Nature Reviews Microbiology | AOP, published online 9 June 2008; doi:10.1038/nrmicro1919

SYSTEMS MICROBIOLOGY

Systems biology of persistent infection: tuberculosis as a case study

Douglas Young[*‡], *Jaroslav Stark*[*§] *and Denise Kirschner*[||]

Abstract | The human immune response does an excellent job of clearing most of the pathogens that we encounter throughout our lives. However, some pathogens persist for the lifetime of the host. Despite many years of research, scientists have yet to determine the basis of persistence of most pathogens, and have therefore struggled to develop reliable prevention and treatment strategies. Systems biology provides a new and integrative tool that will help to achieve these goals. In this article, we use *Mycobacterium tuberculosis* as an example of how systems-biology approaches have begun to make strides in uncovering important facets of the host–pathogen interaction.

Infections that evolve over the course of prolonged persistent interactions between the host and the pathogen present a major challenge for disease control. Important advances have been made in the control of a range of acute infections using interventions that target a single biological focal point — neutralizing diphtheria toxin, for example, or enhancing immune recognition of the capsule of meningitis-causing bacteria — but it is more difficult to predict the impact of a targeted intervention on the complex biology of a persistent infection. Infectious diseases reflect an equilibrium between the host and the pathogen that is established and maintained by a broad network of interactions that occur across scales that range from molecular to cellular, to whole organism and population levels. Maintenance and evolution of these interactions over a prolonged time frame adds further complexity to persistent infections[1] (BOX 1). Experimental approaches that are applied at each of these individual levels generate dense islands of information — for example, in terms of pathogen genome sequences or the global transcriptional response of an infected cell — but conventional approaches cannot integrate information across scales and systems. For the persistent infections with the greatest global health impact, such as HIV-1/AIDS, tuberculosis (TB) and malaria, the ability to understand the interplay of various host–pathogen interactions across different spatial and temporal scales will be of considerable assistance in the rational design of improved tools for disease control and their rational implementation.

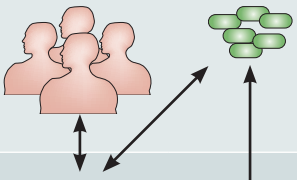
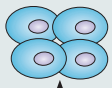
In this Review, we consider how systems biology can contribute to the challenges that are involved in studying persistent infection, and focus on TB as a specific example. Systems biology is an approach to understanding, explaining and predicting biological phenomena that arise from the dynamic interactions of more than one component. These components could be molecules, cells, organs or whole organisms[2]. The systems-biology framework combines mathematical modelling and simulation to complement traditional empirical and experimental approaches to biomedical research. These models and simulations are driven by empirical observations and generate specific, explicitly testable predictions that enable refinement of the models in response to experimental validation. This iterative development of models and experiments is a crucial feature of a systems-biology research approach[2,3].

There are two ways in which a systems-biology approach can be used to address persistent infections. First, at a single biological scale, current whole-genome technologies produce datasets that far exceed the analytical capacity of traditional reductionist reasoning. By constructing, validating and analysing mathematical and computational models, systems biology offers an opportunity to identify key networks of interactions, to suggest their functional properties and to predict the most informative sets of future experiments. A second, and more challenging, role for systems biology is to exploit the common language that is inherent in mathematical formulations to forge links between models that reflect different scales, thereby allowing us to explore how properties at one scale affect phenomena at other scales. For example, the rational design of a new antimicrobial requires understanding at the molecular level of drug action against a microbial target; at the host level of drug distribution and pharmacokinetics; and at the population level of the effect of resistance mutations on microbial

*Centre for Integrative Systems Biology at Imperial College (CISBIC), [§]Department of Mathematics, Imperial College London, London SW7 2AZ, UK.
[‡]Division of Mycobacterial Research, National Institute for Medical Research, London NW7 1AA, UK.
[||]Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan, USA.
Correspondence to D.Y.
e-mail:
d.young@imperial.ac.uk
doi:10.1038/nrmicro1919
Published online 9 June 2008

Box 1 | **Scales of *Mycobacterium tuberculosis* infection**

The outcome of infection with *M. tuberculosis* is determined by interactions that occur over various biological scales that range from molecular and cellular to anatomical and population levels. To formulate an integrated biological understanding of tuberculosis, we need to be able to assess how interactions at one level affect interactions that occur at each of the other levels. For example, how is a change in the sequence of a gene that encodes a drug target in *M. tuberculosis* manifested at the population scale? Mathematical models that arise from systems-biology approaches offer a unique potential to establish quantitative links across multiple biological scales. Different mathematical systems capture biological complexity best at individual scales. Continuous modelling: variables (for example, concentration) in the model are tracked in a continuous manner; discrete modelling: variables (for example, cell number) in the model are tracked discretely; deterministic modelling: dynamics of the model system are completely determined by the input; stochastic modelling: dynamics are not determined, but variability and unpredictable outcomes may arise.

	Host	Pathogen	Scale (time)	Scale (length)	Examples of modelling (dynamics)	Examples of modelling (model type)
Population			Days–years	10^{0}–10^{3} m	Deterministic	Systems of ODEs
Body			10^{5}–10^{6} s	10^{-2}–10^{0} m	Deterministic	Hybrid: ODE and ABM
Tissue			10^{4}–10^{5} s	10^{-5}–10^{-2} m	Stochastic and discrete	Algorithmic: ABM
Single cell			10^{1}–10^{3} s	10^{-5} m	Deterministic and continuous	Mathematical: ODE
Molecule			10^{1}–10^{2} s	10^{-9}–10^{-8} m	Deterministic and continuous	Statistical

ABM, agent-based model; ODE, ordinary differential equation.

