(P1-(5'-Adenosyl)P5-(5'-(3'-azido-3'-Deoxythymidyl) Pentaphosphate) are indicated as anti-cancer. DB02594 is a known human TK2 and deoxycytidine kinase inhibitor, and *E. coli* class B acid phosphatase. DB03233, DB03150, DB03195, DB03845 are already known to interact with human Tmk (P23919) proteins. All these proteins share the same GO annotation thymidylate kinase activity (GO:0004798). DB01799, DB02480, DB03150, DB03165, DB03195, DB03233, DB03845, and DB04170 interact with single targets while DB02594 interacts with 3 targets, DB02745 and DB03723 interacts with only 2 targets each. Another predicted target, dUTPase (Dut) is required for the optimal growth of Mtb [45] plays an important role in the de novo and salvage biosynthesis of dTTP [46]. It maintains the dUTP/dTTP levels low

by cleaving the dUTP to form dUMP and consequently preventing the incorporation of uracil into DNA. In addition, Dut is also responsible for production of dUMP is a precursor for dTTP formation for DNA synthesis. If the *dut* gene is non-functional, the uracil content goes up causing its incorporation into the DNA and consequently cell death [46]. Moreover, dUTPase knockouts are shown to be lethal in *E. coli* and *S. cerevisiae* [47, 48]. Hence, the prioritized Dut can be a viable target for TB. Predicted drug for Dut, DB04685 is an antiparasitic drug interacts with only 2 targets, a human Dut protein and a *P. falciparum* Deoxyuridine 5'-triphosphate nucleotidohydrolase. Both proteins have the same molecular function GO term as that of the prioritized Mtb target for dUTP diphosphatase activity (GO:0004170). According to the work done by Kaur et al., it is possible to predict the vulnerability of a target in silico and to assess the bactericidal based on the NAD/NADH redox ratio [49]. Knockout of some of the genes can lead to increase in this ratio that differentiates whether the mutation is bacteriostatic or bactericidal. The ratio under normal growth is maintained within 0.5–2. If the ratio increases above 5, the cidality increases. Both *tmk* and *dut* were assessed to be bacteriostatic genes [50] and since these genes belong to the same pyrimidine biosynthesis pathway, have no human homologs and are essential for the survival of Mtb, can be targeted together to produce a synergistic bactericidal effect.

Rv0548c (MenB) is part of menaquinone biosynthesis which is essential in the electron transport chain and oxidative phosphorylation for the survival of Mtb. Menaquinone is synthesized by *menA-G* genes. MenB converts O-succinylbenzoyl-CoA to 1,4-dihydroxy-2-naphthoic acid (DHNA). Menaquinone has to be made continuously to maintain the membrane stability during growth [51, 52]. Since Mtb does not have ubiquinone like other bacterial species, it is dependent on menaquinone for its ATP production. Inhibiting MenB will prevent transfer of electrons in ETC and consequently inhibit the ATP production. Thus, a cascading effect will also decrease the activity of various other ATP-dependent efflux pumps. Our analysis predicted 13 drugs for MenB. Seven of these drugs—DB01669, DB01764, DB02516, DB03230, DB01846, DB02039, and DB3912 are antibacterials. DB01783 (Pantothenic acid), is used in treatment of testicular torsion, diabetic ulceration, wound healing, acne, obesity, diabetic peripheral polyneuropathy and DB03612 (3-Hydroxybutyryl-Coenzyme A) is an inhibitor of mitochondrial beta-oxidation. DB03699, DB01856, DB03134 and DB03905 drugs have been reported to target an unknown prokaryotic organism. However, the GO terms for the targets do not match with the predicted target. Since these drugs show low promiscuity, targeting MenB

can therefore be a viable method of disrupting the formation of menaquinone and consequentially hindering bacterial growth [52].

