Tuberculosis 91 (2011) 479–486

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: http://intl.elsevierhealth.com/journals/tube

REVIEW

Open source drug discovery– A new paradigm of collaborative research in tuberculosis drug development

Anshu Bhardwaj [a], Vinod Scaria [b], Gajendra Pal Singh Raghava [c], Andrew Michael Lynn [d], Nagasuma Chandra [e], Sulagna Banerjee [f], Muthukurussi V. Raghunandanan [b], Vikas Pandey [b], Bhupesh Taneja [b], Jyoti Yadav [b], Debasis Dash [b], Jaijit Bhattacharya [g], Amit Misra [h], Anil Kumar [i], Srinivasan Ramachandran [b,**], Zakir Thomas [a,***], Open Source Drug Discovery Consortium [a], Samir K. Brahmachari [a,b,*]

[a] Council of Scientific and Industrial Research, Anusandhan Bhawan, 2 Rafi Marg, New Delhi 110 001, India
[b] Institute of Genomics and Integrative Biology (CSIR), Mall Road, Delhi 110 007, India
[c] Institute of Microbial Technology (CSIR), Sector-39A, Chandigarh 160 036, India
[d] School of Information Technology, Jawaharalal Nehru University, New Delhi 110 067, India
[e] Bioinformatics Centre, Indian Institute of Science, Bangalore 560 012, India
[f] Anna University, K.B. Chandrasekhar Centre, Chromepet, Chennai 600 044, India
[g] Hewlett-Packard, Global Business Park, Mehrauli-Gurgaon Road, Gurgaon 122 002, India
[h] Central Drug Research Institute (CSIR), Lucknow 226 001, India
[i] Department of Chemistry, Sri Sathya Sai University, Prashanti Nilayam 515134, India

ARTICLE INFO

Article history:
Received 22 October 2010
Received in revised form
11 May 2011
Accepted 12 June 2011

Keywords:
Open source
Drug Discovery
Tuberculosis
Malaria
Neglected diseases
Generics

SUMMARY

It is being realized that the traditional closed-door and market driven approaches for drug discovery may not be the best suited model for the diseases of the developing world such as tuberculosis and malaria, because most patients suffering from these diseases have poor paying capacity. To ensure that new drugs are created for patients suffering from these diseases, it is necessary to formulate an alternate paradigm of drug discovery process. The current model constrained by limitations for collaboration and for sharing of resources with confidentiality hampers the opportunities for bringing expertise from diverse fields. These limitations hinder the possibilities of lowering the cost of drug discovery. The Open Source Drug Discovery project initiated by Council of Scientific and Industrial Research, India has adopted an open source model to power wide participation across geographical borders. Open Source Drug Discovery emphasizes integrative science through collaboration, open-sharing, taking up multi-faceted approaches and accruing benefits from advances on different fronts of new drug discovery. Because the open source model is based on community participation, it has the potential to self-sustain continuous development by generating a storehouse of alternatives towards continued pursuit for new drug discovery. Since the inventions are community generated, the new chemical entities developed by Open Source Drug Discovery will be taken up for clinical trial in a non-exclusive manner by participation of multiple companies with majority funding from Open Source Drug Discovery. This will ensure availability of drugs through a lower cost community driven drug discovery process for diseases afflicting people with poor paying capacity. Hopefully what LINUX the World Wide Web have done for the information technology, Open Source Drug Discovery will do for drug discovery.

Introduction

Tuberculosis (TB) still remains a leading cause of deaths world wide despite numerous efforts to control and eradicate. In India alone, the number of deaths average to about 1 person every 1.5 min. Among the first line drugs used are isoniazid, rifampicin, ethambutol, pyrazinamide, streptomycin. Wherever the first line of treatment fails, the second line of therapy is to be given. Examples

* Corresponding author. Tel.: +91 11 2371 0472; fax: +91 11 2371 0618.
** Corresponding author. Tel.: +91 11 2766 6156; fax: +91 11 2766 7471.
*** Corresponding author. Open Source Drug Discovery Consortium. Tel.: +91 11 2331 6763.
E-mail addresses: anshu@csir.res.in (A. Bhardwaj), vinods@igib.in (V. Scaria), raghava@imtech.res.in (G.P.S. Raghava), andrew@mail.jnu.ac.in (A.M. Lynn), nchandra@biochem.iisc.ernet.in (N. Chandra), sulagna.banerjee@gmail.com (S. Banerjee), raghu@igib.in (M.V. Raghunandanan), vikas@igib.in (V. Pandey), btaneja@igib.res.in (B. Taneja), j.yadav@igib.res.in (J. Yadav), ddash@igib.in (D. Dash), jaijit.bhattacharya@hp.com (J. Bhattacharya), amit.cdri@gmail.com (A. Misra), anilk3@gmail.com (A. Kumar), ramu@igib.in (S. Ramachandran), zt@csir.res.in (Z. Thomas), skb@igib.res.in (S.K. Brahmachari).