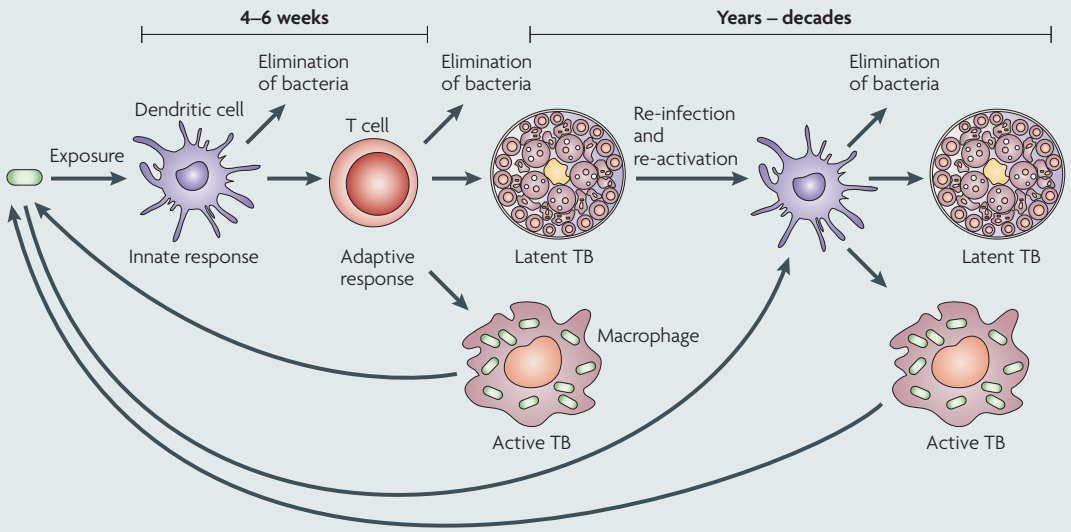
fitness in immunocompetent and immunodeficient hosts. It is impossible to construct a conventional biological framework that encompasses these widely separated disciplines. We would therefore argue that integrative systems biology, particularly when applied across several scales, is the most promising approach for breakthroughs in our understanding of persistent infections.

The interaction between the host and the pathogen occurs on different scales. These range from molecular interactions, including, for example, the recognition of specific molecular patterns on innate immune cells by Toll-like receptors, to interactions between individual cells, which, in turn, can range from the phagocytosis of bacteria by macrophages to the spread of disease through a host population and the emergence of different strains of pathogens in response to different host conditions. Elements at different scales interact, which can determine, for example, the relative success of the pathogen at the population level. This interaction then determines selection pressures at the genetic level in both the host and the pathogen, which are expressed in terms of often subtle changes to molecular interactions. At each biological scale, the appropriate mathematical or computer model to be constructed is chosen. This decision is based, in part, on the dynamics that are under consideration (BOX 1). Mathematical models use equations or simulations to describe biological events, and these are then typically solved using a computer. By contrast, computational models are applied using a computer, as a detailed sequence of rules that are implemented directly in object-oriented programming languages. Hybrid models that use both equations and computational systems are possible. Several excellent texts that describe the use of each of these models in different areas of biology are available[4–10], as well as texts that describe statistical model analysis of biological systems[11–13] and computer-based models[8,14].

Both biological experiments and modelling efforts have been successful at elucidating the properties of a disease at any particular level, but a full understanding requires the integration of all scales. This is a major challenge for systems biology.

In this Review, we explore this approach using TB as a paradigm example of a persistent infection[15]. In addition to the spatial scales outlined in BOX 1, TB highlights

Box 2 | **Temporal scales in tuberculosis infection**



Compared with many other diseases, the timescales involved in tuberculosis (TB) are long and there is large variation between different individuals. *Mycobacterium tuberculosis* is transmitted by aerosol from individuals with active disease. Bacteria that reach the alveoli of the lung are ingested by macrophages, where they can initiate rounds of intracellular replication and cell lysis (see the figure). Macrophages are key effector cells in mycobacterial killing, but can also provide a niche for bacterial multiplication. Dendritic cells engulf bacteria, or bacterial components, circulate to the draining lymph nodes and prime T cells, which then return to the lungs to orchestrate control of the infection[100]. T cells enhance the antibacterial activity of macrophages by releasing cytokines, such as interferon-γ, which generally results in arrest or clearance of the infection. If the T-cell response is insufficient to control the initial infection, clinical symptoms will develop within ~1 year in the form of primary progressive disease. Prior vaccination with bacille Calmette–Guérin (BCG), a live attenuated strain that is closely related to *M. tuberculosis*, establishes a primed population of T cells and reduces severe primary disease in children. Most individuals develop a T-cell response in the absence of any clinical symptoms, which is defined as a latent infection and carries a risk of secondary disease owing to subsequent reinfection or reactivation of the initial infection. Autopsy studies show that latent infection is often associated with persistence of viable bacteria[15]. Bacteria can persist within granulomas (see latent TB in the figure) that function to contain bacterial spread. In adult pulmonary TB, breakdown of granulomas in the lung promotes mycobacterial replication, release of bacteria into the airways and effective aerosol transmission. Transmission is enhanced by the destruction of lung tissue, which is mediated by the same immune cells that are crucial for protection during the earlier stages of infection.

the importance of temporal scales (BOX 2). Signalling pathways that are triggered within the first few minutes of the encounter of macrophages with *Mycobacterium tuberculosis* are crucial in determining the intracellular fate of the bacilli, and the release of cytokines from dendritic cells (DCs) over the first few days is important in programming subsequent T-cell responses. Primed T cells migrate to the lung over a period of weeks. As indicated in BOX 2, pathogenesis and disease evolve over a timescale of years or decades. Attempting to link events over such timescales with molecular and cellular events that occur in minutes or hours cannot be achieved by conventional microbiological or immunological experimental approaches, particularly considering the different spatial scales that are involved and problems in integrating spatial and temporal scales. To surmount such challenges, a multi-scale systems-biology approach is the most feasible.

In this article, we describe events at various scales that have key roles in determining the outcome of *M. tuberculosis* infection. We briefly highlight some of the systems-biology modelling that has been carried out at individual scales, but focus mainly on attempts to integrate across scales. We also discuss how these attempts have already contributed to uncovering important facets of the host–pathogen interaction and indicate the most important directions for future research.

A tour of the scales

As indicated in BOX 1, all of the spatial scales of the host–pathogen interaction are interlinked and it is not easy to collapse the complex network of interactions into a logical linear narrative. The scale at which one starts the description is also somewhat arbitrary. From the perspective of human health, two key objectives serve as a guiding principle.

