



Editors-in-Chief

Toren Finkel – National Heart, Lung and Blood Institute, National Institutes of Health, USA
Charles Lowenstein – University of Rochester Medical Center, Rochester, NY.

Mycobacterial infections

Systems biology approaches to understanding mycobacterial survival mechanisms

Helena I.M. Boshoff^{1,*}, Desmond S. Lun²

¹Tuberculosis Research Section, LCID, NIAID, NIH, Building 33, 9000 Rockville Pike, Bethesda, MD 20892, United States

²Phenomics and Bioinformatics Research Centre, School of Mathematics and Statistics, and Australian Centre for Plant Functional Genomics, University of South Australia, Mawson Lakes Boulevard, Mawson Lakes, SA 5095, Australia

The advent of high-throughput platforms for the interrogation of biological systems at the cellular and molecular levels has allowed living cells to be observed and understood at a hitherto unprecedented level of detail and has enabled the construction of comprehensive, predictive *in silico* models. Here, we review the application of such high-throughput, systems-biological techniques to mycobacteria – specifically to the pernicious human pathogen *Mycobacterium tuberculosis* (MTb) and its ability to survive in human hosts. We discuss the development and application of transcriptomic, proteomic, regulomic, and metabolomic techniques for MTb as well as the development and application of genome-scale *in silico* models. Thus far, systems-biological approaches have largely focused on *in vitro* models of MTb growth; reliably extending these approaches to *in vivo* conditions relevant to infection is a significant challenge for the future that holds the ultimate promise of novel chemotherapeutic interventions.

Section editor:

Helena Ingrid Boshoff – NIAID, National Institutes of Health, Bethesda, MD, United States

Introduction

The pathogenesis of MTb infection is determined by a complex interplay of response of host cells to bacilli that invade and replicate in macrophages and to extracellular bacilli as well as response of the pathogen to the changing environments in the host. Global approaches to understanding the interaction between host and pathogen offer the prospect of a more integrated understanding of the dynamic behavior of the system. This requires integration of large-scale data that can be used to build a model of this interaction which in turn can answer disease relevant questions and generate testable hypotheses. To unravel the network of such a system, whether regulatory, structural or catalytic, requires heterogeneous datasets to be integrated with such datasets collected from high-throughput techniques, more recently referred to as the 'omics' approaches. To date, most of the data have consisted of transcriptional data superimposed on reconstructed metabolic pathways based on genomic information of both the host and the pathogen [1–6]. However, systems approaches benefit from a broad range of large scale data collection including proteins, metabolites, small stable RNAs, lipids, protein–protein interactions and protein–nucleic acid interactions. We will discuss the omics approaches that have been applied to MTb and finally how some of the data

*Corresponding author.: Helena I.M. Boshoff (hboshoff@niaid.nih.gov)

collected from these omics approaches have been applied to reconstruct metabolic models of MTb metabolism.

Transcriptomics

Early studies with knockout mutants indicated the importance of iron, nutrient, cation and carbon source (reviewed in [7]), but the underlying networks were unknown. DNA microarrays have to date proven the most sensitive tool for unraveling responses of pathogens to their host and of host to pathogen (reviewed in [8]). Transcriptional profiling under *in vivo* relevant conditions has been used to understand the regulatory networks and potential metabolic pathways that may be active during host pathogenesis. Thus, microarray analysis under defined *in vitro* conditions such as exposure to lung surfactant, starvation, nitrosative stress, iron restriction, low oxygen and acidic pH [9–19] has indicated the importance of certain transcriptional pathways for adaptation or survival under these *in vivo* pertinent conditions. Transcriptional profiling of MTb during parasitism of macrophages has enabled the regulatory and catalytic pathways that are relevant during infection, growth and survival of these host cells to be explored and to be correlated to previous *in vitro* transcriptional analyses [20,21]. Thus, the importance of the dormancy response regulator DosR in orchestrating the metabolic slow-down that occurs when respiration on oxygen as terminal electron acceptor becomes limiting [22,23], of WhiB3 in maintaining redox homeostasis [24,25], of the Rv3574-encoded transcriptional regulator in controlling cholesterol metabolism [26], of the RshA-PknB-SigH regulatory pathway in adaptation to stress [27], of IdeR in regulating iron assimilation (reviewed in [28]), of RelA in coordinating the stringent response [11] and of PhoP in regulating the expression of pathogenic lipids and other virulence factors [29] *in vivo* was elegantly demonstrated by a combination of *in vitro* transcriptional profiling, mutant knockout studies and analysis of gene expression in infected hosts. Transcriptional profiling along with analysis of stored lipids in the form of triacylglycerides of MTb expectorated in patient sputum has indicated that these bacilli are probably derived from an environment that allows slow replication based on the observation of an upregulated dormancy response and a high content of triacylglycerides which are associated with stress adaptation [30]. Microarray analyses of MTb derived from other *in vivo* sources such as chronically infected mice [16,17] and a variety of lesions from human TB patients [31] have corroborated the relevance of *in vitro* findings but also pointed to the environmental heterogeneity that MTb experiences *in vivo*. The multitudes of environmental cues that MTb experiences *in vivo* are not only apparent by analysis of different lesions [31] but also evident within one lesion [32] which makes simple extrapolation of data from pooled *in vivo* samples impossible. The transcriptional profiling under defined *in vitro* conditions has been and will be useful

in finding indicator genes, proteins or metabolites of environmental cues that can be used as a sensor to explore the microenvironment of MTb during human pathogenesis [33].