1472-9792/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tube.2011.06.004

of second line drugs are fluroquinolones, ethionamide, cycloserine, para-aminosalicylic acid, capreomycin, kanamycin and amikacin.[1] The emergence of MDR-TB (Multidrug resistance TB) and XDR-TB (Extensively drug resistance TB) have caused significant concern in eradicating TB.[2] New highly potent and fast acting drugs with short treatment regimen are essentially required for treatment of TB.[3] However, between 1975 and 2004 only 3 out of 1556 new chemical entities arrived in the market for TB treatment.[4]

Several organizations like National Institutes of Health, U.S.A., European Union Framework, Bill & Melinda Gates Foundation and pharmaceutical industries such as Novartis, GlaxoSmithKline, AstraZeneca, Sanofi-Aventis, Johnson & Johnson have funded programs on TB drug discovery. These initiatives have resulted in 6 compound types in preclinical development and 12 compound types in different phases of clinical trials including Gatifloxacin and Moxifloxacin in Phase III.[5] However, in comparison to cancer drug development these efforts appear minimal and limited. The approaches followed in most of these efforts are not available in the open. The need for openly available information on pharmaceutical expertise, compounds, research tools, screenings and analyses is being felt necessary to stimulate research in this neglected diseases area by scientists. In this context an Open Source Drug Discovery (OSDD) approach was proposed by Council of Scientific and Industrial Research (CSIR), India, for tackling TB in 2006 and the project was launched for global participation in 2008.[6,7] The major progress over the last 2 years has made OSDD as an alternate model of Intellectual Property Rights (IPR) protected closed-door drug discovery.[8] The successful implementation of the Human genome project and contribution of open source software and World Wide Web gives sufficient confidence that OSDD model for neglected diseases is a robust alternative.

Vision

The vision of OSDD is to provide affordable healthcare to the developing world by providing a global platform where the best minds can collaborate & collectively endeavor to solve the complex problems associated with discovering novel therapies for neglected tropical diseases like tuberculosis, malaria, leshmaniasis, etc. To achieve this goal, OSDD aims to reduce the risks in the discovery stage by facilitating collaborations between scientists, doctors, technocrats and students through a collaborative platform.

Rationale and approach

Conventional approaches to discovering therapies for TB has met limited success, thus, demanding a more open and innovative strategy to capture the experience of experts and enthusiasm of young researchers at a global scale. To overcome this crucial bottleneck, OSDD is designed to function through a web-based platform where experts from academia, industry including individuals from varied subject areas can interact and contribute to solve complex problems associated with discovering novel therapies. The course from drug discovery to development is carefully planned, yet decentralized in nature. OSDD exemplifies the power of distributed co-creation using web as an organization.

Drug discovery

OSDD emphasizes following the path of integrative science and on using modern tools of communication. All data of OSDD are shared with the entire community through the SysBorg TB (Systems Biology of organisms) portal.[9] Sharing of data among investigators reduces duplication of efforts while giving appropriate credits to the contributors. All data, methods, procedures, algorithms and scripts are available for use, reuse and modification for further activities within the purview of OSDD License.[10] The information available online on the OSDD portal is regarded as a 'Protected Collective Information' by the OSDD sign-in license. Anyone is free to contribute to or use this community property but with the obligation that all improvements and value additions are contributed back to the community. This facilitates the integration of different facets, namely, computational biology, bioinformatics, systems biology, molecular biology, chemo-informatics, medicinal chemistry, experimental pharmacology of the drug discovery process without time delay. This also allows online review of work done and sharing the information on both successful and failed experiments among the members of the community. As the methods and results are already available, other investigators can modify or improve them in order to achieve better results.

The process of integration in OSDD is structured and is referred to as 'baton passing'.[11] It enables scientists to deliver on their core competence and let the results be carried forward by others with respective competences down the pipeline. An example is the case of the drug target glmU (UDP-N-acetylglucosamine pyrophosphorylase, Rv1018c). This was identified as a drug target by a bioinformatics group.[12] Experimentalists are developing assays and screening. A chemo-informatics group is carrying out the task of prediction of ligands, which is being followed up by experimental chemists for synthesizing the molecules. This sequence of work is linked by two research organizations and an academic university located in different parts of India. This web-based collaborative model in the manner of Science 2.0[13] holds great promise for future global open innovative collaboration.

All contributions on the portal are date and time-stamped. This enables downstream researchers to cite references to earlier work. All contributions are bound by OSDD license where the investigators bind themselves to give credit to the works of others and to deliver back to the community all improvements or value additions made to existing information contributed by other members of the community. The agreement to OSDD license confers a responsibility on the individual to return back the property to the OSDD community including data or methods with appropriate value addition. Investigators are encouraged to post new ideas, concepts and challenges to propel the movement towards drug discovery.

The students and scientists would be rewarded for developing novel algorithms, finding drug targets, lead identification and other novel contributions. The recent technological advances of Web 2.0 such as blogging, tagging and social networking have facilitated online publication, editing and collaboration. The SysBorg2.0[9] is built incorporating the tools of Web 2.0 and allows OSDD projects to be implemented in the manner of Science 2.0[13]. The SysBorg2.0 provides the following facilities: Project management system, Laboratory information management system, Workflow system, Learning management system, Data store in RDF (Resource Description Framework), Portlets services, Application Programming Interfaces (APIs) and Web services, 3rd party web services, Grid computing system, Semantic search engine, and Micro-attribution. The SysBorg2.0 was developed using the following technologies: Liferay 5.2.3, CAS 3.0.1, MySQL 5.1, Java OpenID Server, Apache tomcat 6.0.18, Data Vision 1.2.0, Galaxy, dotProject 2.1.2, Moodle 1.9.4, Jena 2.6.0.[14]

One of the salient features of the portal is its micro-attribution system and algorithm to assign credits to contributors. To start with all contributions to the system are tagged by the date-stamp, the time-stamp and the contributor-stamps. Each contribution is conceptualized as a node in a semantic network. Once a contribution is added, the RDF store also updates the semantic links to different other nodes already in the system. The credit points are calculated automatically by a micro-attribution algorithm. Each

semantic connection between nodes adds to the micro-credit to a particular node. The credit points are then updated by the algorithm depending on the semantic connections each node makes, and thus updates the credit points to each contributor node. The micro-attribution to a contributor is thus philosophically the amount of valuable nodes contributed directly or indirectly as evidenced by other nodes connecting to that piece of contribution. The points can be accrued over time for all the contributions, which may be converted into rewards.