First, at the host population level, we aim to predict the epidemiological effects of various treatments, including drugs, vaccines and public health measures (such as the isolation of infected individuals, a common strategy in the past for TB), and predict the spread of new emerging strains (perhaps drug resistant) into populations with particular characteristics.

Second, although a good epidemiological model might accurately predict the statistical likelihood of an

individual becoming infected with TB, we aim to predict the outcome of the disease at the individual level, and therefore prescribe the best possible treatment for that individual.

The interplay between these two objectives highlights a basic theme that will be apparent throughout this Review. At the epidemiological level, we can initially treat the population as homogeneous, using averaged population properties for the host, the pathogen and their interaction. However, such homogeneous models often provide a poor fit to observed data and make poor predictions. This is because individual variation is important and the interactions between different sub-groups of both hosts and pathogens can lead to significantly different outcomes compared with a homogeneous model. Such heterogeneity, in turn, affects the overall population behaviour. In attempting to obtain more accurate population-level models, we are therefore inevitably drawn towards understanding host–pathogen interactions at the individual level — that is, in reality, the host and pathogen populations only interact via individuals (BOX 1). To explore this further, we first describe modelling at the population level.

Population level. Interest in population-level modelling long precedes the recent enthusiasm for systems biology, and indeed represents one of the earliest applications of mathematics in biology. Fibonacci set a calculation of the exponential growth of a dividing population as an exercise in the Liber Abaci in 1202 (REF. 16). Although this calculation was applied specifically to breeding rabbits, the same model can also describe the initial unchecked growth of an epidemic or of a population of bacteria. Daniel Bernoulli used a modelling approach to influence public health policy for smallpox in 1760 (REFS 17–19), and models of the dynamics of epidemics began to be developed in the mid-nineteenth century[20]. Of particular note is the work of Ronald Ross, who was interested in why some diseases, such as cholera, produce rapid epidemics followed by periods in which they almost disappear, whereas others, such as malaria, can persist indefinitely in the population. He developed a mathematical model based on a set of ordinary differential equations that described the dynamics of the number of infected individuals in a population[21] and, together with Hilda Hudson, analysed their dynamic properties under various conditions[22,23]. This approach was refined and extended by Kermack and McKendrick[24–26] into the now-fundamental SIR (susceptible–infected–recovered) model, which is at the root of almost all existing models of the dynamic spread of disease through a finite population.

For TB, it is important to divide the infected group into latently infected and actively infected individuals, as these contribute particularly differently to the spread of the disease[27–29]. Additional categories, such as infectious and non-infectious, which are distinguished by the presence or absence of bacteria in the sputum, can also be introduced to provide more refined models or represent potential treatments[27,30–32]. By computing the frequency at which patients with active disease transmit infection to the susceptible population, these models can be used to predict the impact of different interventions on the overall dynamics of disease. It is predicted for TB, for example, that enrolment of 70% of the most infectious patients in a chemotherapy programme with an 85% cure rate would affect R_o (the number of secondary cases arising from each primary case) and reduce the epidemic[27,33]. Modelling can similarly be applied to predict the effect of different vaccination approaches and assess the potential impact of combining treatment and prevention[34,35]. These models provide the cornerstone of global TB control programmes (see Further information for a link to The Global Plan to Stop TB 2006–2015).

From population to individual: the effect of heterogeneities. The standard SIR-type models described above assume that all individuals are identical. However, it is well known that population heterogeneities can have an important role. Contrary to the assumptions of an SIR-type model, an infected individual is not equally likely to come into contact with every susceptible individual in the population. This has led to models that more accurately reflect the spatial and social organization of the host population. The value of these models was illustrated by the foot-and-mouth epidemic in the United Kingdom and the subsequent debate about different vaccination strategies[36–38]. Such models made use of comprehensive farm-by-farm data on the spread of the epidemic. For TB, a number of recent models have incorporated household structure by incorporating different transmission rates within and between households of different sizes[39,40]. The resulting network models more accurately reflect transmission patterns and the prevalence of TB in response to changing social organization[41].

It is also essential to consider the influence of biological heterogeneities, in both the host and the pathogen. An example of heterogeneity in the host was provided by classical twin studies which showed that human genetic diversity has a major influence on susceptibility to TB[42]. Two types of study have provided insights into the genetic control of TB. The first approach, which focused on rare mutations that confer a Mendelian form of hypersusceptibility to mycobacterial infection, demonstrated an essential role for genes that are involved in the T-helper-1 (T_H1)-type pathway for macrophage activation[43]. The second approach, which was based on case-control studies of candidate genes and genome-wide linkage analysis of affected siblings, identified much smaller effects of more common polymorphisms in genes that were implicated in microbial recognition, phagosome biology and antigen presentation[44–50]. The weak effects that were associated with individual polymorphisms suggest that the profound effects observed in twin studies arise from combinations of genes. Inclusion of population heterogeneity in the basic SIR model predicts important changes in transmission parameters that may influence strategies for optimal disease control[28,31]. This is an important example of how systems biology can provide a connection between research at the scale of cellular immunology with population-biology research.

Ordinary differential equation
A system of equations that is based on the rates (derivatives) of change of dependent variables with respect to time. Most of the interesting differential equations are nonlinear and, with a few exceptions, cannot be solved exactly. Approximate solutions are determined using computer simulations.

Mendelian
Genetic inheritance of disease susceptibility through a single gene.

T_H1
After priming by exposure to signals from antigen-presenting cells, T cells undergo a process of maturation to their final effector phenotype. Cytokines produced by T_H1 cells (for example, interferon-γ) enhance the antimicrobial activity of macrophages and have an important role in protection against *Mycobacterium tuberculosis*. Cytokines produced by T_H2 cells (for example, interleukin-4) are important in promoting antibody responses. Cells that have not committed to the T_H1 or T_H2 lineages are referred to as T_H0.

Linkage analysis
A test for co-inheritance of genetic markers along with disease susceptibility in family groups.

Box 3 | **Within-host modelling of tuberculosis**

Mathematical models have been developed that track the course of infection within a single host. These models typically track numbers of bacterial and host cell populations, as well as signalling molecules (for example, cytokines). Here, we present a series of models that have been developed over the past decade in our laboratories that explore host dynamics of infection. These models serve as a paradigm for the way in which such models can be approached and how they evolve, and highlight the importance of an iterative interaction with experimental observations.