The regulatory networks that control *Corynebacterium glutamicum* transcription have been deduced based on transcriptional profiling data and analyses of individual regulators [34]. *Corynebacterium* is taxonomically related to the mycobacteria and the transcriptional networks that control MTb gene expression have been deduced from conserved genes and binding sites upstream of these genes between these two organisms [35]. A large-scale transcriptional network of MTb was compiled based on gene regulatory interaction from the literature, conserved transcription regulatory pairs between *Escherichia coli* and MTb as well as the complete list of MTb operons with the regulation of all genes in an operon assumed to occur as a single transcriptional unit affected by a transcription factor [36]. The response of this transcriptional network during MTb adaptation to hypoxia or starvation was explored using published microarray data which showed the temporal responses of transcriptional subnetworks at various stages of the transition to non-replicating persistence [36]. This systems levels analysis is useful for probing the *in vivo* responses of MTb to similar environments.

The information that is gained from microarray analyses of gene expression is of course only as good as the functional annotation of the pathogen's genome. Sequencing of genomes of clinical isolates has revealed some differences in open reading frames between strains in terms of both differences in annotation and presence of unique genes [37]. In addition, polymorphisms and deletions in genome sequence have pointed to the relevance or lack thereof of certain metabolic pathways for certain aspects of TB pathogenesis [38–40]. The functional significance of genetic diversity between strains has also been explored by transcriptome analyses [41]. In these analyses, transcriptional profiles during *in vitro* growth could be correlated to the clade that the isolates belonged to, with certain gene clusters being commonly up- or down-regulated in one or more of the 5 global phylogeographical lineages included in the study. The variation of expression observed in the basal transcriptomes of the clinical strains was unexpectedly high if the absolute genomic diversity of MTb strains relative to other bacteria is taken into consideration [42]. The variation in transcription profiles during *in vitro* growth is, however, an indicator of the functional consequences of the high-proportion of non-synonymous single nucleotide polymorphisms that have been observed in the genome sequences between various clinical strains [41,42]. Growth in macrophages of these clinical isolates was associated with clade-, genotype- and strain-specific fitness profiles and transcriptional profiles of these strains during parasitism of resting or activated phagocytes allowed correlations to be made between gene expression signatures and intracellular phenotypes [41]. A core transcriptome asso-

ciated with intracellular growth of a panel of MTb strains was derived consisting of metabolic pathways such as hypoxia, iron storage, oxidative and nitrosative stress, cell wall remodeling, and fatty acid metabolism that were previously reported to be important for macrophage survival [20,21,41]. Although intra-macrophage transcriptomes were more divergent than *in vitro* transcriptomes, the correlation was still higher between strains of the same lineage than inter-genotypic expression profiles and corroborated genotype-specific differences in virulence gene expression that were known to play a role during survival in the host.

The microarrays that have been used to date to analyze transcription of MTb have not probed the regulatory elements associated with intergenic regions such as small stable RNAs. Tiling arrays have been used to analyze the *Listeria* transcriptome including RNAs derived from intergenic regions in response to *in vitro*, *ex vivo* and *in vivo* growth which highlighted the importance of small stable RNAs, anti-sense RNAs and 5' and 3' untranslated regions including riboswitches in its metabolism [43]. The role of small RNAs in controlling gene expression in MTb has only recently started to be explored [44].

Genome wide gene essentiality

Whole genome transposon mutagenesis has allowed the role of genes during growth or survival under *in vivo* relevant conditions to be explored. The essentiality of genes for symbiosis of *Bacteroides* in the gut [45], *Salmonella typhi* tolerance to bile [46], *Haemophilus influenza* in the mouse lung [47] and *in vitro* fitness of *Streptococcus pneumonia* [48] was analyzed by insertion sequencing, high-throughput insertion tracking by deep sequencing or high-throughput parallel sequencing of transposon mutants before and after exposure to the respective stresses. This has allowed every gene in the genomes of these pathogens to be explored for their role under the conditions of interest in a systems-wide approach. Initial saturating transposon mutagenesis studies of MTb assessed the contribution of genes to growth of MTb during growth in rich medium *in vitro* [49,50]. The importance of genes that were non-essential for *in vitro* growth in rich medium for survival in infected mice [51,52] as well as growth on *in vivo*-relevant carbon sources [53] could also be assessed using genome-wide transposon mutagenesis libraries of MTb.