Drug development

Once the discoveries are made it is imperative to carry them forward in the development pipeline which can be done in an affordable way if the development work is carried out in countries where the disease is endemic, but where the framework for such development at internationally acceptable standards exists. The benefit that OSDD has is the leadership of CSIR. The partnering institutes of CSIR will be involved in managing the projects with product development mandate. CSIR institutions like the Central Drug Research Institute (CDRI), an OSDD partner, are dedicated to drug discovery and have rich expertise in this area. A public private partnership with these institutions and CRO's involved in drug development as well as pharmaceutical companies interested in partnering in TB drug development will bring together the best minds and the best processes in the development phase.

CSIR has set up a project directorate of OSDD. The product development projects are monitored by the project directorate, the partnering institution and creative inputs from the industry partner. In addition, the results of the projects are available online for community inputs. Also, for tuberculosis, India has the experience in conducting clinical trials. The clinical trials that laid the groundwork for the development of the currently globally accepted Directly Observed Treatment, Short-course (DOTS) was done in a public funded institution of India, Tuberculosis Research Centre. In the past decade a number of CROs have come up in India with specialization in clinical trials. There are Indian pharmaceutical industries that have clinical trial experience on TB drugs. The CSIR-IIIM developed anti TB drug Risorine[15] in partnership with M/s Cadilla Pharmaceuticals. Also, M/s Lupin Pharmaceuticals have been collaborating with CSIR in conducting clinical trials of another anti TB drug.

A study by London School of Economics on neglected diseases reports that major pharma companies prefer to have public partners for clinical development.[16] The pharmaceutical companies rarely have developing country experience to conduct clinical trials and are wary of undertaking risky clinical trials in neglected diseases. The only alternative then is for public funded organizations to take up this responsibility. Therefore, OSDD approach is that public funding should meet substantial expenditure of clinical trials. Such trials in public funded hospitals in collaboration with doctors, who are willingly joining the project like OSDD, will bring costs down. Such trials will be conducted in collaboration with CRO's specialized in clinical trials so that all internationally accepted standards are maintained. OSDD's clinical trial efforts will be monitored by an independent ethical committee which shall ensure that all ethical standards are followed in clinical trials. One distinguishing feature of OSDD programs is that it will not maintain data exclusivity and the data from clinical trials will be made public, bringing openness to clinical trial process. In OSDD model, contributors can apply and seek intellectual property protection on the condition that these are made available to the developing world through a non-exclusive license. In India, business models exist where drugs are made available at affordable prices through market based competition without the exclusivity of intellectual property rights leading to competitive pricing.

Project management

An investigator with interest in participating in OSDD needs to post a project online on the SysBorg TB portal.[9] Thereafter, it is open to review by members within the OSDD community. In order to facilitate this process, some experts of the subject area are invited to provide comments. This open peer review is coordinated by a group of scientist, both from within CSIR and outside. In addition, the proposal is open to comments from other community members. The project details, comments and subsequent communications during the peer review process are open to the community. Further round of discussions are held at this stage in order to scrutinize the costs involved, to check the alignment to the objectives of 10 Work Packages[17] [Supplementary File 1] and to match the best skills of the leader and participants in order to facilitate performance through synergizing the project activity. The entire process of posting and approving proposals in OSDD is carried out in full view of the community with the provision that at any stage any member with relevant information can participate to enrich the proposal. This transparency is hallmark of open peer review process of OSDD. It was also important to study the potential bottlenecks and gaps in implementing and monitoring various projects in OSDD. Towards this, a project was carried out by National Institute of Pharmaceutical Education and Research (NIPER), Mohali, India, as they worked out the gaps and possible solutions in the OSDD supply chain model. This project primarily dealt with understanding the potential bottlenecks and gaps in connecting predictions of computational projects to experimental validation. It also reviewed the resources that are available for the same and provided workflows for connecting these resources through an online portal.

Open project space and open lab notebook drive breakthrough innovation and rule based processes

The OSDD model has two major zones – Innovation driven and Process driven (Figure 1). The innovation driven zone aims to excel in breakthrough innovations and is somewhat free space to exchange ideas, results and collaborate on difficult problems to find solutions faster. No specific order is necessarily observed while making the connections. The process driven zone links with the free zone by carrying out the standardized scientific activities either in collaborative or in out-sourcing mode. Individuals working in the free zone would post projects in Open Project Space and carry on their works by placing the results in Open Lab Notebooks. Projects can be carried out by academic institutions in collaboration with industries as well. In all cases, the results are shared in realtime with the OSDD community. Once again as the results are posted in Open Lab Notebook, the individuals in the free zone can now pick new problems to solve towards advancement. This Collaboration Innovation cycle is facilitated by the SysBorg portal, which is divided into multiple conceptual Webs, namely Open Ideas, Open Project Space and Open Lab Notebooks. Other interesting facets of this portal are forums associated with each web, blogs and wall posts, where the community can share latest research news and resources on TB.

