Metabolic Modeling

Regulatory Modeling

Immune System Modeling

Protein-Protein Interaction Networks

Computational Epidemiology

Finding Drug Targets

Finding Drugs

Diagnostic Biomarker Discovery

Clinical Deployment Modeling

Setting the Research Agenda

Where Tuberculosis Meets Computation:

10 Points of Intersection

By Katharine Miller

Computational approaches to tuberculosis are unavoidable.

The growing threats of multi-drug resistant (MDR) and extensively drug resistant (XDR) tuberculosis (TB) are spurring worldwide interest in faster and more innovative research approaches, such as computation offers. And, as in other areas of biomedicine, high-throughput experiments are yielding a data deluge: The bug's bacterial genome was sequenced a decade ago and more than 26 public databases are now accumulating vast and varied information about the disease—all of it ripe for analysis.

In addition, computation makes an appealing complement to experimentation. In the lab, because the bacterium (*Mycobacterium tuberculosis* or *Mtb*) grows slowly (replicating only once a day), one experiment might require months to complete. By contrast, a virtual experiment might take seconds—and doesn't require rigorous safety precautions.

Computation also has the capacity to address important questions in TB research. Simply flipping through the TB Research Roadmap (published by the World Health Organization's Stop TB Partnership in 2011) reveals the many ways computation can contribute to developing TB drugs, diagnostics and vaccines.

In addition, notes the Roadmap, systems biologists are uniquely situated to study one of *Mtb*'s big mysteries: how it can survive inside the human lung for years—seemingly and inexplicably protected by the very immune system that should wipe it out. Only about 10 percent of the 2 billion people infected worldwide develop active disease from the get-go (and will die if not treated); the rest develop latent disease, which they control but cannot clear. And about 10 percent of latent infections will transition to active disease later in a person's life. By studying *Mtb* as a whole—rather than by looking at its individual parts—systems biologists can tease out how the bug manages these stunts.

But is TB research really benefiting from computation's promise?

Here we've highlighted 10 ways computation is currently making a difference to this problem of global significance. It's a non-exhaustive sampler, designed to whet your appetite. But the exercise of finding key points of intersection between an important infectious disease and computation is instructive in its own way: It provides a window into a disease often described as a black box and suggests novel ways to gain insight about this mysterious pathogen.

Systems Biology of TB Metabolism

To kill a bacterium, researchers try to determine what genes it needs to survive. In the wet lab, experimental biologists identify essential genes by knocking them out one at a time and then observing the result: Does the bug thrive or die? Systems biologists do this same exercise *in silico*, building metabolic models and using them to identify essential genes.

In 2007, both **Bernard Palsson**'s systems biology lab at the University of California, San Diego, and **Johnjoe McFadden**'s molecular genetics lab at the University of Surrey

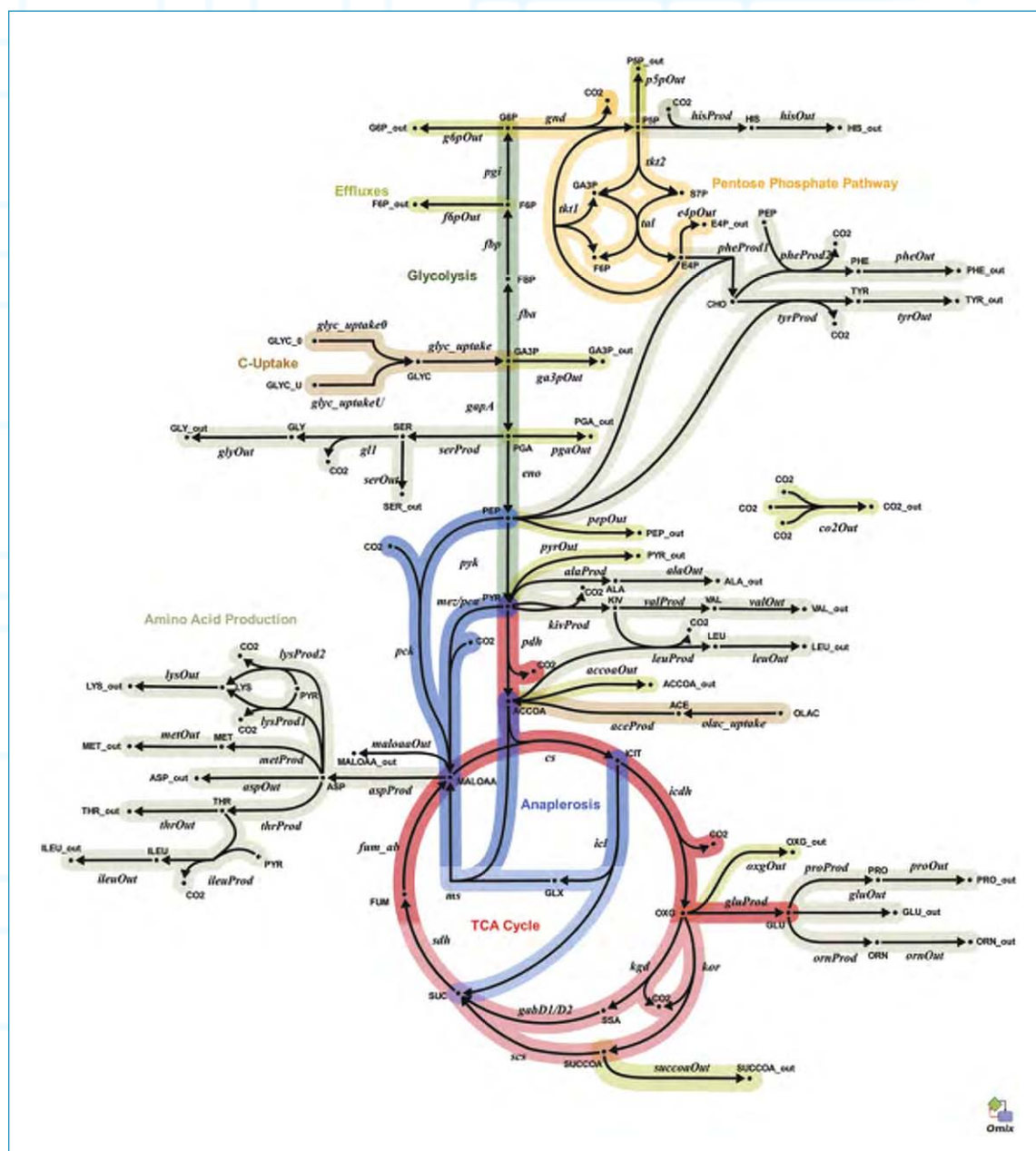published genome-scale network models of *Mtb* metabolism. The McFadden team's model involved 726 genes, 849 reactions, and 739 metabolites and was calibrated by growing *Mycobacterium bovis*, a close relative of *Mtb*, in a steady state. The researchers then used metabolic flux analysis to simulate the flow of metabolites through the network.