Proteomics

Proteomics has lagged behind transcriptomics due to instrumental and sensitivity problems although recent technical advances especially in the field of tandem mass spectrometry have made it possible to study the pathogenesis of organisms by analyzing the spectrum of secreted proteins, the proteome of virulence mutants in regulatory proteins or small RNAs, the proteome of the pathogen in response to the host or to *in vivo* relevant environmental cues (reviewed in [54]). Differ-

ential proteomics where proteins synthesized under an environmental condition of interest are fluorescently or isotopically labeled or tagged to compare to a basal condition, have allowed subtle variations in protein expression as well as temporal changes in expression to be unraveled [54]. The proteome of MTb in response to deletion of the pyruvate kinase gene, a gene deletion observed in certain members of the MTb complex, was determined and used to define the metabolic pathways that were altered and the associated metabolic consequences deduced from the altered proteome [55]. These studies indicated that loss of pyruvate kinase was associated with a shift in metabolism to growth on lipids for energy generation as seen by increased expression of isocitrate dehydrogenase and beta-oxidation enzymes with concomitant decreased expression of isocitrate lyase and gluconeogenic proteins. The proteome of cell membrane associated proteins has been determined which has given further insight into the localization and associated role of a variety of plasma membrane-associated proteins [56]. Proteome analysis of mycobacterial proteins probably targeted for proteasomal degradation by conjugation to a ubiquitin-like protein (PUP) has revealed that there were 103 candidate PUPylation targets of which 52 were confirmed [57]. An analysis of the pathways containing PUPylated proteins indicated a high preponderance of targets on metabolic pathways associated with rapid growth as well as an over-representation of proteins from the MTb S-nitroso proteome [57,58] suggesting the importance of proteasome degradation for proteasome remodeling during metabolic growth arrest as well as stress. Proteome studies of non-replicating as opposed to replicating MTb have indicated that there are very few differences in the presence of components of the proteome, despite large-scale transcriptional differences. However, the discrepancy between transcriptional and proteomic studies can be addressed by studies of protein turnover which has revealed the subtle dynamics of the response to stresses such as iron starvation [59,60].

Proteomic analyses have been used to develop models of protein-protein interaction networks which can be based on interaction of proteins to form a functional enzymatically active complex or interaction through their role in metabolic or signaling pathways. Wet-lab experiments to experimentally identify protein-protein interactions are plagued by high numbers of false-negatives but especially false-positives. Improvements in methods to identify protein interactions by for example using dual affinity tagged bait proteins that are used for *in vivo* interaction experiments followed by *in vivo* cross-linking with formaldehyde or using dual affinity purification under fully denaturing conditions followed by liquid chromatography-tandem mass spectrometric identification have allowed networks of proteins that interact with virulence factors in *Salmonella* to be revealed [61]. Protein interaction networks that govern mycobacterial metabolism

have been constructed and have led to novel insights of the function of proteins with previously unknown function as well as the presence of a potential signaling pathway [62]. The protein–protein interaction network of MTb was a crucial component in the creation of a database for MTb drug target identification which also takes into account flux balance analysis, network analysis, information on protein essentiality, comparative genomics with the host to assess uniqueness of the target to the pathogen as opposed to its host, as well as structural analysis of proposed ligand-binding sites [63].

Regulomics

High-throughput chromatin immunoprecipitation (ChIP) techniques for regulomic analysis of mycobacteria have been developed and applied. Rodrigue *et al.* [64] have applied ChIP-chip, where DNA fragments bound by immunoprecipitated proteins are identified using microarrays, for various σ factors in *M. bovis* BCG. More recently, ChIP-seq, where DNA fragments are instead identified using next-generation sequencing, has been applied to MTb [65]. ChIP-seq offers a number of advantages over ChIP-chip including lower cost, higher resolution, and a lower requirement for input material [66]. Because of the very small genomes of mycobacteria compared to the mammalian genomes for which ChIP-seq was first developed, it is easy to obtain very high sequencing coverage and hence very high resolution of the location and degree of interaction at protein–DNA interaction sites. For the DosR transcription factor of MTb, ChIP-seq followed by bioinformatics analysis with the computational method CSDeconv [65] allows accurate identification of individual binding sites separated by as few as 40 bp.

Metabolomics

Metabolomic analyses of pathogens are extremely challenging. For MTb in particular, quenching of metabolism in such a way as to preserve the relative metabolite concentrations at a specific point in time without causing leakage of metabolites from cells is extremely challenging. The half-life of cellular metabolites is generally in the order of milliseconds or less thus immediate quenching requires addition of a vast excess of subzero temperature solvent and confirmation that the quenching method does not lead to cell leakage. In MTb metabolite analyses have been done on monolayers of cells grown on nitrocellulose filters to overcome the excessive dilution and cell leakage issues of quenching cell suspension [67,68]. These metabolite analyses have highlighted the role of gluconeogenesis in MTb during growth on beta-oxidation substrates and the importance of this pathway during macrophage infection as well as the unique flow of metabolites through the two halves of the tricarboxylic acid cycle. Substrate availability influences intermediary metabolite concentrations and this in turn affects the production of end-

products such as cell wall lipids. Jain *et al.* [69] demonstrated that growth on propionate and other uneven chain length fatty acids which are expected to increase intracellular pools of methylmalonyl-CoA, affected the production as well as the length of methyl-branched lipids such as sulfolipid-1 and phthiocerol dimycocerosate. The chain length of phthiocerol dimycocerosate was then used as an indicator of carbon substrate pools in MTb infected mice which suggested that MTb may oxidize odd-chain length fatty acids *in vivo* as evidenced by the increased chain length of this cell wall lipid [69].