The starting point was a simple model of a mycobacteria–macrophage interaction in the lung[101]. This model tracked the evolution over time of the numbers of three different sub-populations of macrophages, bacterial sub-populations, three T-helper-cell populations and four key cytokines, and predicted that latency and active disease, as well as clearance, can be observed under different host conditions. The model was extended to explore the role of $CD8^+$ T cells by mimicking experimental protocols, such as depletion and deletion of specific cytokines and cell types[102]. These studies predicted a different contribution for effector $CD8^+$ T cells that are cytotoxic compared with those that produce interferon-γ, and suggested that a minimum level of effector memory cells of each T-cell subset ($CD4^+$ and $CD8^+$) is required to provide protection following vaccination. This type of analysis is only available through *in silico* studies, as the level of detail that is needed to perform these experiments is not currently tractable.

A further extension allowed study of the role of tumour-necrosis factor α (TNFα) in protection during infection[103]. The model simulated two commonly used anti-inflammatory therapies that antagonize TNFα, which allowed us to study why those therapies lead to different tuberculosis reactivation rates. This model predicted that only a small proportion (<2%) of TNFα is needed to control acute infection and maintain latency.

Equally important is genetic variation in the pathogen population. Advances in genome sequencing and related technologies have uncovered an unexpected diversity in the global phylogeny of *M. tuberculosis*. Initial interest focused on a series of rapidly changing genetic markers, such as the movement of the IS*6110* insertion sequence[51]. These markers can facilitate tracking of local transmission chains, which could be used to refine the network transmission models referred to above[39,40,52,53]. At a deeper level of phylogeny, six major lineages of *M. tuberculosis* have been identified, which has prompted speculation about their historical and geographical relationships and the possible contribution of genotypic variation to the biology of disease[54]. These can be envisaged as a set of 'ancient' African and Indo–Oceanic lineages that are being progressively displaced by more aggressive 'modern' strains represented by the Euro–American, North Indian and Beijing lineages. One attractive hypothesis is that strains differ in their tendency towards primary disease relative to latent infection and reactivation (BOX 2). Models such as those described above could be used to explore the impact of such phenotypic differences on the spread of TB in populations with different densities and life histories. This, in turn, might help to identify important differences in selective pressures under which different lineages evolved. In the future, it will be important to link models of changing bacterial phylogeny with host population models to understand how genetic and environmental changes in both populations interact with each other. This should lead to a better understanding of which population groups are the most susceptible to which pathogen strains[55]. An understanding of the phenotypic properties of different strains and lineages is also central to modelling the spread of drug-resistant organisms[56].

Although TB and HIV-1 co-infection is of global concern, only a few models have been introduced that address the dynamics of these two diseases simultaneously. At the tissue scale, Kirschner[57] developed a cellular model that described the co-infection of HIV-1 and TB and predicted treatment effects. Naresh and Tripathi[58] developed a model that was based on the population being divided into four sub-classes and then studied the transmission dynamics of HIV-1 in settings in which TB infection is treatable. West and Thompson[59] performed numerical simulations to predict the future transmission trends of TB, whereas Porco, Small and Blower[60] developed a model that predicted the impact of HIV-1 infection on TB outbreaks. Some of these studies also examined treatment[27,30,31].

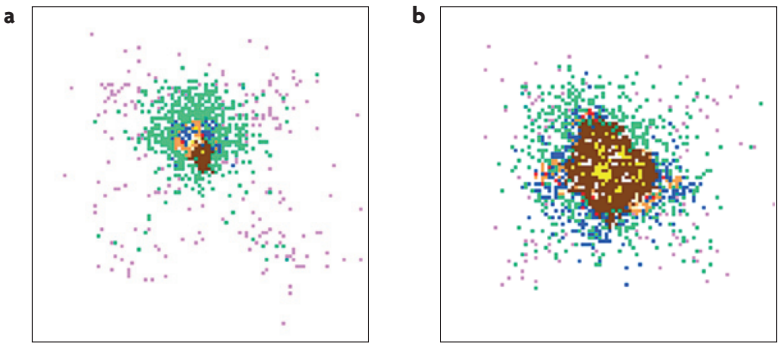
Moving down the scales: the immune system. The above discussion emphasizes how, to accurately predict population-level epidemic behaviour, we need to start incorporating aspects of individual behaviour. So far, this has been encapsulated through a few summary parameters, such as an individual's susceptibility to infection or propensity towards active disease; ultimately, however, we need to link these parameters with information at the cellular and genetic levels. The first step in this direction is to model the dynamic progression of infection within a single host.

TB can affect any organ of the body, and progressive primary disease can occur at extra-pulmonary sites. The lungs of an individual patient typically contain a diversity of lesions with differing overall structures that change over time. Correlating spatio–temporal variations with breakdown or resolution of lesions is crucial to our understanding of the disease process. However, experimental studies are predominantly restricted to measuring immune parameters in the blood, and there is a problem in reconciling systemic observations with local anatomical diversity. The use of models can help overcome these difficulties, and even the simple models that are currently available already yield important insights (BOX 3).

An important extension to these simple models is the incorporation of a spatial component by tracking cell populations and cytokines separately in the lung and draining lymph nodes[61,62]. This approach has been used to explore DC trafficking between the two compartments and the role of antigen presentation. The combination of such modelling approaches with data from modern non-invasive imaging techniques[63] provides a powerful framework to study the whole-body dynamics of TB.

From immune system to cell: structured population models. The models described in BOX 3 make the assumption that each cell population is homogeneous. For example, all T_H0 cells are assumed to have the same properties and to differentiate into T_H1 or T_H2 cells at the same rate. In reality, however, any such cell population will exhibit variation — for example, in the exact expression levels of various genes and the levels of various metabolites. Similarly, individual bacteria probably

Box 4 | **Agent-based model (ABM) of granuloma formation**



Our two-dimensional model[104] simulates both the spatial and temporal events that are relevant to the formation and maintenance of a granuloma. Cell types and known biological rules of how these cells interact and behave (that is, rules such as proliferation, recruitment, effector actions and death) were used. The environment (lattice) represented 2 mm x 2 mm of lung tissue, which is approximately the average size of a granuloma, and vascular sources, where cells can enter lung tissue, were also included. As agents (see the figure), we included T cells (pink), and macrophages that are resting (green), activated (blue), infected (orange) or chronically infected (red). Bacteria are shown in yellow and caseum in brown. The model permits spatio-temporal simulations of cells, bacteria and chemokines.