McFadden says he thinks of fluxes as traffic through an island road network where the bacillus is the island—the United Kingdom, say—and substrates enter at the ports and are transported through cities (various chemical reactions). "If the rate of traffic going through the networks is steady," he says, "you can go to a port at Plymouth where some product is being produced, measure the rate of production and infer the fluxes inside the network using linear algebra."

The researchers then looked at what happens to the fluxes when various genes are knocked out. "Once you have the model, it becomes a virtual cell," McFadden says. "So you can do experiments instantaneously that would take months and months in the lab." If a gene is essential, the fluxes through the network change, creating blockades that make it impossible for the bacterium to survive. McFadden's lab's analysis of which *Mtb* genes were essential found a 75 to 80 percent matchup between model predictions and lab results.

McFadden concedes that the metabolic models remain incomplete, and that it's still the early days for TB systems biology. But, look-

starting to open it up. We're collecting the data to map the innards of TB and using that to create predictive models."

Metabolic network of the central metabolism of Mycobacterium bovis *BCG. Reprinted from Beste DJ, et al., ^{13}C metabolic flux analysis identifies an unusual route for pyruvate dissimilation in mycobacteria which requires isocitrate lyase and carbon dioxide fixation.* PLoS Pathog (2011) 7(7):e1002091.

ing under the hood of TB has the potential to give researchers a more accurate and predictive view of how TB works.

"TB is a black box," says **James Galagan**, **PhD**, associate director of microbial genome analysis at the Broad Institute of MIT and Harvard. "But we're starting to open it up. We're collecting the data to map the innards of TB and using that to create predictive models."

2 Combining Metabolic Models with Gene Regulatory Models to Get At Latency

Researchers would like to understand how TB survives its veiled existence in the lung. "If we can better understand latency, it might change our minds about how to treat latency," Galagan says.

To understand latency, researchers must combine gene regulatory models with metabolic network models. "There are probably 1,000 papers on how to model regulatory networks and probably 1,000 on metabolic networks" says **Nathan Price**, **PhD**, of the Institute for Systems Biology in Seattle. "But there are very few that do them together in an integrated way."

In a paper published in *PLoS Computational Biology* in June 2011, McFadden's team explored changes in metabolites when *Mtb* is grown inside a host cell (a macrophage). Their method, called differential producibility analysis (DPA), uses a metabolic network to extract metabolic signals from transcriptome data, allowing a glimpse at which pathways *Mtb* is using inside the host cell, what substrates it is eating, and what products it is generating.