In silico metabolic modeling

In silico metabolic modeling of living organisms offers the lucrative prospect of integrating data from multiple omics platforms to allow fast, inexpensive *in silico* experiments and fundamental understanding of function and survival. Such modeling falls into two major classes: kinetic modeling, where chemical reaction dynamics are modeled using coupled differential equations, and constraint-based modeling, where only the stoichiometric constraints of metabolic reactions are imposed and predictions of metabolic behavior are made by assuming that organisms regulate metabolic fluxes so as to achieve an objective such as growth maximization. Kinetic modeling is more detailed but suffers from the requirement of knowing many parameters, such as reaction rate constants, which cannot currently be measured on a large scale, and its application is therefore restricted to specific, well-studied pathways. For MTb, kinetic modeling has been applied to describe the dynamics of the tricarboxylic acid (TCA) cycle and glyoxylate bypass [70].

Constraint-based modeling, by contrast, can be achieved by simply knowing reaction stoichiometries and can therefore be applied on a genome-wide scale. Indeed, genome-scale constraint-based modeling has proven to be powerful tool for understanding and accurately predicting metabolic behavior in a large variety of organisms [71]. An example of such a model is *iAF1260* [72], which models the metabolism of *E. coli* and accounts for the action of 1260 enzyme-coding ORFs and for the presence of 2077 reactions and 1039 metabolites.

Using the method of flux-balance analysis (FBA) [73,74], constraint-based models can be used to give predictions of metabolic behavior, such as growth rates and metabolite consumption and production rates. To make predictions using FBA, the assumption needs to be made that organisms regulate metabolic fluxes to achieve particular objectives, such as growth maximization or maximization of ATP production. Such an assumption, although seemingly artificial, nevertheless produces high predictive capability. For example, FBA can be used to predict whether or not an organism grows under genetic modifications such as knockouts and in the presence or absence of particular nutrients and, using

Table 1. Constraint-based models of MTb

Model	Number of reactions	Number of metabolites	Number of ORFs	Reference
MAP	219	197	28	[77]
GSMN-TB	849	739	726	[78]
iNJ661	939	828	661	[79]

iAF1260, FBA predicts *E. coli*'s ability to grow under glucose aerobic conditions with any single knockout of the 1260 modeled ORFs with 92% accuracy and its ability to grow on over 300 different sources of carbon, nitrogen, phosphorous, and sulfur with approximately 76% accuracy [72]. FBA also predicts specific growth rates and metabolite consumption and production rates for *E. coli* with high accuracy over a variety of nutrient sources and over time [75,76].

For MTb, three constraint-based models have been constructed: the mycolic-acid pathway (MAP) model, and two genome-scale models – GSMN-TB and iNJ661 (see Table 1). The MAP model is a relatively small model focusing only on the mycolic acid synthesis pathway of MTb. This pathway is significant because mycolic acids are known to be important for the growth, survival, and pathogenicity of MTb, and InhA – a key enzyme in the pathway – is the target of the front-line anti-tubercular drugs isoniazid and ethionamide. Using this model, it is possible to identify other potential targets for anti-tubercular drugs inhibiting mycolic acid production [77].

GSMN-TB and iNJ661, by contrast, are genome-scale models in that they both model the effect of a significant fraction of metabolic genes in the MTb genome. Both the models give good agreement with experimentally observed growth rates and metabolite consumption rates. Moreover, both the models were validated by comparing experimental measurements of essentiality using transposon site hybridization (TraSH) [50] with *in silico* predictions and achieved high levels of accuracy (78% and 65% correct predictions, respectively [78,79]).

Given genome-scale metabolic models, it becomes possible to interrogate these models to simulate and understand MTb. In particular, it becomes possible to predict, *in silico*, gene or gene combinations that are necessary for the survival of MTb in defined growth media or in an environment reflecting that of host infection. targetTb [63], an *in silico* pipeline for tuberculosis drug target identification, uses predictions obtained from FBA to identify potential targets.