Rv0819 (MshD) is an acetyltransferase and is a part of the mycothiol biosynthesis. 11 drugs were predicted for MshD. DB01669, DB01764 (Dalfopristin), DB03230, DB03912, DB01846, DB02516 are all antibacterials. DB01783 has been investigated for its hypolipidemic effects and as cholesterol lowering agent. Quinupristin-Dalfopristin is a known inhibitor of *Enterococcus faecalis* and is bacteriostatic against the bacteria. Also, when treated against *Mycobacterium marinum* was less active than clarithromycin [53]. However, when given in combination with other drugs like doxycycline and ampicillin, the synergistic effect enhances the bactericidal activity [54]. DB03134, DB03699, DB03905 have been known to target unknown prokaryotic organisms.

Rv2150c (FtsZ) is part of the cell division process and standard anti-tubercular drugs are shown to be bacteriostatic towards FtsZ [49]. With the exception of DB08185, which is interacting with 27 different targets in the DTI, all the 17 predicted drugs for FtsZ have a very low promiscuity. Known targets for DB04723, Db06835, DB06921, DB07136, DB07157, DB07182, DB07269 share the molecular function GO term with FtsZ up to the hydrolase activity after which FtsZ branches out to further annotation downstream. Moreover, known targets of DB02082 and DB02623 share the same depth of GO annotation as FtsZ suggesting that the predicted drugs could be explored for repurposing against FtsZ.

3,4-Dihydroxy-2-butanone 4-phosphate of pentose phosphate pathway is converted to 6,7-Dimethyl-8-ribyllumazine by Rv1416 (RibH). Since Mtb is not capable of taking riboflavin from the environment and riboflavin biosynthesis is absent in humans, RibH makes a great target for anti-tb drug therapy. Predicted drug DB02452 is promiscuous and interacts with 10 targets in the DTI. However, according to the GO annotations, the known targets of DB02452 shared only single depth of the GO annotation as having a catalytic activity. Rv3602c (PanC) is part of the (R)-Pantothenate biosynthesis. A gene silencing study performed on *panC* reported that silencing *panC* in vitro was bacteriostatic in nature [55]. Predicted drugs for PanC, DB03255, DB03215, and DB07706, are antibacterial drugs and interact with only one target each hinting of high specificity. Rv2773c (DapB) is a part of lysine biosynthesis that feeds into peptidoglycan biosynthesis. Predicted drug for DapB, DB03969 is known to interact with *E. coli* DapB and can be repurposed for Mtb DapB.

Recently, there has been a renewed interest in Rv1484 (InhA) which is already a known anti-tb target. Mtb has a single copy of enoyl reductase making *inhA* an essential

gene. However, known anti-tb drug, isoniazid is an inhibitor of InhA that acts via an indirect mechanism of action as a prodrug. It needs to be first activated by KatG. Since, there have been known mutations in KatG that cause resistance towards isoniazid, new direct inhibitors of InhA have recently been reported [56]. Predicted drugs for InhA in our study, DB04007, DB04393, DB08605 are antibacterials, DB05291 is an anesthetic, DB07453 and DB08607 have unknown indications. Known targets for DB07452, DB04047 and DB08608 share some functional GO based association with InhA. All these targets show oxidoreductase activity.

Polypharmacology is a preferred strategy to target more than one protein in Mtb. However, it is also important that there are no anti-targets or off-targets of the predicted drugs in the host. In the combined evidence approach, targets that have similarity with the host proteome are systematically eliminated from the prioritized list. So in essence, drugs with multiple targets in Mtb but no or few targets in host may be prioritized for further testing.