One of the most interesting results: According to the network model, growing the bacillus in a macrophage caused it to go quiet. It stopped making DNA, RNA and amino acids and focused on one job: rebuilding its cell wall. "It seems that the TB bacillus has realized it's inside a host immune cell whose job is to kill it," McFadden says. "So the *Mtb* hunkers down, shuts
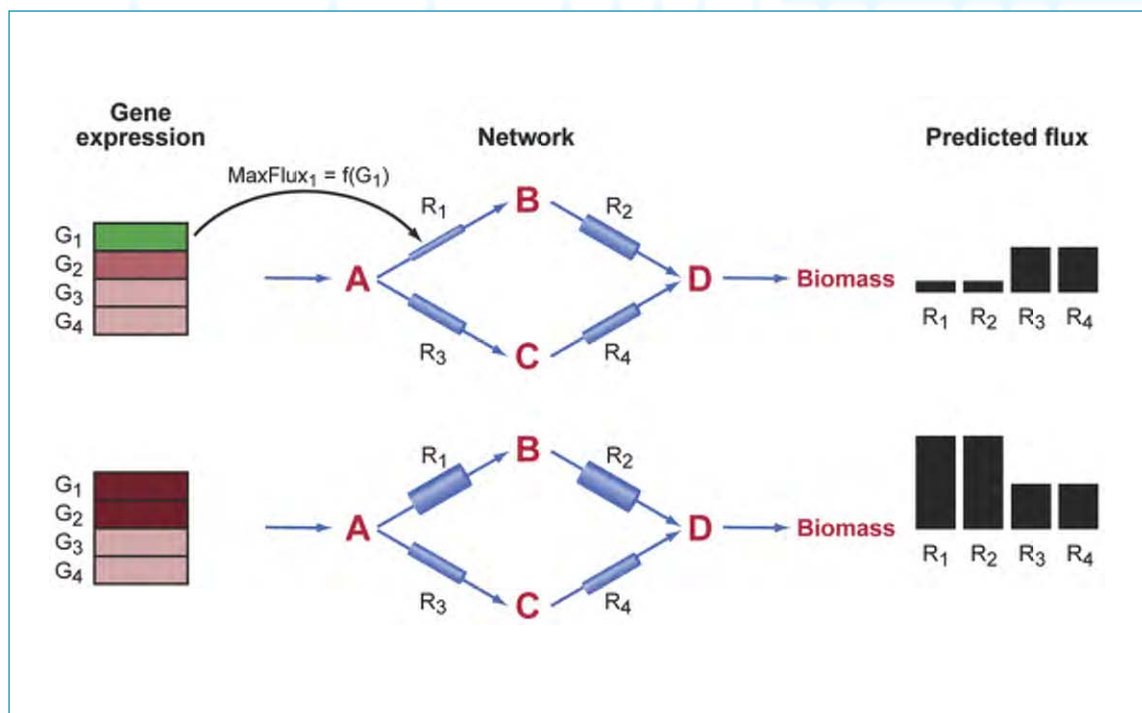
Galagan E-Flux. *E-flux uses gene expression to set maximum flux constraints on individual metabolic reactions. This can be illustrated as pipes of different widths around each reaction as illustrated here in a simple model of four metabolites (A–D), four internal reactions, an uptake reaction for A, and a reaction converting D to biomass. On the left are simulated gene expression data for four genes whose enzymes catalyze the four internal reactions (green = lower expression; red = higher expression). Where G1 is poorly expressed (top panel), a thin pipe is illustrated around reaction 1. In the bottom panel, G1 and G2 are highly expressed, corresponding to a wider pipe for these reactions. Under conditions in which uptake of A is not limiting, we would predict more flux through R1 and R2 in the bottom panel relative to the top panel and R3 and R4, as shown by the bars on the right. Reprinted from Colijn C, Brandes A, Zucker J, Lun DS, Weiner B, et al. (2009) Interpreting Expression Data with Metabolic Flux Models: Predicting* Mycobacterium tuberculosis *Mycolic Acid Production. PLoS Comput Biol 5(8): e1000489. doi:10.1371/journal.pcbi.1000489.*

down central processes and creates a more effective barrier for itself." Although this is not a novel insight—researchers already knew that *Mtb* shifts to cell wall building during latency—seeing it in a model was both impressive and instructive. Some of the details of the model might help researchers develop drugs to kill *Mtb* more easily, McFadden says.

Previously, when researchers have looked at transcriptomes, McFadden says, they've picked a favorite gene and looked at what that gene does. "It's like throwing a thousand stones in a pond and producing a thousand ripples but looking at just one of those ripples in an attempt to understand what's going on," he says. "What we do with DPA is look at all the ripples and put them together to get a picture of what's going on throughout the entire system."

Price and graduate student **Sriram Chandrasekaran** at the University of Illinois took a different approach to creating a unified model of *Mtb* gene regulation and metabolism. Their model, called PROM (probabilistic regulation of metabolism), was published in *Proceedings of the National Academy of Sciences* (*PNAS*) in 2010 and integrates work from Palsson's lab on metabolic networks.

PROM calculates the probability that expression of a particular transcription factor will result in the expression of a particular metabolic enzyme. These probabilities act as constraints on the metabolic model, like a dimmer switch. For example, researchers can ask how a knockout of a particular transcription factor will affect the abundance of metabolic enzymes and, in turn, the flux through a particular reaction. "We're trying to link changes in transcriptional regulation to what's going to happen in terms of a metabolic phenotype," Price says. And, like the metabolic model on its own, PROM was able to identify essential genes. "PROM picked up really well which transcription factors are essential to optimal growth in tuberculosis," Price says.

Another approach, called E-flux, brings gene expression together with metabolic models in yet another way. Developed by Galagan and his colleagues at the Broad Institute, E-flux uses gene expression data to constrain the metabolic model—essentially setting the width of the pipes leading to and from particular reactions in the metabolic network. Because the gene expression data comes from *Mtb* grown under a variety of conditions—including under exposure to 75 different substances and conditions such as hypoxia (which is akin to what the bug experiences in latency)—the results could help researchers understand how existing drugs work and identify other drugs that might also be effective.