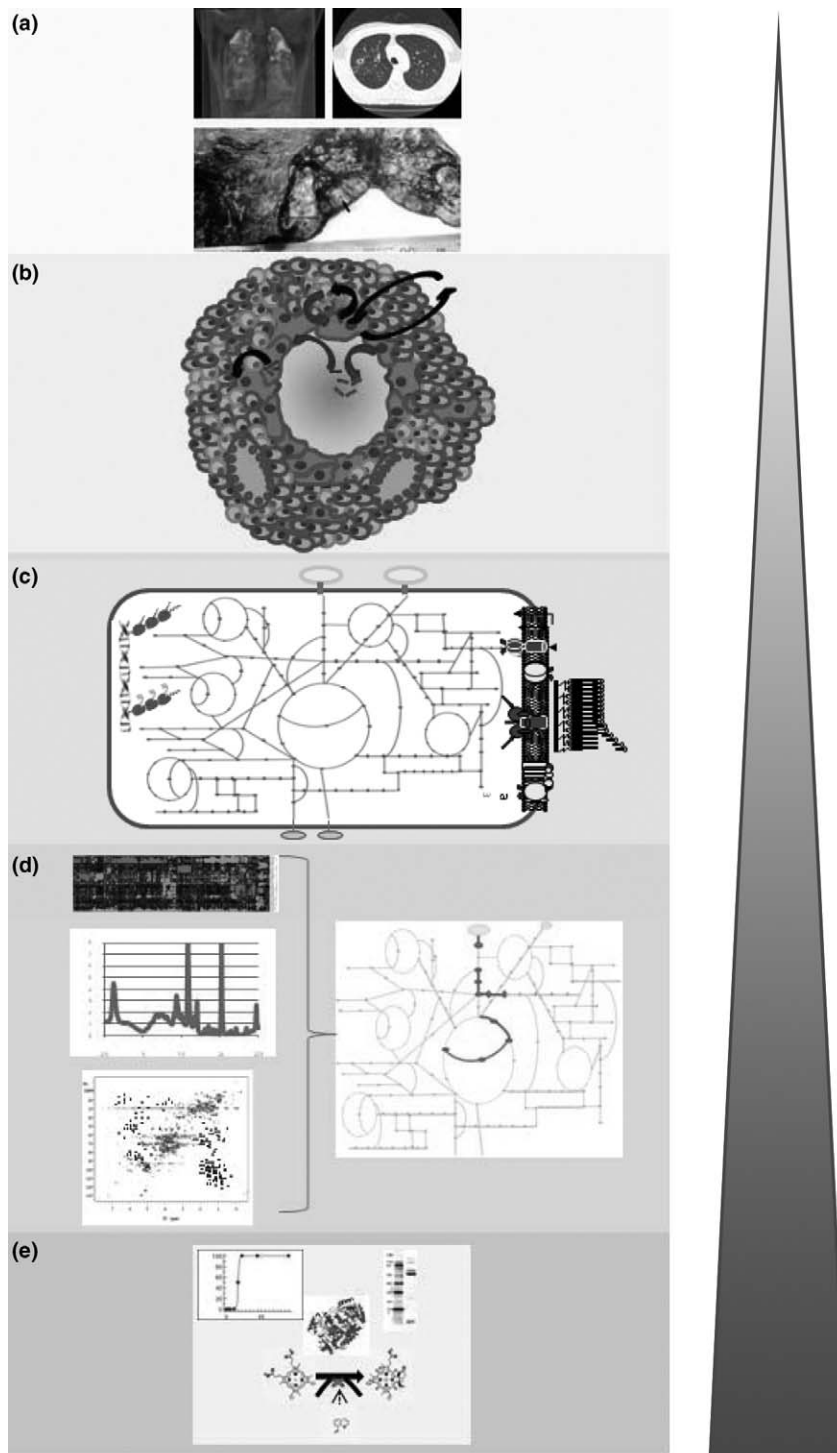
It becomes possible also to understand mycobacterial metabolic capability and how particular perturbations, such as the application of a drug, affect this capability and the organism's ability to survive. In a novel method for coupling FBA with measurements of gene expression called E-flux, Colijn *et al.* [80] use the GSMN-TB and MAP models coupled with over 400 genome-wide expression measurements

[19,81] to identify drugs, drug combinations, and nutrient conditions that inhibit mycolic acid production and the specificity of this inhibition. E-flux correctly predicts seven of the eight known fatty acid inhibitors in the tested expression compendium, including isoniazid and ethionamide, and makes accurate predictions regarding the specificity of the action of these compounds. Further, E-flux predicts several additional potential modulators of mycolic acid biosynthesis in MTb. Thus, E-flux demonstrates the capability of *in silico* modeling to characterize the mechanism of drugs with unknown mechanism.

In addition to employing either kinetic modeling or constraint-based modeling, it is possible to simultaneously employ both approaches, using kinetic modeling for detailed modeling of particular pathways of interest and using constraint-based modeling to extend beyond these pathways to the genome scale. Fang *et al.* [82] have employed such a combined approach to quantitatively model the response of MTb to the growth inhibitors 3-nitropropionate and 5'-O-(N-salicylsulfamoyl) adenosine under particular nutrient conditions, and achieved good correspondence with experimentally measured dose-response curves.

Conclusion

The ultimate goal of systems biology of pathogens such as MTb is to understand the metabolism of the organism in the context of infection (Fig. 1). In the past few years much progress has been made in the various omics approaches aimed at understanding MTb metabolism. Clearly, the biggest challenge will be applying these high-throughput methods to *in vivo* models of this disease where the vast excess of host proteins, DNA, RNA and metabolites makes identification of the bacterial derived products extremely challenging. Some transcriptional data are available for mouse and even human tissues infected with MTb [16,17,31] but converting this data to a form that allows modeling of metabolism has proven difficult. Thus, the most robust data are currently only available for various *in vitro* models of MTb growth. An understanding of the environment in which MTb resides in the host is, however, required to reconstruct the metabolic models that would be most pertinent to *in vivo* pathogenesis. Thus, determination of extracellular metabolites as well as oxygen concentration in the various types of granulomas would allow flux through the metabolic pathways to be modeled. A recent study measured the oxygen concentration



Drug Discovery Today: Mechanisms

Figure 1. Development of systems approaches for understanding MTb infection requires integrating information from all levels of understanding of disease. At the highest level **(a)**, information is gathered from tuberculosis patients ranging from understanding the pathology of disease from clinical data and imaging studies and the epidemiology of human genetic variation in disease susceptibility and progression. Modeling of the dynamics of infection at the level of the granuloma **(b)** builds on information acquired from host immunology as well as pathogen biology. Genomic information of MTb as well as clinical strains of MTb has allowed metabolic networks **(c)** to be constructed that mirror the metabolic capabilities of this organism as well as the production of metabolites that play a role in growth, survival and virulence. Data gathered from 'omics' studies such as transcriptomics, metabolomics and proteomics have allowed reconstruction of dynamic metabolic networks that capture information of metabolic fluxes that occur under various environmental conditions **(d)**. The majority of work to date has been performed at the base level which includes focused studies of proteins, their function and structure and their interaction with small molecules **(e)**.