The model predicted three distinct and robust infection outcomes: a granuloma that was tightly packed, small and showed little necrosis and that contained bacterial growth (see the figure, panel **a**); a granuloma that was larger and more diffuse with a much greater caseous area that failed to restrict bacterial dissemination (see the figure, panel **b**); and a granuloma that cleared the bacteria load altogether and then dispersed (not shown). The key question that arises is how are these different outcomes obtained? Using the sensitivity analysis techniques developed specifically for use with ABMs[105], we determined that of the more than 27 parameters (rules) that govern the interactions in our model, only 7 were statistically significant in determining these different granuloma structures. A major feature of the sensitivity analysis is that these significances can be determined over time, allowing us to determine which parameters are important early in infection dynamics (for example, at 30 days) and which have a role later in the dynamics (for example, at 500 days). This analysis suggested that both T cells and chemokines are important in determining granuloma outcome. See Further information for a link to the full simulations (Time Lapse Simulations of Agent Based Models).

vary in their replication rate and antigen expression in response to differing local microenvironments. This situation is the same as that discussed above for modelling host-population behaviour and, in the same way, we need to link models of the whole population of immune and pathogen cells to models of individual decision making within each cell.

One attractive approach is to use structured population models[64], which are generalizations of age-structured population models in ecology and epidemiology. Such models combine internal dynamic variables for each individual with a description of the interactions of a population of such individuals. An example is a multi-scale model that combined mutual inhibition of two key transcription factors with an earlier cell-population model[65] to describe T_H-cell differentiation[66]. This particular model did not incorporate pathogens or other cell types but, in principle, such a structured population framework would provide a powerful approach to linking models at different scales.

Markov chain
A discrete-time stochastic process with the property that the next state solely depends on the present state, but not on the previous states. If a sequence of states has the Markov property, then every future state is conditionally independent of every prior state.

From cell to tissue: agent-based modelling. In linking different levels, we have so far started with a model at the larger scale, and incorporated increasingly detailed information about individual behaviour at the smaller scale. An alternative is to model a population of discrete individuals, before using it to predict overall population characteristics at higher scales. This requires a different mathematical approach compared with the models we have considered so far, which have used continuous variables to describe numbers of individuals or cells, or concentrations of relevant signalling molecules, leading to models that use ordinary differential equations[67–70]. Such equations provide a powerful modelling framework and have the advantage that a huge suite of analytical tools are available. Thus, rather than just simulating such models, we can both fit them to data[71,72] and determine qualitative properties, such as stability and robustness[73].

Models that use continuous variables, however, are only appropriate if the number of entities (such as organisms, cells and molecules) with identical behaviour is large. They are unsuitable in situations in which overall numbers are small or if we need to model discrete individual behaviour. If such behaviour is simple, Markov chain models are typically used[72,74], which, if they incorporate space, take the form of cellular automata and their generalizations[75]. As the behaviour of the individual entities becomes more sophisticated, agent-based models (ABMs), which offer a powerful approach to integrating the behaviour of individuals with the next scale above, become more widely used. BOX 4 illustrates an ABM model of granuloma formation that links the behaviour of individual cells with that of tissue-level structure.

Molecular systems biology: from sequence to cell. A large proportion of current efforts in systems biology is focused on integrating genome-wide data on the abundance of various classes of molecules (for example, mRNAs, proteins and metabolites) into a coherent picture of cellular behaviour. In the context of this Review, this can be done for both the pathogen and various cell types in the host.

Transcriptome analysis using whole-genome microarrays has been performed for *M. tuberculosis* in a range of different culture conditions and during interactions with host macrophages[76] (see Further information for a link to the TB Database). The proteome of *M. tuberculosis* under different growth conditions has been characterized by two-dimensional gel analysis[77], although only limited information is available about protein–protein interactions. Furthermore, high-throughput transposon mutagenesis has been used to identify genes that are essential for growth in culture in macrophages and in mice[78–81]. Whereas all of the 'omics' datasets include some degree of experimental error, they provide a useful framework on which to build models of bacterial physiology[82]. Jamshidi and Palsson[83] have begun to develop an *in silico* strain of *M. tuberculosis* (*iNJ*661) by incorporating a partial metabolic map, and have used this to help identify novel drug targets. An alternative approach for the identification of drug targets that is based on differential weighting of information from multiple 'omics' datasets has been

incorporated into a web-based tool (*web*TB; see Further information) by the TB Structural Genomics Consortium. Other websites that are relevant to the modelling of *M. tuberculosis* at the molecular level include genome data (for example, the pathogen website of the Wellcome Trust Sanger Institute and the Comprehensive Microbial Resource of the J. Craig Venter Institute), proteomic data (for example, the Proteome 2D-PAGE Database) and information on mutant strains (for example, TARGET (Tuberculosis Animal Research and Gene Evaluation Taskforce) (see Further information).

Host–pathogen interactions at the cellular and molecular levels. Classical studies of *M. tuberculosis* infection of macrophages showed that live, but not killed, bacteria can inhibit the process by which internalized micro-organisms are exposed to the toxic environment of an acidic phagolysosome, which provides an attractive explanation for their ability to transform these cells from lethal enemies to a safe haven[84]. However, the observation that the presence of antibodies reverses this inhibition without affecting intracellular survival[85] suggests an additional level of complexity. Many subsequent studies have added detailed characterization of the immature phagosome that is occupied by mycobacterial pathogens[86]. This is a complex and stochastic process: mycobacteria are distributed between early and late phagosomal compartments and, under some conditions, can also become free in the cytoplasm[87]. The dynamics of mycobacterial killing and survival are substantially altered in macrophages that have undergone a maturation process that is triggered by a combination of microbial ligands and cytokines (mainly by IFN-γ that is delivered by natural killer cells in the early stage of infection and by T cells following the onset of adaptive immunity). The molecular mechanism of phagosomal arrest, as well as the mechanism and location of mycobacterial killing, remains to be clearly defined. Jordao *et al.*[88] have attempted to capture the complexity of the fate of the bacille Calmette–Guérin (BCG) vaccine in J774 macrophages in the form of a simple differential equation model that describes the distribution of live and dead organisms in early and late phagosomal compartments. Their careful analysis suggests cycles of killing and replication, with late phagosomes providing the location for rapid killing and clearance of bacteria.