In a 2009 publication in *PLoS Computational Biology*, Galagan's team applied E-flux to existing metabolic models of the *Mtb* mycolic acid pathway. Mycolic acids are good drug targets because they are critical components of the *Mtb* cell wall, do not exist in humans, and are the target of several existing antibiotics used to treat TB. E-flux predicted seven of eight known inhibitors of the mycolic acid pathway and identified several novel compounds not previously known to inhibit mycolic acid biosynthesis. The model also mimicked the ineffectiveness of first-line TB drugs against dormant tuberculosis.

But Galagan and his team have more ambitious goals for E-Flux: They want to improve the method so that it can make more refined interpretations of what *Mtb* is doing in latency—or at other key points in the disease process, Galagan says. His plans include building a metabolic model that isn't constrained to a steady state, by using mass spectrometry to measure metabolites directly.

In recent work with E-flux, his team observed—as McFadden did—that when TB goes dormant, most of the genes are ramped down. "But if you turn off the lights, things could go haywire," Galagan notes. "TB has to handle the process in a way that doesn't kill itself." His team observed that as *Mtb* scales down its activities, the bacterium makes a series of adaptations to the toxins that build up. "Those could be a weak link," he says. "Perhaps we could muck with that—target those processes that are important for *Mtb* not dying as it goes to sleep," Galagan says.

Ultimately, McFadden says, researchers need to link *Mtb* metabolic and regulatory models together with models that look at how *Mtb* and host cells interact. "That's really the challenge for the future," he says.

Ultimately, McFadden says, researchers need to link *Mtb* metabolic and regulatory models together with models that look at how *Mtb* and host cells interact. "That's really the challenge for the future," he says.

3 Multiscale Models of *Mtb*-Host Interactions

The granuloma, a spherical conglomeration of immune cells, bacteria, and tissue that walls off *Mtb* bacteria inside the human lung, offers another piece of the TB latency puzzle. How does the immune system force TB into an inactive state? Why do some people contain and wall off TB infections better than others? And why do
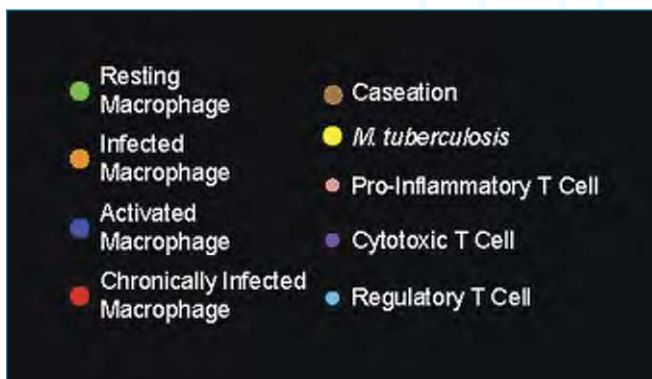
granulomas sometimes break down to reactivate TB infection?

Denise Kirschner, **PhD**, professor of microbiology and immunology at the University of Michigan Medical School, has been modeling the granuloma for about 11 years. Her team uses a multitude of host-pathogen response data gathered from granulomas in monkeys to build models at multiple scales: Ordinary differential equation models at the molecular scale link to an agent-based model at the cellular scale that reads out at the tissue scale.

Kirschner's models are stochastic, meaning they contain probabilities that certain events will or will not occur: One hundred simulations will generate one hundred different answers, simulating the sorts of variability one would find in human hosts. Kirschner refines the agent-based models until they produce outcomes that are relatively stable—akin to TB in its latent state. "Even with slight perturbations, the models still go to the same place in the end," she says. It then becomes possible to run *in silico* experiments that perturb the models. "We can look at the finest scale and say how it's impacting at the largest scale and vice versa," Kirschner says. And they can virtually "knock out" various parts of the host immune system instantaneously to see the effect on the granuloma. Ultimately, Kirschner would like to understand what brings *Mtb* out of the stable state—transitioning the disease from latency to reactivated disease. But for now, her model predictions are helping to focus the next series of animal experiments.

Since 2004 when she published her first model of the granuloma, Kirschner's team

Legend:
- Resting Macrophage
- Infected Macrophage
- Activated Macrophage
- Chronically Infected Macrophage
- Caseation
- *M. tuberculosis*
- Pro-Inflammatory T Cell
- Cytotoxic T Cell
- Regulatory T Cell