in MTb- and *M. bovis*-infected rabbit granulomas to be measured which demonstrated the extreme hypoxic environment in which MTb can reside [83]. However, these studies do not address the spectrum of oxygen concentrations and associated substrate concentrations that can be expected in the various heterogeneous granulomas that are observed in human TB patients. Metabolomic analyses of granulomas would be key to developing systems biology approaches to MTb in the context of human pathogenesis. This would allow us to fully exploit metabolic models for the identification of metabolic chokepoints *in vivo* that could be targeted for chemotherapeutic intervention.

Acknowledgements

This work was supported in part by the Intramural Research Program of the NIH, NIAID. We thank Clifton Barry 3rd for helpful suggestions.

References

- Rizzetto, L. and Cavalieri, D. (2010) A systems biology approach to the mutual interaction between yeast and the immune system. *Immunobiology* 215 (9–10), 762–769
- Napolitani, G. *et al.* (2005) Selected Toll-like receptor agonist combinations synergistically trigger a T helper type 1-polarizing program in dendritic cells. *Nat. Immunol.* 6, 769–776
- Kirschner, D.E. *et al.* (2010) Tuberculosis: global approaches to a global disease. *Curr. Opin. Biotechnol.* 21 (4), 524–531
- Sullivan, D.E. *et al.* (2010) Data integration for dynamic and sustainable systems biology resources: challenges and lessons learned. *Chem. Biodivers.* 7 (5), 1124–1141
- Wang, A. *et al.* (2010) A systemic network for *Chlamydia pneumoniae* entry into human cells. *J. Bacteriol.* 192 (11), 2809–2815
- Wang, Y.C. *et al.* (2010) Global screening of potential *Candida albicans* biofilm-related transcription factors via network comparison. *BMC Bioinform.* 11, 53
- Hingley-Wilson, S.M. *et al.* (2003) Survival perspectives from the world's most successful pathogen, *Mycobacterium tuberculosis*. *Nat. Immunol.* 4, 949–955
- Sturdevant, D.E. *et al.* (2010) Host–microbe interaction systems biology: lifecycle transcriptomics and comparative genomics. *Future Microbiol.* 5 (2), 205–219
- Bacon, J. *et al.* (2007) Lipid composition and transcriptional response of *Mycobacterium tuberculosis* grown under iron-limitation in continuous culture: identification of a novel wax ester. *Microbiology* 153 (Pt 5), 1435–1444
- Betts, J.C. *et al.* (2002) Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol. Microbiol.* 43, 717–731
- Dahl, J.L. *et al.* (2003) The role of RelMtb-mediated adaptation to stationary phase in long-term persistence of *Mycobacterium tuberculosis* in mice. *Proc. Natl. Acad. Sci. U. S. A.* 100, 10026–10031
- Deb, C. *et al.* (2009) A novel *in vitro* multiple-stress dormancy model for *Mycobacterium tuberculosis* generates a lipid-loaded, drug-tolerant, dormant pathogen. *PLoS One* 4, e6077
- Fisher, M.A. *et al.* (2002) Microarray analysis of the *Mycobacterium tuberculosis* transcriptional response to the acidic conditions found in phagosomes. *J. Bacteriol.* 184, 4025–4032
- Rustad, T.R. *et al.* (2008) The enduring hypoxic response of *Mycobacterium tuberculosis*. *PLoS One* 3, e1502
- Schwab, U. *et al.* (2009) Transcriptional responses of *Mycobacterium tuberculosis* to lung surfactant. *Microb. Pathog.* 46, 185–193
- Talaat, A.M. *et al.* (2004) The temporal expression profile of *Mycobacterium tuberculosis* infection in mice. *Proc. Natl. Acad. Sci. U. S. A.* 101, 4602–4607
- Talaat, A.M. *et al.* (2007) *Mycobacterial bacilli* are metabolically active during chronic tuberculosis in murine lungs: insights from genome-wide transcriptional profiling. *J. Bacteriol.* 189, 4265–4274