Funding

Government of India in the first phase has committed upto Rs. 50 crores (approx. US $11 million) for TB research in open source projects. OSDD also proposes to raise funding from multilateral or

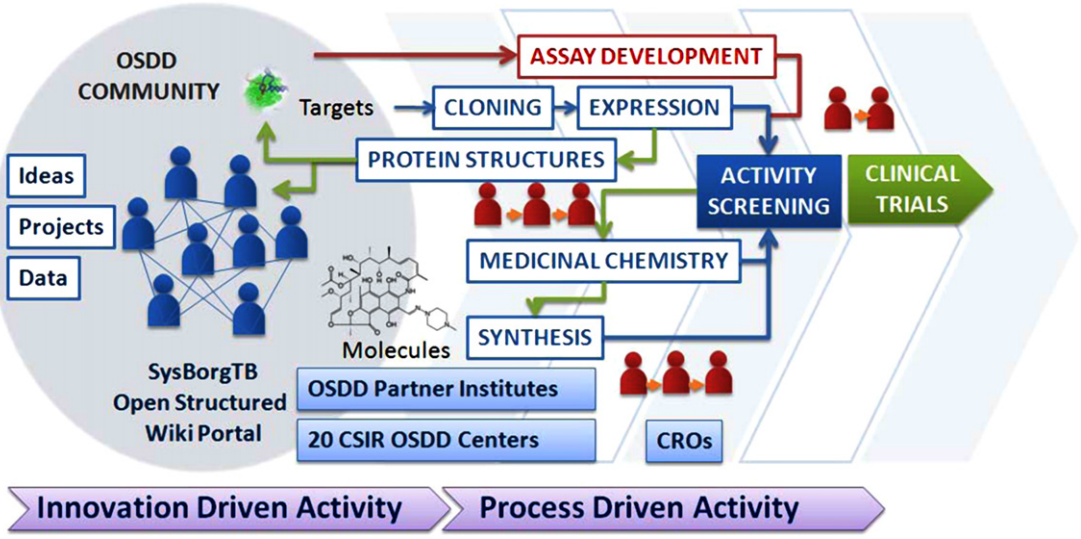

Figure 1. OSDD model conception with dual characteristic of supporting both breakthrough innovation and rule based process drives. This model ensures that activities which need creativity flourish through the online platform in a somewhat free zone. This free innovative zone is connected and integrated in a structured manner by standardized rule based process driven exercises. For example, the free zone may include a set of academic individuals generating various ideas and working on them. The industry partners would fall in the rule based standardized process drive area. For success in drug discovery, activities in both zones are harmonized and working together through the OSDD web portals.[9]

bilateral agencies and philanthropic organizations. OSDD is a major opportunity for Industries to discharge their corporate social responsibility by participating in research into neglected diseases. Public Private Partnership is now an accepted business model and OSDD is using this business model. OSDD is already working with SUN Microsystems, Hewlett-Packard (HP), Infosys Technologies Limited, Jubilant Chemsys and Premas Biotech on various projects ranging from portal design to lead optimization, etc. Another model of collaboration with private sector is illustrated by OSDD's collaborative project of designing and developing the OSDD portal with Infosys, the major global IT giant. Infosys developed this semantic web enabled portal at no cost. OSDD draws upon the strengths of India, namely, large IT community and students, researchers and doctors working on TB, access to patients, well organized Contract Research Organizations (CROs) for drug development and experience in generics.[18]

OSDD, a growing community

Based on the experience so far, sharing and collaborative environment of OSDD is expected to grow. The OSDD model would exploit the system of monetary and non-monetary rewards based on meritocracy to engage biomedical researchers.[18] Since its global launch in September 2008 many partners and collaborators joined this initiative. Collaborators engaged in various activities such as *in silico* biology for drug target identification, chemical synthesis of molecules, designing new assays, have joined this initiative and are actively participating. Several algorithms and strategies have been developed with emphasis on prioritizing non-toxic drug targets. The work done on OSDD is closely monitored by a team of Project investigators (PIs) who posts projects online and ensuring deliverables are met and timelines followed. The deliverables get finalized during the open peer review.

By the end of 2010, OSDD membership has grown beyond 4500 registered participants. The registrants include students, researchers, clinicians, teachers, from all across the globe. They are represented from more than 130 countries with overwhelming majority from India (81%). The other countries are represented in various proportions. A few representative examples are USA (~4%), European Countries (~2%), UK (~0.46%), and Canada (~0.4%). The most striking feature of the community is the diverse representation from many countries across different continents with potential that substantial representation will continue to grow globally. According to the Apache Software Foundation,[19] communities that are open, diverse and meritocratic are more robust than closed ones. The OSDD community has posted more than 180 different projects and most of them are linked to Open Lab NoteBook entries. The projects cover a wide range of different subject areas including genome annotation, screening for lead molecules against selected targets, creating repositories for DNA, preparing clones for protein expression, literature corpus for TB. It is to be noted that this range of projects is characteristic of drug discovery.[20] As part of long reach effort and to provide participating channels to younger students at the college (undergraduate) level, OSDD has adopted 21 university-colleges based on their interest and educational profile across India. Each of these is termed as 'CSIR Center for OSDD.[21] Based on the activities, the OSDD community may be expected to display a fair degree of robustness. It is to be noted that while the number of 4500 OSDD members is indicative of their interest in OSDD, the number of active contributors is variable. It is observed that contributions depend on multiple factors including challenges, merits of investigators, and demands of skill types.