We identified similarity between the known and predicted drugs of the targets. FolP1, Tmk, FtsZ and MenB predicted drugs showed over 85% structural similarity (Tanimoto similarity) with their known drugs. The predicted drugs, were subjected to FAF drug filter [57] to predict their ADME/tox properties. We used the XLOGP3 method as the logP computation program with following parameters as true: PPIHitProfiler, Filter undesirable substructures moieties, Filter Pan Assay Interference Compounds (PAINS) Filter (A, B, C), Lilly MedChem Rules (regular). Out of the 56 predicted drugs, 4 drugs, namely, DB01764, DB07453, DB04196 (Pteric Acid) and DB03969 passed all the FAF filters. Pteric acid interacts with only one target in the DTI suggesting that the drug is not promiscuous. The drugs passed the FAF filter with desirable solubility and bioavailability suggesting that these compounds can be repurposed for the Mtb. The predicted compounds therefore, offer a new set of diverse chemical space that can be tested further for their anti-tb activity.

Based on the analysis done so far, it was observed that some targets of the predicted drugs fall in metabolic pathways that have common precursors and may be inhibited simultaneously for synergistic effect as shown in Fig. 4.

The structure based drug discovery process is based on the availability of three-dimensional structures of the target protein. Although there are 2052 Mtb structures in PDB, only ~700 representative structures for Mtb proteins are present. Therefore, alternate methods are needed to complement the structure-based prioritization of drug target space. For the predicted DTIs, we

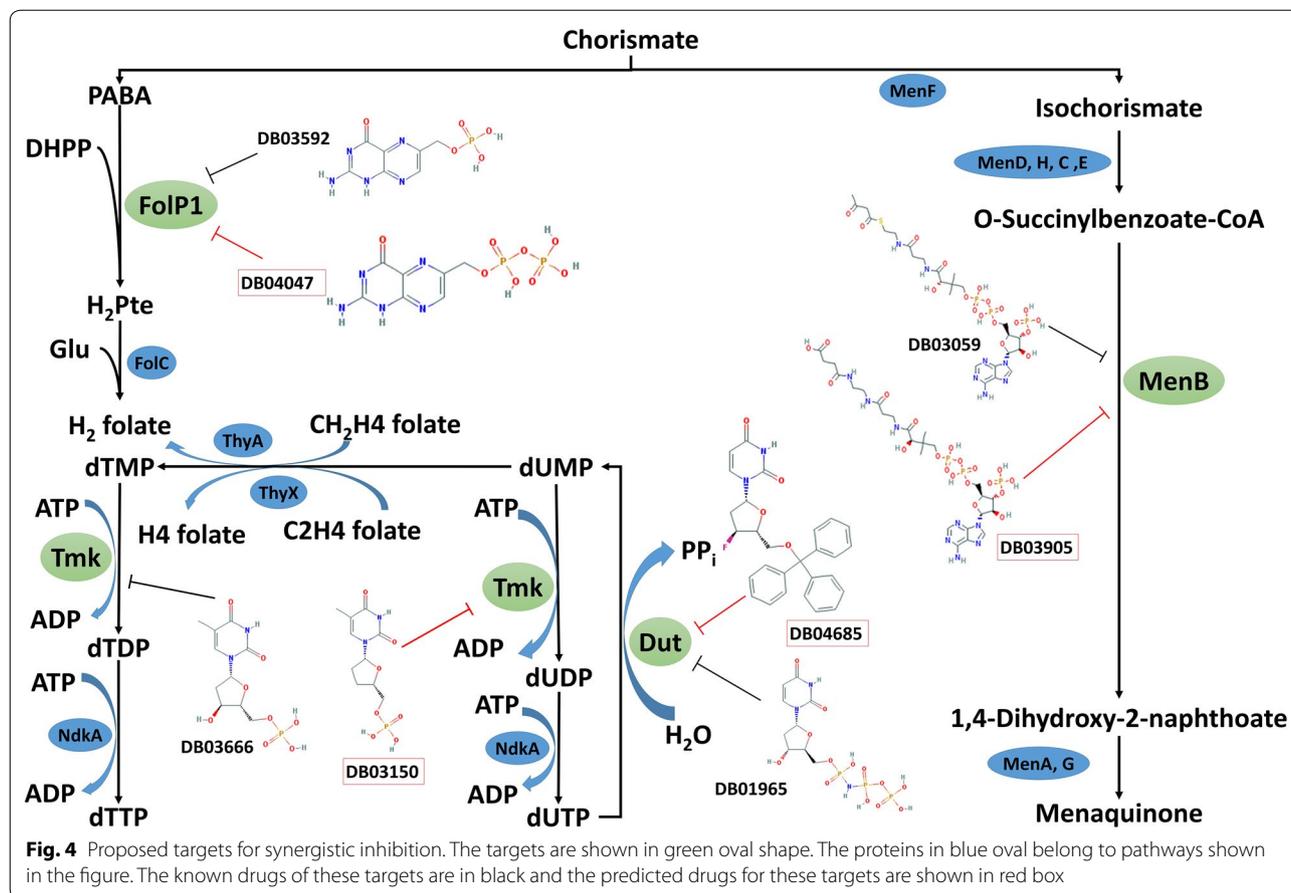
compared our results with those of other groups [20, 21] which are primarily based on the extent of structural information available for the Mtb proteins. Although there were no common predicted DTIs among the studies, there were common predicted targets such as InhA, PanC, MshD, FolP1, and MenB. Our work, therefore, complements published methods by providing new DTIs based on the molecular function associations between a drug and its known target. We have tried to incorporate GO information to the highest level of annotation for each target. Since, most proteins have broader GO terms, the functional aspect of DTI network is still a work in progress and as the GO terms for the targets become more descriptive, the GO mappings will be more precise leading to better predictions. We believe that RepTB may complement the efforts towards alternate process of drug discovery by predicting hidden drug-target associations.