- Voskuil, M.I. *et al.* (2003) Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. *J. Exp. Med.* 198, 705–713
- Boshoff, H.I.M. *et al.* (2004) The transcriptional responses of *Mycobacterium tuberculosis* to inhibitors of metabolism. *J. Biol. Chem.* 279, 40174–40184
- Rohde, K.H. *et al.* (2007) *Mycobacterium tuberculosis* invasion of macrophages: linking bacterial gene expression to environmental cues. *Cell Host Microbe* 2, 352–364
- Schnappinger, D. *et al.* (2003) Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. *J. Exp. Med.* 198, 693–704
- Leistikow, R.L. *et al.* (2010) The *Mycobacterium tuberculosis* DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy. *J. Bacteriol.* 192 (6), 1662–1670
- Park, H.D. *et al.* (2003) Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol. Microbiol.* 48, 833–843
- Singh, A. *et al.* (2009) *Mycobacterium tuberculosis* WhiB3 maintains redox homeostasis by regulating virulence lipid anabolism to modulate macrophage response. *PLoS Pathog.* 5, e1000545
- Banaiee, N. *et al.* (2006) Regulation of *Mycobacterium tuberculosis* whiB3 in the mouse lung and macrophages. *Infect. Immun.* 74, 6449–6457
- Kendall, S.L. *et al.* (2007) A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *Mol. Microbiol.* 65, 684–699
- Park, S.T. *et al.* (2008) Regulation of the SigH stress response regulon by an essential protein kinase in *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* 105, 13105–13110
- Rodriguez, G.M. (2006) Control of iron metabolism in *Mycobacterium tuberculosis*. *Trends Microbiol.* 14, 320–327
- Gonzalo-Asensio, J. *et al.* (2008) PhoP: a missing piece in the intricate puzzle of *Mycobacterium tuberculosis* virulence. *PLoS One* 3, e3496
- Garton, N.J. *et al.* (2008) Cytological and transcript analyses reveal fat and lazy persistor-like bacilli in tuberculous sputum. *PLoS Med.* 5, e75
- Rachman, H. *et al.* (2006) Unique transcriptome signature of *Mycobacterium tuberculosis* in pulmonary tuberculosis. *Infect. Immun.* 74, 1233–1242
- Ryan, G.J. *et al.* (2010) Multiple *M. tuberculosis* phenotypes in mouse and guinea pig lung tissue revealed by a dual-staining approach. *PLoS One* 5 (6), e11108
- Timm, J. *et al.* (2003) Differential expression of iron-, carbon-, and oxygen-responsive mycobacterial genes in the lungs of chronically infected mice and tuberculosis patients. *Proc. Natl. Acad. Sci. U. S. A.* 100, 14321–14326
- Baumbach, J. (2007) CoryneRegNet 4.0 – a reference database for corynebacterial gene regulatory networks. *BMC Bioinform.* 8, 429
- Krawczyk, J. *et al.* (2009) From *Corynebacterium glutamicum* to *Mycobacterium tuberculosis* – towards transfers of gene regulatory networks and integrated data analyses with MycoRegNet. *Nucleic Acids Res.* 37, e97
- Balazsi, G. *et al.* (2008) The temporal response of the *Mycobacterium tuberculosis* gene regulatory network during growth arrest. *Mol. Syst. Biol.* 4, 225
- Reddy, T.B. *et al.* (2009) TB database: an integrated platform for tuberculosis research. *Nucleic Acids Res.* 37 (Database issue), D499–508
- Caws, M. *et al.* (2008) The influence of host and bacterial genotype on the development of disseminated disease with *Mycobacterium tuberculosis*. *PLoS Pathog.* 4, e1000034
- Tsolaki, A.G. *et al.* (2005) Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* 43, 3185–3191
- Tsolaki, A.G. *et al.* (2004) Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains. *Proc. Natl. Acad. Sci. U. S. A.* 101, 4865–4870
- Homolka, S. *et al.* (2010) Functional genetic diversity among *Mycobacterium tuberculosis* complex clinical isolates: delineation of