Modelling at this level can also begin to unravel the ways in which the pathogen manipulates signalling by host cells, and microarray snapshots have been generated of the host cell response to infection[89]. Analysis of microarray time courses using ordinary differential-equation models is proving to be a powerful approach to unravelling innate immune responses[90]. For example, such analysis emphasized the role of the transcription factor ATF3 and its associated network of interactions in signalling through Toll-like receptors, and subsequently identified 92 transcription factors that have a role in this macrophage response[91]. This approach highlights the huge challenge in understanding such complex networks. Differences in pathogenesis among *M. tuberculosis* strains are likely to result, at least in part, from differences in innate immune recognition[92,93]. For example, the highly virulent phenotype expressed by a Beijing-family strain in a mouse model was associated with a difference in cytokine profile, which was linked to the expression of a phenolic glycolipid surface molecule[92].

M. tuberculosis also subverts adaptive immune responses by downregulating antigen presentation in macrophages. A mathematical model that tracks levels of various molecules, including complexes of peptide–class II major histocompatibility complex (MHC) on the cell surface, can allow the comparison of different hypotheses. This approach showed that mechanisms which target class II MHC expression are effective at inhibiting antigen presentation, but only after a delay of at least 10 hours after *M. tuberculosis* infection. By comparison, mechanisms that target other cellular processes have an immediate effect, but may be attenuated under certain conditions. Therefore, targeting multiple cellular processes might represent an optimal strategy by which *M. tuberculosis* maintains continuous inhibition of antigen-presentation events[94].

Conclusion: integrating scales

Persistent infections present a complex problem for scientific understanding and treatment at both the individual and population levels. Using TB as an example, we have shown how it is important to consider interactions at scales that range from the molecular to the population levels (BOX 1). Implicit to this problem is the need to integrate events that occur at widely disparate timescales, ranging from molecular and cellular interactions that occur in seconds or minutes to the potentially decade-long development of the disease. This presents a huge challenge, but also an exciting opportunity to develop fundamentally novel insights by establishing links between disparate areas of biology.

Multi-scale models have recently been applied to a number of tissues, including cardiac cells[95] and growing tumours[96], as well as to the field of tissue engineering[97]. Recent work in our laboratories[66,98,99] has begun to explore a multi-scale modelling approach that describes aspects of immunity, and we propose that this approach is well-suited for understanding the host–pathogen interaction in *M. tuberculosis* infection.

We have summarized a number of ways in which models are beginning to be used to link events at different scales. Typically, this involves the development of a model at a single level that considers a homogeneous population of individuals, cells or molecules, and then incorporates variation between individual members of this population. Given the diversity of modelling approaches that are available, such multi-scale integration presents a major challenge. Some of the models will be rich in experimental detail, whereas others will be highly speculative extrapolations from a sparse database. The success of the models should be judged by their ability to provide a logical framework that predicts the most productive areas for future experimental efforts, rather than their ability to include every detail of current knowledge.

Cellular automata
Discrete models that consist of a regular grid of cells, each of which has a finite number of states. The state of a cell at time t is a function of the states of a finite number of cells (called its neighbourhood) at time $t-1$. Every cell has the same rule for updating that is based on the values in this neighbourhood. Each time the rules are applied to the whole grid, a new generation is created.

Granuloma
A roughly spherical structure that comprises a focus of infection that is surrounded by immune cells. Dead cells at the centre of the granuloma may decompose, leaving a 'cheesy' residue that is referred to as caseum.

A key step in building a multi-scale model will be to identify points at which informative data can be relayed between models and to focus experimental attention on strengthening these links. Molecular models of mycobacterial gene expression have the potential to provide information about the availability of ligands that are involved in recognition by the innate and adaptive immune responses; for example, by providing important parameters that influence the dynamics of granuloma formation. Resulting differences in the local microenvironment of the pathogen, in turn, provide parameters that can be fed back into metabolic models to predict bacterial growth rates and susceptibility to drug- and immune-mediated killing.

In this Review, we have focused on TB as an example. The same themes, however, arise in any other persistent disease, such as HIV/AIDS or malaria. Each of these diseases would provide specific details about the interactions between scales (such as the complex life cycle of the malaria parasite) that would require the development of customized modelling approaches. Nevertheless, the basic principles highlighted here will have an equally important role in understanding other diseases. Ultimately, as experience with multi-scale integrative systems biology grows, we hope to identify generic techniques that lead to novel and efficient means of controlling these global diseases.