Conclusion

It is an established fact that TB is a global pandemic. There is an increasing demand for not only new drugs but also for innovative drug discovery strategies. In this work we created RepTB, a GO based DTI network to identify potential repurposed candidates for TB. RepTB uses combined evidence approach to prioritize Mtb targets. In addition, we also propose new chemical space that may be taken further for validation. Our method relies heavily on the availability of well-defined GO annotations and we believe that our method will be more accurate with improved annotations of the Mtb proteome. We, therefore, conclude that we have predicted diverse compounds that can potentially act as anti-tb. To the best of our knowledge, this is the first recommendation system to predict DTIs for TB through NBI and can also be applied for any other infectious agent.

Methods

1. Data acquisition and network generation
Known drug target interactions were taken from DrugBank [28–31] (version 4 downloaded in October 2015 and updated in February 2016). The original list contained 6630 unique drugs and 4083 unique targets. The total DrugBank DTIs were 15824. The data files, codes and associated documents are available through the GitHub link: <http://ab-openlab.csir.res.in/gitlab/openlab/reptb/tree/master>
2. Gene Ontology (GO) mapping for network enrichment
We downloaded the molecular function GO mappings from QuickGO [32, 33] database of EBI. The data was filtered based on evidence codes mentioned in Additional file 3. Evidence codes IEA (Inference from Electronic Annotation), NAS (Non-traceable



Author Statement) and ND (No Biological Data) were excluded as these three evidence codes either are not curated or do not have supporting reference. We mapped the leaf nodes of the GO annotations to DrugBank targets. If two targets shared the same GO code, a new DTI edge was created (<http://ab-openlab.csir.res.in/gitlab/openlab/reptb/tree/master>).

3. Drug-Target Network Projection using Network-Based Inference

A drug-target interaction can be represented as a bipartite graph $G(D, T, E)$, where drug set $D = \{d_1, d_2, \dots, d_m\}$, target set $T = \{t_1, t_2, \dots, t_n\}$ and $E = e_{ij} : t_i T, d_j D$. An edge is formed if a drug is associated with a target. This graph can also be represented as an $n \times m$ adjacent matrix $\{a_{ij}\}$, where $a_{ij} = 1$, if t_i and d_j are linked and $a_{ij} = 0$ if t_i and d_j are not linked.