These six snapshots from a time-lapse simulation of a granuloma forming after infection with Mycobacterium tuberculosis *show a 2 mm by 2 mm slice of lung tissue. The simulation covers 200 days of infection dynamics (days 0, 50, 75, 100, 150, 200). Once the infected macrophage cell takes up bacteria and initiates infection, it begins to recruit additional immune cells that arrive via vascular sources distributed on the parenchyma. This simulation represents a controlled infection where the granuloma is able to physically contain and immunologically restrain the bacteria. Courtesy of Denise Kirschner.*
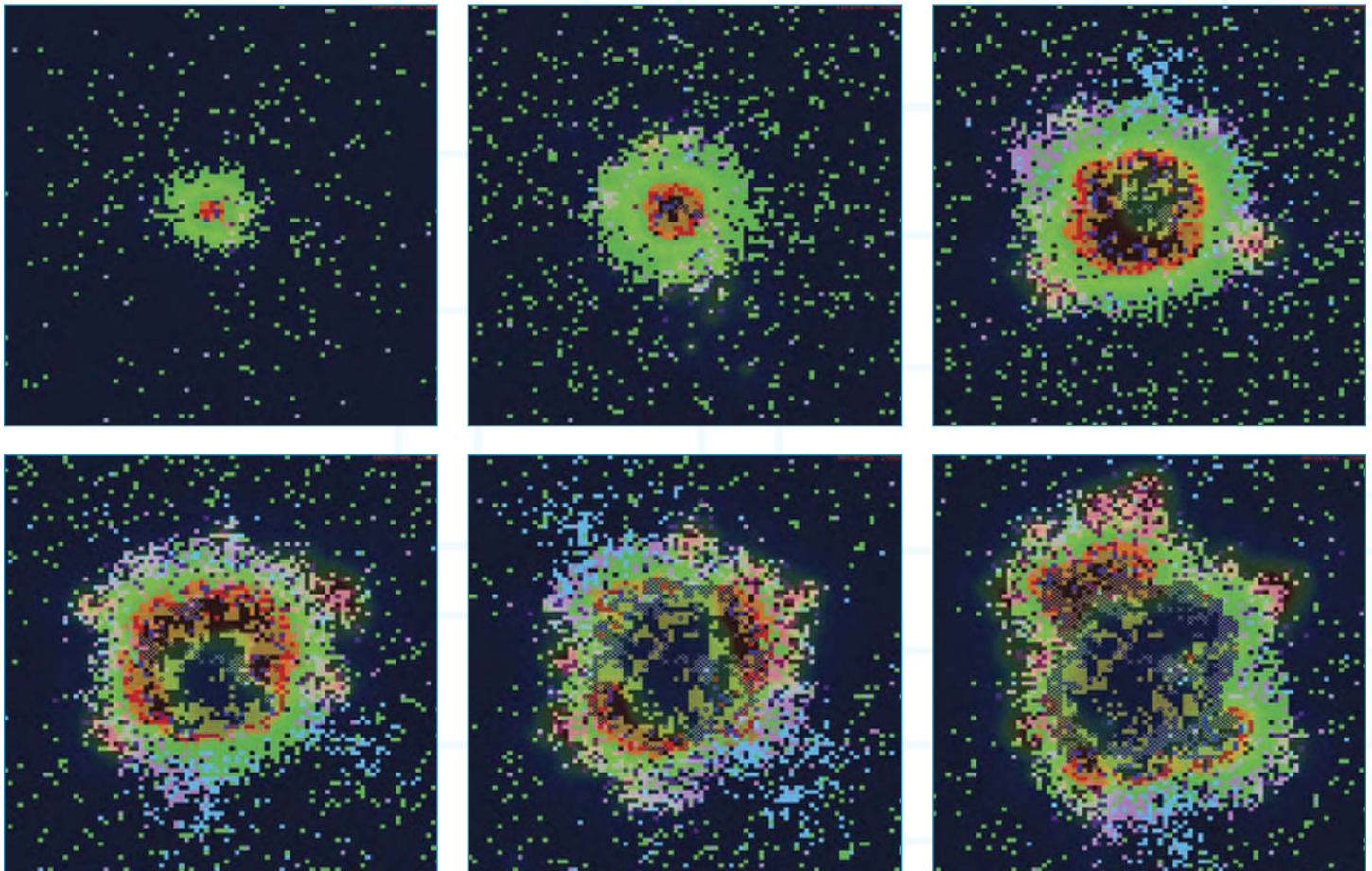
has refined and improved the models. They can now be run in 3-D and incorporate compartments outside the lungs, including the lymph nodes and blood. She's also taking the model to a new scale: populations. Working with a team in Italy, she says, they now have an agent-based model of TB epidemiology. The people in the population-scale model each have an immune-scale model of a granuloma running inside them, Kirschner says. These models predict how events at the smallest scales can influence epidemic outcomes, and thus can be used to test vaccine and treatment strategies.

Others studying the systems biology of TB latency and reactivation in the host are still at the early stages of their projects. For example, **Henry Boom**, **MD**, vice chair for research and director of the tuberculosis research unit at Case Western University Medical School, is looking at whether patients at different stages of progression from latent TB to active TB can

be distinguished by looking at protein-protein interaction sub-networks inside certain host immune cells. Watch for results in the future.

4 Protein-Protein Interactions and TB Drug Resistance

Emergence of drug-resistant TB strains is the biggest health challenge facing TB researchers. "Each time you administer a drug you are selecting for the organisms that are drug resistant," notes **Karthik Raman**, **PhD**, assistant professor of biotechnology at the Indian Institute of Technology Madras in Chennai, India.

While he was a graduate student in the lab of Dr. Nagasuma Chandra at the Indian Institute of Science, Raman and his colleagues used a systems biology approach to unravel the different mechanisms by which

TB drugs trigger resistance. His team created an *Mtb* protein-protein interaction network (from the STRING database of protein-protein interactions). They then merged that network with *Mtb* gene expression data gathered under exposure to seven different TB drugs, allowing an analysis of the possible routes leading to resistance.

One of their key aims was to identify what Raman calls "co-targets"—proteins that could be inhibited along with a primary drug target to reduce the likelihood that the *Mtb* bacillus will become drug resistant. "Some proteins were more important than others in terms of their strategic location in the network," Raman says. "Our hypothesis is that one could try to disable these proteins and not just the target proteins and it could probably reduce resistance."