1. Blaser, M. J. & Kirschner, D. The equilibria that allow bacterial persistence in human hosts. *Nature* **449**, 843–849 (2007).
2. Kitano, H. Systems biology: a brief overview. *Science* **295**, 1662–1664 (2002).
3. Yates, A., Chan, C. C., Callard, R. E., George, A. J. & Stark, J. An approach to modelling in immunology. *Brief Bioinform.* **2**, 245–257 (2001).
4. Edelstein-Keshet, L. *Mathematical Models in Biology* (Random House, New York, 1988).
5. Murray, J. D. *Mathematical Biology* (Springer-Verlag, New York, 1989).
6. Keener, J. P. & Sneyd, J. *Mathematical Physiology* (Springer, New York, 1998).
7. Segel, L. A. & Cohen, I. R. *Design Principles For the Immune System and Other Distributed Autonomous Systems* (Oxford Univ. Press, 2001).
8. Grimm, V. & Railsback, S. F. *Individual-Based Modeling and Ecology* (Princeton Univ. Press, 2005).
9. Lauffenburger, D. A. & Linderman, J. L. *Receptors: Models for Binding, Trafficking and Signaling* (Oxford Univ. Press, 1993).
10. Brauer, F. & Castillo-Chāavez, C. *Mathematical Models in Population Biology and Epidemiology* (Springer, New York, 2001).
11. Armitage, P., Berry. G. & Matthews, J. N. S. *Statistical Methods in Medical Research* (ed. Malden, M. A.) (Blackwell Science, Oxford, 2001).
12. Lund, O. *Immunological Bioinformatics* (The MIT Press, Cambridge, Massachusettes, 2005).
13. Ewens, W. J. & Grant, G. R. *Statistical Methods in Bioinformatics: an Introduction* (Springer, New York, 2005).
14. DeAngelis, D. L. & Gross, L. J. *Individual-Based Models and Approaches in Ecology: Populations, Communities, and Ecosystems* (Chapman & Hall, New York, 1992).
15. Stewart, G. R., Robertson, B. D. & Young, D. B. Tuberculosis: a problem with persistence. *Nature Rev. Microbiol.* **1**, 97–105 (2003).
Review of the biology of persistent infection with *M. tuberculosis*.
16. Fibonacci, L. & Sigler, L. E. *Fibonacci's Liber Abaci: a Translation into Modern English of Leonardo Pisano's Book of Calculation* (Springer, New York, 2002).
17. Dietz, K. & Heesterbeek, J. A. Bernoulli was ahead of modern epidemiology. *Nature* **408**, 513–514 (2000).
18. Dietz, K. & Heesterbeek, J. A. Daniel Bernoulli's epidemiological model revisited. *Math. Biosci.* **180**, 1–21 (2002).
19. Blower, S. & Bernoulli, D. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. *Rev. Med. Virol.* **14**, 275–288 (2004).
20. Farr, W. On the Cattle Plague. *J. Soc. Sci.* **1**, 349–351 (1866).
21. Ross, R. An application of the theory of probabilities to the study of a priori pathometry. Part I. *Proc. R. Soc. Lond. A* **92**, 204–230 (1916).
22. Ross, R. & Hudson, H. P. An application of the theory of probabilities to the study of a priori pathometry. Part II. *Proc. R. Soc. Lond. B* **89**, 507 (1917).
23. Ross, R. & Hudson, H. P. An application of the theory of probabilities to the study of a priori pathometry. Part III. *Proc. R. Soc. Lond. B* **89**, 507 (1917).
24. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* **115**, 700–721 (1927).
25. Kermack, W. O. & McKendrick, A. G. Contributions to the mathematical theory of epidemics. II. The problem of endemicity. *Proc. R. Soc. Lond. A* **138**, 55–83 (1932).
26. Kermack, W. O. & McKendrick, A. G. Contributions to the mathematical theory of epidemics. III. Further studies of the problem of endemicity. *Proc. R. Soc. Lond. A* **141**, 94–122 (1933).
27. Blower, S. M. *et al.* The intrinsic transmission dynamics of tuberculosis epidemics. *Nature Med.* **1**, 815–821 (1995).
28. Murphy, B. M., Singer, B. H., Anderson, S. & Kirschner, D. Comparing epidemic tuberculosis in demographically distinct heterogeneous populations. *Math. Biosci.* **180**, 161–185 (2002).
29. Waaler, H., Geser, A. & Andersen, S. The use of mathematical models in the study of the epidemiology of tuberculosis. *Am. J. Public Health Nations Health* **52**, 1002–1013 (1962).
30. Blower, S. M., Small, P. M. & Hopewell, P. C. Control strategies for tuberculosis epidemics: new models for old problems. *Science* **273**, 497–500 (1996).
Developed a model for designing effective control strategies for TB that was used to assess how suboptimal programmes can contribute to the development of drug resistance.
31. Murphy, B. M., Singer, B. H. & Kirschner, D. On treatment of tuberculosis in heterogeneous populations. *J. Theor. Biol.* **223**, 391–404 (2003).
32. Dye, C., Garnett, G. P., Sleeman, K. & Williams, B. G. Prospects for worldwide tuberculosis control under the WHO DOTS strategy. Directly observed short-course therapy. *Lancet* **352**, 1886–1891 (1998).
33. Murray, C. J. & Salomon, J. A. Modeling the impact of global tuberculosis control strategies. *Proc. Natl Acad. Sci. USA* **95**, 13881–13886 (1998).
34. Young, D. & Dye, C. The development and impact of tuberculosis vaccines. *Cell* **124**, 683–687 (2006).
Used epidemiology modelling to predict the impact of the combined effects of drug treatment and vaccination on TB control.
35. Dye, C. & Williams, B. G. Eliminating human tuberculosis in the twenty-first century. *J. R. Soc. Interface* **5**, 653–662 (2008).
36. Keeling, M. J., Woolhouse, M. E., May, R. M., Davies, G. & Grenfell, B. T. Modelling vaccination strategies against foot-and-mouth disease. *Nature* **421**, 136–142 (2003).
37. Keeling, M. J. *et al.* Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science* **294**, 813–817 (2001).
38. Tildesley, M. J. *et al.* Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. *Nature* **440**, 83–86 (2006).
39. Aparicio, J. P., Capurro, A. F. & Castillo-Chavez, C. Transmission and dynamics of tuberculosis on generalized households. *J. Theor. Biol.* **206**, 327–341 (2000).
40. Colijn, C., Cohen, T. & Murray, M. Emergent heterogeneity in declining tuberculosis epidemics. *J. Theor. Biol.* **247**, 765–774 (2007).
41. Grassly, N. C. & Fraser, C. Mathematical models of infectious disease transmission. *Nature Rev. Microbiol.* **6**, 477–487 (2008).
42. Comstock, G. W. Tuberculosis in twins: a re-analysis of the Prophit survey. *Am. Rev. Respir. Dis.* **117**, 621–624 (1978).
43. Alcais, A., Fieschi, C., Abel, L. & Casanova, J. L. Tuberculosis in children and adults: two distinct genetic diseases. *J. Exp. Med.* **202**, 1617–1621 (2005).
44. Barreiro, L. B. *et al.* Promoter variation in the DC-SIGN-encoding gene CD209 is associated with tuberculosis. *PLoS Med.* **3**, e20 (2006).
45. Bellamy, R. *et al.* Tuberculosis and chronic hepatitis B virus infection in Africans and variation in the vitamin D receptor gene. *J. Infect. Dis.* **179**, 721–724 (1999).
46. Hoal-Van Helden, E. G. *et al.* Mannose-binding protein B allele confers protection against tuberculous meningitis. *Pediatr. Res.* **45**, 459–464 (1999).
47. Khor, C. C. *et al.* A Mal functional variant is associated with protection against invasive pneumococcal disease, bacteremia, malaria and tuberculosis. *Nature Genet.* **39**, 523–528 (2007).
48. Malik, S. *et al.* Alleles of the NRAMP1 gene are risk factors for pediatric tuberculosis disease. *Proc. Natl Acad. Sci. USA* **102**, 12183–12188 (2005).
49. Marquet, S. & Schurr, E. Genetics of susceptibility to infectious diseases: tuberculosis and leprosy as examples. *Drug Metab. Dispos.* **29**, 479–483 (2001).
50. Thuong, N. T. *et al.* A polymorphism in human TLR2 is associated with increased susceptibility to tuberculous meningitis. *Genes Immun.* **8**, 422–428 (2007).
51. Van Soolingen, D. Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements. *J. Intern. Med.* **249**, 1–26 (2001).
52. Tanaka, M. M. & Rosenberg, N. A. Optimal estimation of transposition rates of insertion sequences for molecular epidemiology. *Stat. Med.* **20**, 2409–2420 (2001).
53. Tanaka, M. M., Small, P. M., Salamon, H. & Feldman, M. W. The dynamics of repeated elements: applications to the epidemiology of tuberculosis. *Proc. Natl Acad. Sci. USA* **97**, 3532–3537 (2000).
54. Gagneux, S. & Small, P. M. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect. Dis.* **7**, 328–337 (2007).
55. Gagneux, S. *et al.* Variable host–pathogen compatibility in *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **103**, 2869–2873 (2006).
56. Cohen, T. & Murray, M. Modeling epidemics of multidrug-resistant *M. tuberculosis* of heterogeneous fitness. *Nature Med.* **10**, 1117–1121 (2004).
57. Kirschner, D. Dynamics of co-infection with *M. tuberculosis* and HIV-1. *Theor. Popul. Biol.* **55**, 94–109 (1999).
58. Naresh, R. & Tripathi, A. Modelling and analysis of HIV–TB co-infection in a variable size population. *Math. Model. Anal.* **10**, 275–286 (2005).
59. West, R. W. & Thompson, J. R. Modeling the impact of HIV on the spread of tuberculosis in the United States. *Math. Biosci.* **143**, 35–60 (1997).
60. Porco, T. C., Small, P. M. & Blower, S. M. Amplification dynamics: predicting the effect of HIV on tuberculosis outbreaks. *J. Acquir. Immune Defic. Syndr.* **28**, 437–444 (2001).
61. Marino, S. & Kirschner, D. E. The human immune response to *Mycobacterium tuberculosis* in lung and lymph node. *J. Theor. Biol.* **227**, 463–486 (2004).
62. Marino, S. *et al.* Dendritic cell trafficking and antigen presentation in the human immune response to *Mycobacterium tuberculosis*. *J. Immunol.* **173**, 494–506 (2004).