We denoted $f_0(o) = a_{io}, o \in \{1, 2, \dots, m\}$ as the initial resource of drug d_o , for a target t_i , and $f(j)$ as the final resource of drug d_j . The final resource allocation can be depicted in a matrix form as $f' = W f_0'$, where f_0' is the column vec-

tor of f_0 and W is a weight matrix depicted as

$$W = \{w_{pq}\}_{m \times m} = \left\{ \frac{1}{k(tq)} \sum_{l=1}^n \frac{a_{pl} a_{ql}}{k(d_l)} \right\}_{m \times m}, \text{ where } k(d_l) = \sum_{s=1}^m a_{sl}$$

represents the number of targets that interact with drug d_l .

4. Drug Target prioritization using recommendation system

Predicted targets for each drug were prioritized by sorting in descending order of the NBI scores. DTIs with prediction scores of zero were removed and targets which had scores greater than 20% difference from the top score were taken as prioritized drug targets.

5. Prioritization of Mtb targets through combined evidence approach

A target prioritization matrix was created using data from four different research studies to further prioritize Mtb targets. If a target from our study is also prioritized in any of the other studies, a "yes" is marked in that cell, otherwise a "no" is marked. The matrix also takes into account the essentiality of the gene, whether Nonsyn/Syn changes observed in clinical isolates, and whether a human homolog is

present or not for that target. We performed a blastp sequence similarity search between the Mtb targets and the human proteome. The search parameters were based on e-value $1E-4$, % identity of <math><35\%</math> and query coverage of $\geq 60\%$. We also matched our set of prioritized targets with a study done by Ramakrishnan et al. [21] in which they predict DTIs by using sequence and structural analysis for understanding the evolutionary relationships between Mtb proteins and FDA approved drugs.

6. Near neighbor analysis to find chemically similar compounds

We used ChemAxon's JChem (Version 16.5.30.0) to find structurally similar compounds between known and predicted drugs for each of the target. Structural fingerprint for each of the drugs was calculated using the `generatemd` command from JChem as mentioned below:

```
generatemd c<input_smiles_file> -g -k CF -f 1024 -D -o <output_filename>
```

Using this fingerprint, a tanimoto similarity coefficient was calculated to assess the structurally similar drugs using the `nneib` command:

```
nneib -Xmx819 -f 1024 -t 0.3 -g <input_descriptor_file> >> output_near_neighbor_file -v
```

Cytoscape v3.5.1 [58] was used to represent the chemical space profiling for the above analyses.

Additional files

Additional file 1. Combined DTI edge list.

Additional file 2. 49 Mtb targets with known and predicted DTIs along with components of combined evidence approach.

Additional file 3. GO Evidence Codes.

Authors' contributions

AP, DJW and AB conceived the project. AP compiled and curated the data. AP, NKR and AB carried out the data analysis. AP and AB prepared the manuscript. All authors read and approved the final manuscript.

Author details

¹ Bioinformatics Centre, Institute of Microbial Technology, Council of Scientific and Industrial Research, Chandigarh 160036, India. ² Academy of Scientific and Innovative Research, Council of Scientific and Industrial Research, Training and Development Complex, CSIR Campus, CSIR Road, Taramani, Chennai, Tamil Nadu 600113, India. ³ School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN 47405, USA.

Acknowledgements

We acknowledge Rakesh Kumar for setting up the GitLab repository for RepTB. We would like to acknowledge Dr. Abhik Seal, Senior Scientist, AbbVie Inc.,

USA for his critical inputs in the study. We would like to acknowledge the support of CSIR-SRF and USIEF-Fulbright-Nehru Doctoral Research (FNDR) Fellowship for providing the funding to complete this work. We would also like to acknowledge the support of the OSDD Community.

Competing interests

The authors declare no competing financial interests.

Availability of data and materials

All data and material associated with the manuscript is provided as supplementary files.