Data sharing

From the beginning, OSDD members adopted basic standards used by the TB community for data submission. For genes, either the Rv identification numbers or accepted standard gene symbols[22] are used. The Rv ids are used more often than gene symbols. Chemical compounds are represented as SMILE (Simplified Molecular Input Line Entry) specifications. The IUPAC system of nomenclature is followed for all biomolecular sequences. In the case of literature, PubMed IDs are considered. As all data conform to these basics, it is easy for members to communicate smoothly and review each other's contributions. Considering the diversity of data types, the data would be stored in RDF format and semantic

searches could be carried out. Additionally we are currently developing and expanding SysBorg data through the **R**[23] programming environment. The data available in SysBorg has been packaged into **R** environment by the OSDD community members.[24]

Data submitted by members are scrutinized through a laid down process check for conforming to the basic data standards. Entries requiring changes are reported to investigators whereupon the submitted page is revised. All revisions are stored in order to keep track and also to serve as repository for investigators to check history. The OSDD portal is managed by a team of system administrators, who keeps a close watch on any spam that is reported by the community and ensures the deactivation of the user who posted the spam. Any other activity of improving or rectifying errors in the content of the ideas/projects/lab notebooks is encouraged as all versions of the respective webs are maintained.

Resources

The OSDD community data although available through SysBorg, needs supplementing from various resources. The existing heterogeneity limits cross-talk between data across multiple resources. In order to overcome this difficulty an integrated platform called TBrowse[25] was developed. Fifty different resources encompassing more than a million data points were integrated by converting the data to a standard format (Generic Feature Format – GFF). This integrated platform provides the largest resource on mycobacterial genomics in a standard interoperable format. Underlying the platform is the GMOD architecture with display of integrative genomics map of *Mycobacterium tuberculosis* (Figure 2). Another significant component of the OSDD federated resource is Computational Resources for Drug Discovery (CRDD).[26] The CRDD Web provides computational resources related to drug discovery through a single site such as CRDD Forum, KiDoq, BIAdb, MycoTb, and Drugpedia. CRDD provides computational resources for researchers in the field of computer-aided drug design and maintains a Wikipedia related to drug discovery. The members of the OSDD community and others may contribute to or host their database or web server on the CRDD portal. CRDD has started a novel initiative called Indipedia to focus specifically on Indian Researchers and their contributions. A related component of this activity is to create a platform for designing *in silico* workflows which facilitates automated and high-throughput computational analysis. Members of the OSDD community have developed Web services using Web Services Description Language (WSDL) standards. The OSDD community is free to use and add more tools to this open source workflow engine. This rich set of resource building was possible because of community participation. The power of community participation was demonstrated recently where the launch of the project Connect 2 Decode has achieved a complex task of re-annotation of *M. tuberculosis* genome by cutting down the time required by a huge margin.[20] This project harnessed the talent pool of undergraduate students who carried out a well planned project using the Web 2.0 tools and completed the annotation in globally accepted standard format in five themes: Gene ontology, Interactome and Pathway, Protein structure and folds, Immunome, Glycomics.[27]

Drug targets

Several drug targets identified using novel approaches such as invariant peptide based functional signature and structural determinant approach,[28–30] intrinsically disordered essential protein (IDEP) approach,[12] adhesin score,[31] interactome and structure pocketomes,[32] flux balance and network analysis[33–36] are being actively pursued in OSDD. All investigators have taken integrative approaches for identifying targets. An example of target TB is shown in Figure 3. Both metabolic and cell wall targets are assigned priority. A few examples include UDP-N-acetylglucosamine pyrophosphorylase (Rv1018c, glmU),[12] Napthoate Synthase (Rv0548c),[37] Fatty Acyl Adenylate Ligase (Rv2941).[38] In all these cases, both academia and the industry partners are working together. OSDD is engaging industries and academic partners to produce soluble proteins in large scale followed by developing assays and then engaging another industry partner for medicinal chemistry for identifying leads. The UDP-N-acetylglucosamine pyrophosphorylase (Rv1018c, glmU) was identified as an IDEP that has a unique C-terminal disordered tail specific to the family of Mycobacteria. The glmU is a bifunctional protein involved in peptidoglycan synthesis comprises of an

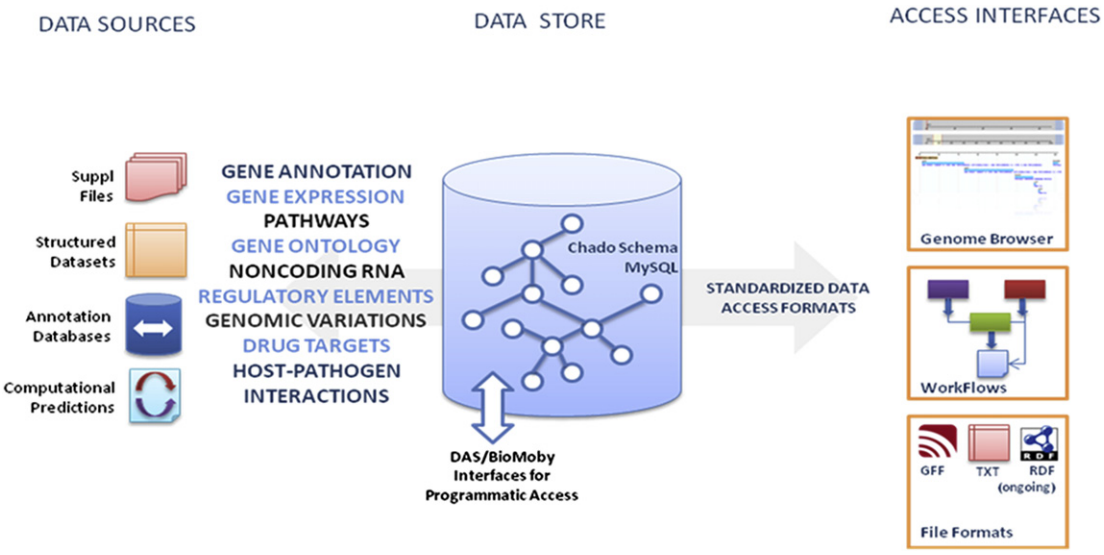

Figure 2. TBrowse data flow: Sourcing from heterogenous resources.