Further work in this area is needed, Raman says. "It's probably an issue we'll never conquer." Yet developing novel strategies, such as the co-target concept, could help.

This portion of the TB protein-protein interaction network involved in drug resistance shows a tight cluster of cytochrome proteins (green nodes) and Rv0892 (blue node), a potential co-target that links the cytochrome clusters to proteins in the mycolic acid pathway. The nodes (individual proteins) are sized in proportion to the number of MAP drugs that induce their upregulation. The thickness of an edge is proportional to the number of times a shortest path is traversed through that edge. Reprinted with permission from Karthik Raman and Nagasuma Chandra, Mycobacterium tuberculosis *interactome analysis unravels potential pathways to drug resistance,* BMC Microbiology 8:234 (2008).

5 Computational Epidemiology and the Emergence of Drug Resistant TB

Computational models are also proving useful in exploring the population-level causes of drug resistant TB. Surprisingly, a recent statistical model shows that TB multiple drug resistance can evolve spontaneously—it is not necessarily caused by mono-therapy or by patients failing to complete a course of antibiotic treatment. This is a fundamental shift in our understanding of how combination drug resistance can emerge, says **Ted Cohen, MD, MPH, DPH,** assistant professor in epidemiology at Harvard University. "And it may help explain how highly drug resistant forms of TB have independently emerged in many settings."

Cohen is also working out the order in which drug resistance mutations occur and their probabilities of occurring. Similar work has proven helpful in understanding and treating HIV. But exploring the order of mutations over time typically requires longitudinal genotypic data—which is often not available for *Mtb*. So Cohen and his colleagues decided to determine whether this information could be inferred from phenotypic data (e.g., *in vitro* tests of drug resistance) gathered at one point in time. Using branching trees, a special kind of Bayesian network that makes it possible to infer past and future events, they were able to infer some possible patterns in which TB drug resistance phenotypes arise, Cohen says. "It's a promising approach and should be even more useful as genetic and genomic data becomes available and we can look not only at drug resistance phenotypes but also at the actual resistance-conferring mutations."

Cohen is also doing work in South Africa to try to understand the phenomenon of complex TB infections—multistrain infections that arise either from multiple infections by unrelated strains or from within-host evolution. He and his colleagues developed a modeling framework to investigate mechanisms of strain competition within hosts and to assess the long-term effects of such competition on the ecology of strains in a population. His initial modeling efforts suggest that the presence of mixed strains in a single host can increase the likelihood that drug-resistant strains will persist and potentially evolve.

"For me, modeling is a cycle," Cohen says. "We've used models to identify the gaps in our understanding of TB that most limit our ability to project trends or design effective interventions. Then we try to conduct studies that reduce this uncertainty so we can refine the models and improve our understanding of how best to intervene."

> "For me, modeling is a cycle," Cohen says. "We've used models to identify the gaps in our understanding of TB that most limit our ability to project trends or design effective interventions. Then we try to conduct studies that reduce this uncertainty so we can refine the models and improve our understanding of how best to intervene."

6 Using Computation to Find TB Drug Targets

The National Institute of Allergy and Infectious Diseases (NIAID), part of the National Institutes of Health, supports a portfolio of computational approaches for modeling TB drug targets and determining how promising drugs bind to these targets, says **Karen Lacourciere, PhD,** program officer for tuberculosis and other mycobacterial diseases at NIAID.

Computation can simplify the entire drug discovery process. For example, Raman developed a pipeline (TargetTB) for pinpointing which essential *Mtb* genes (and their protein products) are also good drug targets. The pipeline starts by identifying essential genes as predicted by both existing metabolic models (including McFadden's and Palsson's) and by experiments. Next, the pipeline filters out proteins with structural similarities to human proteins because targeting such proteins can produce negative side effects. This computationally intensive step involves exhaustive pairwise comparisons of several thousand pockets on more than 750 *Mtb* proteins with more than 70,000 sites on more than 15,000 human proteins. The pipeline filters the resulting short list using additional criteria and also prioritizes the proteins' importance based on their level of expression during TB latency. The work, published in *BMC Systems Biology* in 2008, also examines several known

> "Ten to twelve years back [TB] was considered a third-world disease, but with the appearance of drug resistant forms, pharmaceutical companies are showing more interest," Raman says.

and predicted drug targets based on their filters, and postulates why many known targets may produce adverse drug reactions.

About 400 potential drug targets have emerged from Raman's pipeline. "It's a beautiful place to start, those 400," Raman says. And pharmaceutical companies have shown some preliminary interest in pursuing these targets. "Ten to twelve years back it was considered a third-world disease, but with the appearance of drug resistant forms, pharmaceutical companies are showing more interest," he says.

McFadden's metabolic model is also being used to better understand known drug targets, says **Desmond Lun, PhD**, associate professor of computer science at Rutgers University. For example, Harvey Rubin at the University of Pennsylvania asked Lun to work out the mechanism of action of NDH-2. Rubin hypothesized that NDH-2 could be a lethal knockout in *Mtb* because removing the gene from a related bacterium leaves a non-viable organism. Using a metabolic model that consolidates two others, Lun identified a possible mechanism by which knocking out NDH-2 would be lethal in *Mtb*. "It takes a long time to develop drugs," Lun says, "so you want to know in significant detail what the mechanism is and the possible effect on the host and related organisms."