Ethics approval and consent to participate

Not applicable.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 16 October 2017 Accepted: 23 April 2018

Published online: 21 May 2018

References

1. Manjelienskaia J, Erck D, Piracha S, Schragger L (2016) Drug-resistant TB: deadly, costly and in need of a vaccine. *Trans R Soc Trop Med Hyg* 110(3):186–191
2. WHO (2016) Global tuberculosis report 2016. World Health Organization, Geneva
3. Kendall EA, Fojo AT, Dowdy DW (2016) Expected effects of adopting a 9 month regimen for multidrug-resistant tuberculosis: a population modeling analysis. *Lancet Respir Med* 5(3):191–199
4. Ramappa V, Aithal GP (2013) Hepatotoxicity related to anti-tuberculosis drugs: mechanisms and management. *J Clin Exp Hepatol*. 3(1):37–49
5. Beebe A, Seaworth B, Patil N (2015) Rifampicin-induced nephrotoxicity in a tuberculosis patient. *J Clin Tuberc Other Mycobact Dis* 1:13–15
6. Berkrot B (2011) Success rates for experimental drugs falls: study. Thomson Reuters, Toronto
7. Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL (2007) Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov*. 6(1):29–40
8. Thomson R (2012) knowledge-based drug repositioning to drive R&D productivity. Thomson Reuters, Toronto
9. Kerber R (2003) Old drugs, new life. The Boston Globe, Boston
10. Konemann S, Dorr M, Felix SB (2017) Cardiac immunomodulation. In: Nusinovitch U (ed) *The heart in rheumatic, autoimmune and inflammatory diseases*. pp 681–714
11. Palomino JC, Martin A (2013) Is repositioning of drugs a viable alternative in the treatment of tuberculosis? *J Antimicrob Chemother* 68(2):275–283
12. Zumla A, Nahid P, Cole ST (2013) Advances in the development of new tuberculosis drugs and treatment regimens. *Nat Rev Drug Discov* 12(5):388–404
13. Carroll MW, Jeon D, Mountz JM, Lee JD, Jeong YJ, Zia N et al (2013) Efficacy and safety of metronidazole for pulmonary multidrug-resistant tuberculosis. *Antimicrob Agents Chemother* 57(8):3903–3909
14. Ramón-García S, González del Río R, Villarejo AS, Sweet GD, Cunningham F, Barros D et al (2016) Repurposing clinically approved cephalosporins for tuberculosis therapy. 6:34293
15. Leegwater-Kim J, Waters C (2007) Role of tolcapone in the treatment of Parkinson's disease. *Expert Rev Neurother* 7(12):1649–1657
16. Maitra A, Bates S, Shaik M, Evangelopoulos D, Abubakar I, McHugh TD et al (2016) Repurposing drugs for treatment of tuberculosis: a role for non-steroidal anti-inflammatory drugs. *Br Med Bull* 118(1):138–148
17. Lechartier B, Rybniker J, Zumla A, Cole ST (2014) Tuberculosis drug discovery in the post-post-genomic era. *EMBO Mol Med* 6(2):158–168
18. Brindha S, Sundaramurthi JC, Velmurugan D, Vincent S, Gnanadoss JJ (2016) Docking-based virtual screening of known drugs against murE of