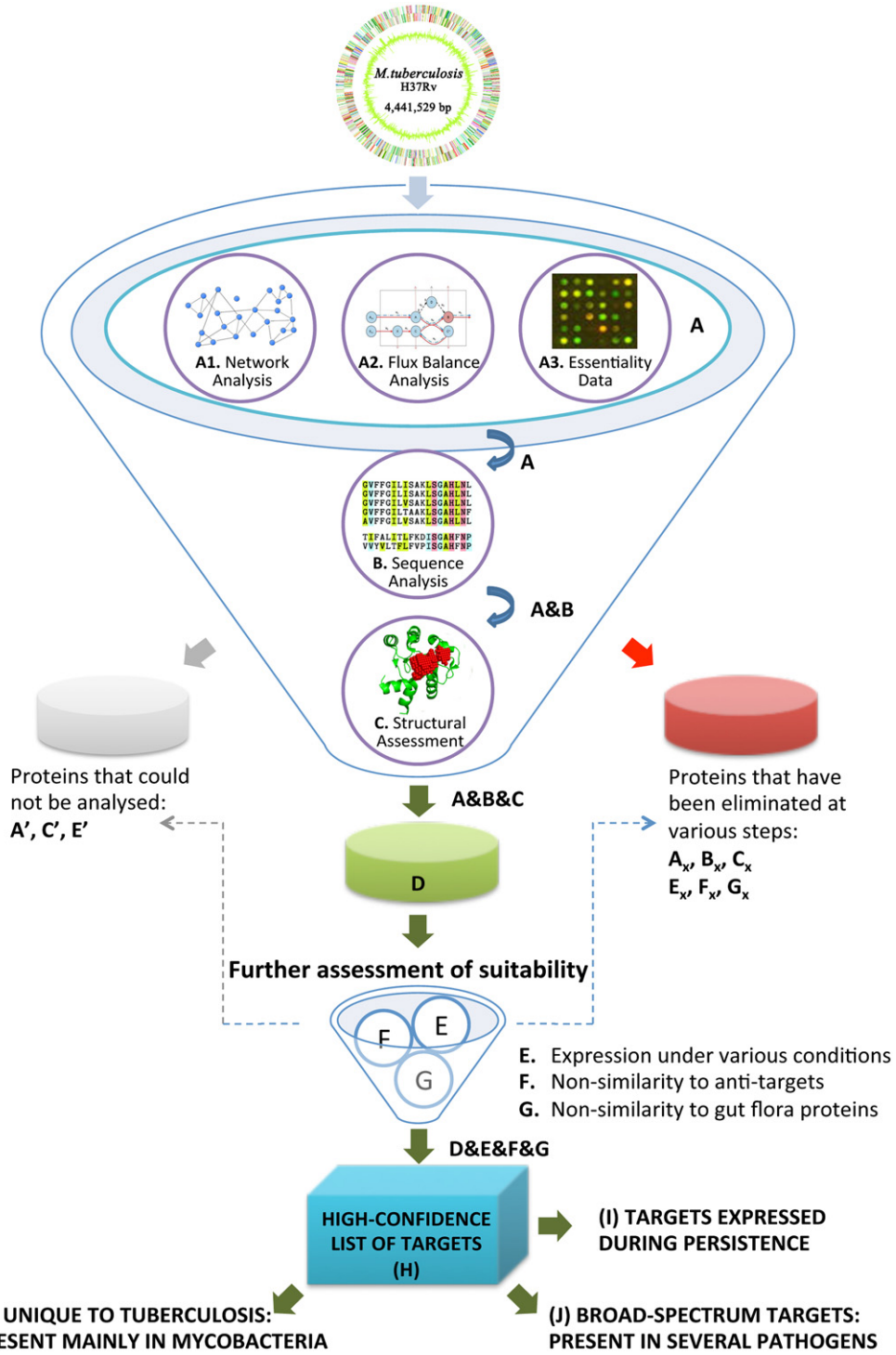

Figure 3. The target TB Target Identification Pipeline.[32] The funnel depicts the order in which the entire proteome of Mtb is considered and analyzed at different layers. A refers to the systems level studies, which includes A1, for network analysis of the interactome; A2, for flux balance analyses of the reactome; and A3, for genome-scale essentiality data determined experimentally as reported by Sassetti et al. Those proteins that passed these filters are indicated as A, and combined with the results of sequence analysis (B), to derive those that passed both filters (depicted as A&B). These were then taken through Filter C, referring to the structural assessment filter, yielding the list of 622 proteins as the D-List (A&B&C). Further steps of filtering are indicated in the smaller funnel as E (expression under various conditions), F (non-similarity to anti-targets) and G (non-similarity to gut flora proteins). Those proteins that pass all the six levels of filtering (indicated as D&E&F&G) form the H-list comprising 451 targets. Additional filters I, J and K used for analyzing the H-List are also indicated. Lists A′, C′ and E′ refer to the set of proteins at A, C and E levels, respectively, that could not be analyzed for lack of appropriate data. Lists Ax, Bx, Cx, Ex, Fx and Gx refer to sets of proteins that failed in that particular filter, but may have passed at other levels.