7 Using Computation to Find TB Drugs

Some chemists are screening for TB drugs using computational approaches, says **Sean Ekins, PhD**, vice president of science at Collaborative Drug Discovery and senior consultant for Collaborations in Chemistry. "People started by thinking we'd do structure-based screening but the molecules that looked good against the target didn't really have the right physical properties to get into the cell," he says. "TB is a pretty tough cookie in terms of the types of molecules it will let in."

Then a few years ago, researchers started whole cell screening—testing lots of compounds to see what would kill *Mtb*. This produced lists of thousands of lethal compounds but gave researchers no hints as to which ones merited further research, Ekins says. So Ekins decided to try a data-mining approach on an NIH database of more than 200,000 compounds with known activity against *Mtb*. The approach uses Bayesian machine learning to cherry-pick the compounds that have a high likelihood of whole cell activity and a low likelihood of

human toxicity, he says. "These models are a step in the right direction: We're leveraging all this data, using it to build models, and using the models to make future experimental decisions for our collaborators."

Indeed, Ekins has used the model to pick groups of compounds for various researchers to test. "That's sort of the acid test, really: backing up your predictions experimentally," he says. "I don't think we've got the magic bullet, but we're giving it a good try."

Using a different approach, **Lei Xie, PhD**, associate professor of computer science at Hunter College, the City University of New York, and **Philip Bourne, PhD**, professor of pharmacology at the University of California, San Diego, sought TB drugs among the pool of currently FDA-approved drugs by creating a genome-scale drug-target network they call the TB drugome. The network compares binding sites on *Mtb* proteins against the known binding targets of FDA-approved drugs (there are only 250). "If two proteins have a similar ligand-binding site,

then our assumption is that they can potentially bind a similar drug," says Xie. Next, they did protein-ligand docking to predict the binding affinity of the drugs to the *Mtb* proteins. The network revealed that about a third of the drugs examined have the potential to be repositioned to treat tuberculosis and that many currently unexploited *Mtb* receptors may be chemically druggable and could serve as novel anti-tubercular targets. The researchers are currently seeking collaborators to validate these findings.

8 Computation to Find Diagnostic Biomarkers

Since the 1940s, the primary diagnostic test for active TB has been the same: Take a sample of sputum (i.e., coughed up mucus) and look at it under a microscope. But the test is insensitive, producing many false-negative results. In addition it takes two
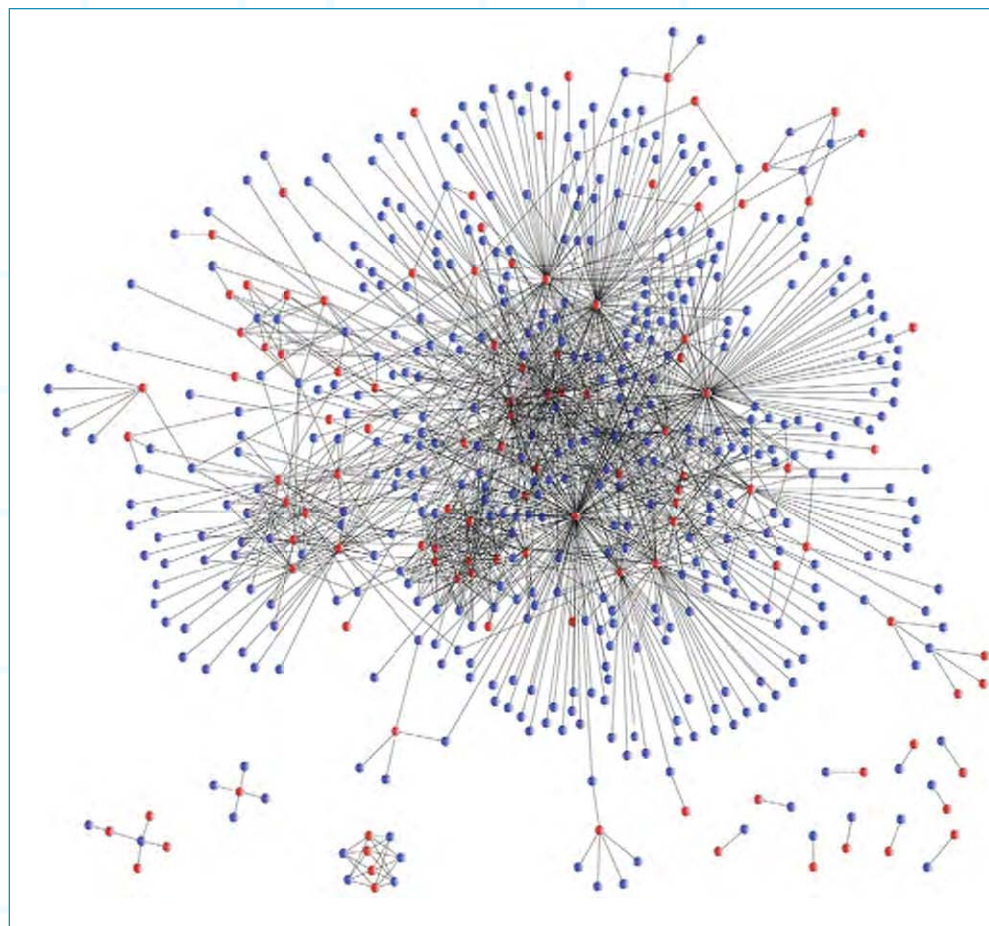
The TB Drugome, a protein-drug interaction network, shows Mtb *proteins (blue circles) connected to drugs (red circles), a single connection indicates binding site similarity between any of the structures of the connected* Mtb *protein, and any of the binding sites of the connected drug. This Drugome is highly connected, indicating that many binding site similarities were observed between* Mtb *proteins and drug targets, even though those proteins had different overall structures. Reprinted from Kinnings SL, et al., The Mycobacterium tuberculosis Drugome and Its Polypharmacological Implications,* PLoS Comput Biol *6(11): e1000976 (2010).*