- Mycobacterium tuberculosis* towards repurposing for TB. *Bioinformatics* 12(8):359–367
19. Brindha S, Vincent S, Velmurugan D, Ananthkrishnan D, Sundaramurthi JC, Gnanadoss JJ (2017) Bioinformatics approach to prioritize known drugs towards repurposing for tuberculosis. *Med Hypotheses* 103:39–45
 20. Kinnings SL, Xie L, Fung KH, Jackson RM, Xie L, Bourne PE (2010) The *Mycobacterium tuberculosis* drugome and its polypharmacological implications. *PLoS Comput Biol* 6(11):e1000976
 21. Ramakrishnan G, Chandra NR, Srinivasan N (2015) Recognizing drug targets using evolutionary information: implications for repurposing FDA-approved drugs against *Mycobacterium tuberculosis* H37Rv. *Mol BioSyst* 11(12):3316–3331
 22. Yamanishi Y, Kotera M, Kanehisa M, Goto S (2010) Drug-target interaction prediction from chemical, genomic and pharmacological data in an integrated framework. *Bioinformatics* 26(12):i246–i254
 23. Seal A, Ahn YY, Wild DJ (2015) Optimizing drug-target interaction prediction based on random walk on heterogeneous networks. *J Cheminform* 7:40
 24. Isik Z, Baldow C, Cannistraci CV, Schroeder M (2015) Drug target prioritization by perturbed gene expression and network information. *Sci Rep* 5:17417
 25. Cheng F, Liu C, Jiang J, Lu W, Li W, Liu G et al (2012) Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput Biol* 8(5):e1002503
 26. Zhou T, Ren J, Medo M, Zhang YC (2007) Bipartite network projection and personal recommendation. *Phys Rev E: Stat, Nonlin, Soft Matter Phys* 76(4 Pt 2):046115
 27. Guo ZR (2011) Drug promiscuity. *Yao Xue Xue Bao* 46(4):361–369
 28. Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y et al (2014) DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res* 42(D1):Database issue–Database issue1097
 29. Knox C, Law V, Jewison T, Liu P, Ly S, Frolkis A et al (2011) DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res* 39(Database issue):D1035–D1041
 30. Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D et al (2008) DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res* 36(D1):D901–D906
 31. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P et al (2006) DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* 34(Database issue):D668–D672
 32. Huntley RP, Sawford T, Mutowo-Meullenet P, Shypitsyna A, Bonilla C, Martin MJ et al (2015) The GOA database: gene ontology annotation updates for 2015. *Nucleic Acids Res* 43(D1):D1057–D1063
 33. Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R (2009) QuickGO: a web-based tool for gene ontology searching. *Bioinformatics* 25(22):3045–3046
 34. Chan S, Segelke B, Lekin T, Krupka H, Cho US, Kim MY et al (2004) Crystal structure of the *Mycobacterium tuberculosis* dUTPase: insights into the catalytic mechanism. *J Mol Biol* 341(2):503–517
 35. Carroll P, Faray-Kele MC, Parish T (2011) Identifying vulnerable pathways in *Mycobacterium tuberculosis* by using a knockdown approach. *Appl Environ Microbiol* 77(14):5040–5043
 36. Bermingham A, Derrick JP (2002) The folic acid biosynthesis pathway in bacteria: evaluation of potential for antibacterial drug discovery. *BioEssays* 24(7):637–648
 37. Baca AM, Sirawaraporn R, Turley S, Sirawaraporn W, Hol WG (2000) Crystal structure of *Mycobacterium tuberculosis* 7,8-dihydropteroate synthase in complex with pterin monophosphate: new insight into the enzymatic mechanism and sulfa-drug action. *J Mol Biol* 302(5):1193–1212
 38. Liu T, Wang B, Guo J, Zhou Y, Julius M, Njire M et al (2015) Role of folP1 and folP2 genes in the action of sulfamethoxazole and trimethoprim against mycobacteria. *J Microbiol Biotechnol* 25(9):1559–1567