Uridyltransfer domain at the N-terminal part and an acetyltransfer domain towards the C-terminal end. The latter function is absent in human and makes it as one of the most preferred target for further investigations. Based on experimentally identified 525 inhibitors[39] and computational docking studies, two compounds 6-(p-Aminoanilino)metanilic acid and 4-chloro-N-[2-chloro-4-[4-[4-[3-chloro-4-[(4-chlorophenyl)sulfonylamino]phenyl]sulfonyl-phenyl]phenyl]sulfonylphenyl]benzenesulfonamide were taken as

starting point for discovering new drugs in OSDD. Other three equally potent maleamide derivatives were dropped after community review as they were pointed out to be unstable and lack specific binding. In the case of Napthoate Synthase, plumbagin derivatives are being tested for their inhibitory activity. Apart from the one target- one drug paradigm, a new challenge is posted to OSDD community to design fused inhibitors containing multiple structural scaffolds that bind to different targets of the same pathway using alternate ligand surfaces. This would likely reduce the undesired toxicity and would be potent at lesser concentration. A similar approach is taken to block at various steps of polyketide synthase pathway. The Fatty acyl-AMP Ligases convert fatty acids to acyl-adenylates. Subsequently these adenylates are acylated on to the acyl carrier proteins of polyketide synthases to synthesize lipid metabolites. The product formed by fatty acyl-AMP ligase is central to synthesize the special types of lipids characteristic of *M. tuberculosis*. Inhibitors have been identified for the Fatty Acyl Adenylate Ligase.[38] Their use impairs the formation of the cell wall of *M. tuberculosis* and results in cell death. OSDD has taken up this project from this starting point and would pursue to the aim of fused inhibitors blocking multiple pathways of lipid metabolite synthesis. In all these cases, analogs are being synthesized and tested in the experimental system. The breakthrough innovation group would contribute by suggesting novel molecules meeting the challenges of designing molecules. The industries with operational excellence and following standard practices complying to regulatory bodies will carry out the medicinal chemistry to develop useful leads.

Conclusion

The CSIR, India, a premier organization for research and product development, holding largest number of patents and played a key role in development of generics drug industries in India[18] has taken a bold step forward in initiating a novel open source approach to drug discovery for TB, which is a disease of the developing world. OSDD uses the process of connecting the best minds to optimize productivity and concomitantly reduce cost of drug discovery. OSDD has both zones of breakthrough innovation and rule based process drives linked in a structured manner. Challenges are being taken up by the innovative members whereas processes to execute and document in compliance with regulatory bodies are taken up by industries. The operational management of OSDD is facilitated among the large number of members through the Web portal SysBorg. Major programmes to identify leads have been initiated from the known inhibitors of selected targets. The OSDD community already has more than 4500 members varying in age from 20 years to above 60 years and we predict that the membership may continue to grow. A chief driving force for this growth stems from the new ideas developed in the Open Project Space. In addition, the anchoring factors serving to bind the members into one large functioning OSDD community are unrelenting focus on the fundamental aim of developing affordable drugs by using the approach of integrative science through collaboration and sharing and use of modern Information Communication Technologies for both informal and formal communication. According to Linus' Law, "Given a large enough beta-tester and co-developer base, almost every problem will be characterized quickly and the fix will be obvious to someone"[40] drawing analogy from software open source experience. This conveys that problems can be solved with open community participation and becomes all the more important for a failure prone and complex process like drug discovery.[40] These factors are already serving to ignite other scientific programmes in a similar open source approach, to adopt open innovation model by pharmaceutical companies interested in the pre-competitive space.

Funding: None.

Competing interests: None declared.

Ethical approval: Not required.

Acknowledgments

The authors thank all the OSDD members for their active participation in OSDD and Council of Scientific and Industrial Research, India for funding (Grant No. HCP0001).

Appendix. Supplementary data

Supplementary data related to this article can be found online at doi:10.1016/j.tube.2011.06.004.