weeks to get results to tests for drug resistance, risking further spread of TB's most virulent forms.

Although newer, more rapid TB diagnostic tests are now hitting the market, many of them require specialized labs and skills not available in areas most stricken by the disease. In addition, the rapid diagnosis of drug resistance remains elusive with one exception: A new test now being deployed in parts of Africa uses the GeneXpert system, a molecular assay that can identify strains of TB that are resistant to treatment with rifampin, a commonly used TB antibiotic. This test became possible because, for rifampin, we know the affected gene and the majority of mutations that result in resistance, says **James Posey**, **PhD**, a research microbiologist in the Center for Disease Control's Division of Tuberculosis Elimination. But for many of the other first- and second-line TB drugs, this information is unknown. Computational approaches are helping fill this gap; for example, Posey's lab is using whole genome sequencing to identify the genes and mutations that cause resistance to additional TB drugs.

Another group of researchers is searching for metabolic changes that could be used to rapidly flag drug-resistant mutants. Lun is working with **Greg Bisson**, **MD**, assistant professor of medicine at the University of Pennsylvania Medical School, on the novel hypothesis that, just as M*tb*'s ancestors—non-pathogenic soil bacteria—respond to assault by making new metabolites, perhaps TB does the same when confronted with drug treatment. Lun is perturbing a consolidated metabolic model using protein abundance data as a way to study that possibility. "This is something we can do to help develop a cheap and easy diagnostic test to work out whether someone has a drug resistant strain," Lun says. "This can make a major difference in public health outcomes."

9 Planning for the Clinic

Even after effective drugs and diagnostics for TB have been developed, deploying them in the clinic can be tricky. For example, when putting a new diagnostic test into action, doctors need to know who should be tested and whether the new tools should replace or complement existing ones. And the answers to those questions might depend on the setting—both the incidence of TB and the availability of resources. To explore these questions, Cohen and his colleagues

"There is a role for modeling to inform local, immediate, real world-type decisions while taking into account detailed knowledge

combined an epidemiological model of TB spread with a health system model. The work shows that, in concept, one can begin to evaluate the operational impact of a diagnostic tool using information not only about how the bug spreads but also about the logistical characteristics of the healthcare system. "That's something many people have simplified out of the problem," he says. "There is a role for modeling to inform local, immediate, real world–type decisions while taking into account detailed knowledge of local conditions." He's now working with others to build on this initial work to look at the potential effects of specific diagnostic tools such as the Gene Xpert system.

There are now at least nine drugs and 12 vaccines in the clinical research pipeline. If they prove medically effective, similar sorts of deployment models will be essential to their ultimate impact.

10 Setting the Clinical Research Agenda

About 10 years ago, the Bill and Melinda Gates Foundation began investing heavily in a tuberculosis research portfolio that included the development of drugs, vaccines and diagnostic tests. Having set that agenda, they then wondered: If we achieve our goals, what will TB morbidity and mortality look like in 2050? So they hired a team headed by **Elizabeth Halloran**, **MD**, **DSc**, professor of biostatistics at the University of Washington and the Hutchinson Research Center, to create a model showing the long-term effect of a successful program.

The work, published in *PNAS* in 2009, produced several interesting insights that could affect funding decisions. For example, introducing a new vaccine for infants would have very little effect by 2050 because people get TB when they are older. But a program of mass vaccinations is much more effective. "So one has to rethink vaccination strategies and clinical trials," she says. The model also showed that finding and curing latent infections—which we currently don't know how to do—would have a very large effect. Halloran notes: "That might affect how you allocate research resources." For example, this might suggest the wisdom of funding systems biology studies of TB latency, which brings us full circle.

TB The Opportunities Are Many

Although these intersections between computation and TB might suggest the field is pretty well picked over, that is not at all the case. The systems models need refinement and must be layered together with other models to gain a multiscale picture of the bug. Host-pathogen interaction research is really in its infancy. Researchers don't know the extent of MDR and XDR TB, let alone how to deal with it. And though there are multiple new diagnostics, drugs and vaccines in the pipeline, no one really knows how to implement them so that they will have the greatest impact.

Thus, each of these intersections between TB and computation suggests more that can be done. Perhaps because of its complexity, modelers haven't flocked to TB research, Cohen says, but he believes that will change: "If you want to ask questions that have global impact to improve the lot of humanity," Cohen says, "I think TB is a great thing to choose to work on." □

"If you want to ask questions that have global impact to improve the lot of humanity," Cohen says, "I think TB is a great thing to choose to work on."