 39. Skold O (2000) Sulfonamide resistance: mechanisms and trends. *Drug Resist Updat* 3(3):155–160
 40. Hammoudeh DI, Zhao Y, White SW, Lee RE (2013) Replacing sulfa drugs with novel DHPS inhibitors. *Future Med Chem* 5(11):1331–1340
 41. Hevener KE, Yun MK, Qi J, Kerr ID, Babaoglu K, Hurdle JG et al (2010) Structural studies of pterin-based inhibitors of dihydropteroate synthase. *J Med Chem* 53(1):166–177
 42. Zhao Y, Hammoudeh D, Yun MK, Qi J, White SW, Lee RE (2012) Structure-based design of novel pyrimido[4,5-c]pyridazine derivatives as dihydropteroate synthase inhibitors with increased affinity. *ChemMedChem* 7(5):861–870
 43. Villella AD, Sanchez-Quitian ZA, Ducati RG, Santos DS, Basso LA (2011) Pyrimidine salvage pathway in *Mycobacterium tuberculosis*. *Curr Med Chem* 18(9):1286–1298
 44. Vanheusden V, Van Rompaey P, Munier-Lehmann H, Pochet S, Herdewijn P, Van Calenbergh S (2003) Thymidine and thymidine-5'-O-monophosphate analogues as inhibitors of *Mycobacterium tuberculosis* thymidylate kinase. *Bioorg Med Chem Lett* 13(18):3045–3048
 45. Sassetti CM, Boyd DH, Rubin EJ (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* 48(1):77–84
 46. Vertessy BG, Toth J (2009) Keeping uracil out of DNA: physiological role, structure and catalytic mechanism of dUTPases. *Acc Chem Res* 42(1):97–106
 47. El-Hajj HH, Zhang H, Weiss B (1988) Lethality of a dut (deoxyuridine triphosphatase) mutation in *Escherichia coli*. *J Bacteriol* 170(3):1069–1075
 48. Gadsden MH, McIntosh EM, Game JC, Wilson PJ, Haynes RH (1993) dUTP pyrophosphatase is an essential enzyme in *Saccharomyces cerevisiae*. *EMBO J* 12(11):4425–4431
 49. Kaur P, Agarwal S, Datta S (2009) Delineating bacteriostatic and bactericidal targets in mycobacteria using IPTG inducible antisense expression. *PLoS ONE* 4(6):e5923
 50. Ramachandran V, Singh R, Yang X, Tunduguru R, Mohapatra S, Khandelwal S et al (2013) Genetic and chemical knockdown: a complementary strategy for evaluating an anti-infective target. *Adv Appl Bioinform Chem* 6:1–13
 51. Black PA, Warren RM, Louw GE, van Helden PD, Victor TC, Kana BD (2014) Energy metabolism and drug efflux in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 58(5):2491–2503
 52. Dhiman RK, Mahapatra S, Slayden RA, Boyne ME, Lenaerts A, Hinshaw JC et al (2009) Menaquinone synthesis is critical for maintaining mycobacterial viability during exponential growth and recovery from non-replicating persistence. *Mol Microbiol* 72(1):85–97
 53. Braback M, Riesbeck K, Forsgren A (2002) Susceptibilities of *Mycobacterium marinum* to gatifloxacin, gemifloxacin, levofloxacin, linezolid, moxifloxacin, telithromycin, and quinupristin-dalfopristin (Synercid) compared to its susceptibilities to reference macrolides and quinolones. *Antimicrob Agents Chemother* 46(4):1114–1116
 54. Eliopoulos GM, Wennersten CB (2002) Antimicrobial activity of quinupristin-dalfopristin combined with other antibiotics against vancomycin-resistant enterococci. *Antimicrob Agents Chemother* 46(5):1319–1324
 55. Evans JC, Trujillo C, Wang Z, Eoh H, Ehrst S, Schnappinger D et al (2016) Validation of CoaBC as a bactericidal target in the coenzyme a pathway of *Mycobacterium tuberculosis*. *ACS Infect Dis* 2(12):958–968
 56. Manjunatha UH, Rao SPS, Kondreddi RR, Noble CG, Camacho LR, Tan BH et al (2015) Direct inhibitors of InhA are active against *Mycobacterium tuberculosis*. *Sci Transl Med* 7(269):269r3
 57. Alland C, Moreews F, Boens D, Carpentier M, Chiusa S, Lonquety M et al (2005) RPBS: a web resource for structural bioinformatics. *Nucleic Acids Res* 33(W1):W44–W49
 58. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498–2504