References

1. Sacchettini JC, Rubin EJ, Freundlich JS. Drugs versus bugs: in pursuit of the persistent predator *Mycobacterium tuberculosis*. *Nat Rev Microbiol* 2008;**6**:41–52.
2. Chan ED, Iseman MD. Multidrug-resistant and extensively drug-resistant tuberculosis: a review. *Curr Opin Infect Dis* 2008;**21**:587–95.
3. Budha NR, Lee RE, Meibohm B. Biopharmaceutics, pharmacokinetics and pharmacodynamics of antituberculosis drugs. *Curr Med Chem* 2008; **15**:809–25.
4. Casenghi M, Cole ST, Nathan CF. New approaches to filling the gap in tuberculosis drug discovery. *PLoS Med* 2007;**4**:e293.
5. Working group on new TB drugs, http://www.newtbdrugs.org/pipeline.php.
6. *Spicy IP*, http://spicyipindia.blogspot.com/2008/03/spicy-ip-interview-with-dr-samir-k.html; March 19, 2008.
7. Singh S. India takes an open source approach to drug discovery. *Cell* 2008;**133**:201–3.
8. Glaxo tries a linux approach, http://online.wsj.com/article/SB10001424052748703341904575266583403844888.html?mod=WSJ_business_IndustryNews_DHC#articleTabs%3Darticle.
9. The OSDD portals, http://www.osdd.net, http://sysborgtb.osdd.net, http://sysborg2.osdd.net.
10. OSDD licence, http://www.osdd.net/faqs/osddlicence.
11. Kak A. Unlocking the secrets of the code. *Express Pharma* 2010;**5**:30.
12. Anurag M, Dash D. Unraveling the potential of intrinsically disordered proteins as drug targets: application to *Mycobacterium tuberculosis*. *Mol Biosyst* 2009;**5**:1752–7.
13. Waldrop MM. Science 2.0: great new tool, or great risk?, http://www.scientificamerican.com/article.cfm?id=science-2-point-0-great-new-tool-or-great-risk.
14. Bhardwaj A, Scaria V, Thomas Z, Adayikkoth S, Brahmachari SK, Open source drug discovery (OSDD) Consortium. *Collaborative computational technologies for biomedical research*;, ISBN 978-0-470-63803-3; 2011. 321–334.
15. Sharma S, Kumar M, Sharma S, Nargotra A, Koul S, Khan IA. Piperine as an inhibitor of Rv1258c, a putative multidrug efflux pump of *Mycobacterium tuberculosis*. *J Antimicrob Chemother* 2010 Aug;**65**(8). 1694–a701.
16. http://eprints.lse.ac.uk/13947/
17. How does OSDD work, http://www.osdd.net/how-does-osdd-work.
18. Taneja B, Yadav J, Chakraborty TK, Brahmachari SK. An Indian effort towards affordable drugs: generic to designer drugs. *Biotechnol J* 2009;**4**:348–60.
19. Meritocracy. Apache software foundation, http://incubator.apache.org/guides/community.html#introduction.
20. Munos B. Can open-source drug R&D repower pharmaceutical innovation? *Clin Pharmacology Ther* 2010;**87**:534–6.
21. http://www.osdd.net/news-updates/csircentersforopensourcedrugdiscovery
22. Tuberculist, http://tuberculist.epfl.ch/.
23. The R project for statistical computing, http://www.r-project.org/.
24. Ramachandran S, Katiyar A, Sinha A, Bharadwaj A, Dutta A, Raman A, et al. [OSDD consortium], *Mycobacterium tuberculosis* systems biology data in R. *Biobytes* 2009;**5**:40–8.
25. Bhardwaj A, Bhartiya D, Kumar N. Open source drug discovery consortium. In: Scaria V, editor. *TBrowse: an integrative genomics map of Mycobacterium tuberculosis. Tuberculosis (Edinb)* 2009;vol. 89. p. 386–7.
26. Computational resources for drug discovery, http://crdd.osdd.net/, http://crdd.osdd.net/drugpedia/, http://crdd.osdd.net/forum/, http://crdd.osdd.net/raghava/kidoq/, http://crdd.osdd.net/raghava/biadb/, http://crdd.osdd.net/raghava/mycotb/, http://crdd.osdd.net/drugpedia/index.php/Guideline_for_creating_Immunotation_of_Mtb.
27. Connect to decode, https://sites.google.com/a/osdd.net/c2d-01/.

28. Prakash T, Khandelwal M, Dasgupta D, Dash D, Brahmachari SK. CoPS: comprehensive peptide signature database. *Bioinformatics* 2004; **20**:2886–8.
29. Prakash T, Ramakrishnan C, Dash D, Brahmachari SK. Conformational analysis of invariant peptide sequences in bacterial genomes. *J Mol Biol* 2005; **345**:937–55.
30. Computer based method for identifying peptides useful as drug targets, http://www.uspto.gov/web/patents/patog/week05/OG/html/1350-1/US07657378-20100202.html.
31. Sachdeva G, Kumar K, Jain P, Ramachandran S. SPAAN: a software program for prediction of adhesins and adhesin-like proteins using neural networks. *Bioinformatics* 2005;**21**:483–91.
32. Raman K, Yeturu K, Chandra N. Target TB: a target identification pipeline for *Mycobacterium tuberculosis* through an interactome, reactome and genome-scale structural analysis. *BMC Syst Biol* 2008;**2**:109.
33. Raman K, Chandra N. Systems Biology of tuberculosis: Insights for drug discovery. In: Dubitzky W, Southgate J and Fuss H, editors. Understanding the dynamics of biological systems: lessons learned from integrative systems biology, 2011, p. 83–110, Springer ISBN 9781441979636, DOI: 10.1007/978-1-4419-7964-3_5.
34. Raman K, Vashisht R, Chandra N. Strategies for efficient disruption of metabolism in *Mycobacterium tuberculosis* from network analysis. *Mol BioSystems* 2009;**5**:1740–51.
35. Raman K, Chandra N. *Mycobacterium tuberculosis* interactome analysis unravels potential pathways to drug resistance. *BMC Microbiol* 2008;**8**:234.
36. Raman K, Rajagopalan P, Chandra NR. Flux balance analysis of mycolic acid pathway: targets for anti-tubercular drugs. *PLoS Comput Biol* 2005;**1**:e46.
37. Johnston JM, Arcus VL, Baker EN. Structure of naphthoate synthase (MenB) from *Mycobacterium tuberculosis* in both native and product-bound forms. *Acta Crystallogr D Biol Crystallogr* 2005;**61**:1199–206.
38. Arora P, Goyal A, Natarajan VT, Rajakumara E, Verma P, Gupta R, et al. Mechanistic and functional insights into fatty acid activation in *Mycobacterium tuberculosis*. *Nat Chem Biol* 2009;**5**:166–73.
39. GlmU assay protocol for 384-well HT, http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=1376&loc=ea_ras#aDescription.
40. Linus' law, http://en.wikipedia.org/wiki/Linus%27s_Law.