- conserved core and lineage-specific transcriptomes during intracellular survival. *PLoS Pathog.* 6 (7), e1000988
- 42 Hershberg, R. *et al.* (2008) High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol.* 6, e311
- 43 Toledo-Arana, A. *et al.* (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* 459, 950–956
- 44 Arnvig, K.B. and Young, D.B. (2009) Identification of small RNAs in *Mycobacterium tuberculosis*. *Mol. Microbiol.* 73, 397–408
- 45 Goodman, A.L. *et al.* (2009) Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe* 6, 279–289
- 46 Langridge, G.C. *et al.* (2009) Simultaneous assay of every *Salmonella typhi* gene using one million transposon mutants. *Genome Res.* 19, 2308–2316
- 47 Gawronski, J.D. *et al.* (2009) Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for Haemophilus genes required in the lung. *Proc. Natl. Acad. Sci. U. S. A.* 106, 16422–16427
- 48 van Opijnen, T. *et al.* (2009) Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods* 6, 767–772
- 49 Lamichhane, G. *et al.* (2003) A postgenomic method for predicting essential genes at subsaturation levels of mutagenesis: application to *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 7213–7218
- 50 Sasseti, C.M. *et al.* (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* 48, 77–84
- 51 Lamichhane, G. *et al.* (2005) Designer arrays for defined mutant analysis to detect genes essential for survival of *Mycobacterium tuberculosis* in mouse lungs. *Infect. Immun.* 73, 2533–2540
- 52 Sasseti, C.M. and Rubin, E.J. (2003) Genetic requirements for mycobacterial survival during infection. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12989–12994
- 53 Pandey, A.K. and Sasseti, C.M. (2008) Mycobacterial persistence requires the utilization of host cholesterol. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4376–4380
- 54 Bhavsar, A.P. *et al.* (2010) Proteomics as a probe of microbial pathogenesis and its molecular boundaries. *Future Microbiol.* 5 (2), 253–265
- 55 Chavadi, S. *et al.* (2009) Global effects of inactivation of the pyruvate kinase gene in the *Mycobacterium tuberculosis* complex. *J. Bacteriol.* 191, 7545–7553
- 56 Sinha, S. *et al.* (2002) Proteome analysis of the plasma membrane of *Mycobacterium tuberculosis*. *Comp. Funct. Genom.* 3, 470–483
- 57 Watrous, J. *et al.* (2010) Expansion of the mycobacterial ‘PUPylome’. *Biosyst.* 6 (2), 376–385
- 58 Rhee, K.Y. *et al.* (2005) S-nitroso proteome of *Mycobacterium tuberculosis*: enzymes of intermediary metabolism and antioxidant defense. *Proc. Natl. Acad. Sci. U. S. A.* 102, 467–472
- 59 Rao, P.K. and Li, Q. (2009) Principal component analysis of proteome dynamics in iron-starved *Mycobacterium tuberculosis*. *J. Proteomics Bioinform.* 2, 19–31
- 60 Rao, P.K. *et al.* (2008) Protein dynamics in iron-starved *Mycobacterium tuberculosis* revealed by turnover and abundance measurement using hybrid-linear ion trap-Fourier transform mass spectrometry. *Anal. Chem.* 80, 6860–6869
- 61 Chowdhury, S.M. *et al.* (2009) A method for investigating protein–protein interactions related to *Salmonella typhimurium* pathogenesis. *J. Proteome Res.* 8, 1504–1514
- 62 Cui, T. *et al.* (2009) Uncovering new signaling proteins and potential drug targets through the interactome analysis of *Mycobacterium tuberculosis*. *BMC Genomics* 10, 118
- 63 Raman, K. *et al.* (2008) targetTB: A target identification pipeline for *Mycobacterium tuberculosis* through an interactome, reactome, and genome-scale structural analysis. *BMC Syst. Biol.* 2, 109
- 64 Rodrigue, S. *et al.* (2007) Identification of mycobacterial σ binding sites by chromatin immunoprecipitation assays. *J. Bacteriol.* 189, 1505–1513
- 65 Lun, D.S. *et al.* (2009) A blind deconvolution approach to high-resolution mapping of transcription factor binding sites from ChIP-seq data. *Genome Biol.* 10, R142
- 66 Mardis, E.R. (2007) ChIP-seq: welcome to the new frontier. *Nat. Methods* 4, 613–614
- 67 de Carvalho, L.P. *et al.* (2010) Activity-based metabolomic profiling of enzymatic function: identification of Rv1248c as a mycobacterial 2-hydroxy-3-oxoadipate synthase. *Chem. Biol.* 17 (4), 323–332
- 68 Marrero, J. *et al.* (2010) Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for *Mycobacterium tuberculosis* to establish and maintain infection. *Proc. Natl. Acad. Sci. U. S. A.* 107 (21), 9819–9824
- 69 Jain, M. *et al.* (2007) Lipidomics reveals control of *Mycobacterium tuberculosis* virulence lipids via metabolic coupling. *Proc. Natl. Acad. Sci. U. S. A.* 104, 5133–5138
- 70 Singh, V.K. and Ghosh, I. (2006) Kinetic modeling of tricarboxylic acid cycle and glyoxylate bypass in *Mycobacterium tuberculosis*, and its application to assessment of drug targets. *Theor. Biol. Med. Model.* 3, 27
- 71 Milne, C.B. *et al.* (2009) Accomplishments in genome-scale *in silico* modeling for industrial and medical biotechnology. *Biotechnol. J.* 4, 1653–1670
- 72 Feist, A.M. *et al.* (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.* 3, 121
- 73 Orth, J.D. *et al.* (2010) What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248
- 74 Raman, K. and Chandra, N. (2009) Flux balance analysis of biological systems: applications and challenges. *Briefings Bioinform.* 10, 435–449
- 75 Mahadevan, R. *et al.* (2002) Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* 83, 1331–1340
- 76 Edwards, J.S. *et al.* (2001) *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* 19, 125–130
- 77 Raman, K. *et al.* (2005) Flux balance analysis of mycolic acid pathway: targets for anti-tubercular drugs. *PLoS Comput. Biol.* 1, e46
- 78 Beste, D.J. *et al.* (2007) GSMN-TB: a web-based genome-scale network model of *Mycobacterium tuberculosis* metabolism. *Genome Biol.* 8, R89
- 79 Jamshidi, N. and Palsson, B.O. (2007) Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the *in silico* strain *iNJ661* and proposing alternative drug targets. *BMC Syst. Biol.* 1, 26
- 80 Colijn, C. *et al.* (2009) Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production. *PLoS Comput. Biol.* 5, e1000489
- 81 Karakousis, P.C. *et al.* (2008) Altered expression of isoniazid-regulated genes in drug-treated dormant *Mycobacterium tuberculosis*. *J. Antimicrob. Chemoth.* 61, 323–331
- 82 Fang, X. *et al.* (2009) A systems biology framework for modeling metabolic enzyme inhibition of *Mycobacterium tuberculosis*. *BMC Syst. Biol.* 3, 92
- 83 Via, L.E. *et al.* (2008) Tuberculous granulomas are hypoxic in guinea pigs, rabbits, and nonhuman primates. *Infect. Immun.* 76, 2333–2340