63. Megason, S. G. & Fraser, S. E. Imaging in systems biology. *Cell* **130**, 784–795 (2007).
64. Cushing, J. M. *An Introduction to Structured Population Dynamics* (Society Industrial and Applied Mathematics, Philadelphia, 1998).
65. Yates, A., Bergmann, C., Van Hemmen, J. L., Stark, J. & Callard, R. Cytokine-modulated regulation of helper T cell populations. *J. Theor. Biol.* **206**, 539–560 (2000).
66. Yates, A., Callard, R. & Stark, J. Combining cytokine signalling with T-bet and GATA-3 regulation in Th1 and Th2 differentiation: a model for cellular decision-making. *J. Theor. Biol.* **231**, 181–196 (2004).
Presented a modelling framework that showed how to integrate transcription-factor dynamics with cytokine signalling in a population of T cells.
67. Stark, J., Chan, C. & George, A. J. Oscillations in the immune system. *Immunol. Rev.* **216**, 213–231 (2007).
68. Tyson, J. J., Chen, K. C. & Novak, B. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr. Opin. Cell Biol.* **15**, 221–231 (2003).
69. Di Ventura, B., Lemerle, C., Michalodimitrakis, K. & Serrano, L. From *in vivo* to *in silico* biology and back. *Nature* **443**, 527–533 (2006).
70. Kholodenko, B. N. Cell-signalling dynamics in time and space. *Nature Rev. Mol. Cell Biol.* **7**, 165–176 (2006).
71. Brewer, D., Barenco, M., Callard, R., Hubank, M. & Stark, J. Fitting ordinary differential equations to short time course data. *Philos. Transact. A Math. Phys. Eng. Sci.* **366**, 519–544 (2008).
72. Jaqaman, K. & Danuser, G. Linking data to models: data regression. *Nature Rev. Mol. Cell Biol.* **7**, 813–819 (2006).
73. Kitano, H. Towards a theory of biological robustness. *Mol. Syst. Biol.* **3**, 137 (2007).
74. Bailey, N. T. J. *The Elements of Stochastic Processes with Applications to the Natural Sciences* (Wiley, New York, 1964).
75. Ermentrout, G. B. & Edelstein-Keshet, L. Cellular automata approaches to biological modeling. *J. Theor. Biol.* **160**, 97–133 (1993).
The first study to use stochastic-type models.
76. Schnappinger, D. *et al.* Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. *J. Exp. Med.* **198**, 693–704 (2003).
77. Pleissner, K. P. *et al.* Web-accessible proteome databases for microbial research. *Proteomics* **4**, 1305–1313 (2004).
78. Sassetti, C. M., Boyd, D. H. & Rubin, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. *Mol. Microbiol.* **48**, 77–84 (2003).
79. Sassetti, C. M. & Rubin, E. J. Genetic requirements for mycobacterial survival during infection. *Proc. Natl Acad. Sci. USA* **100**, 12989–12994 (2003).
80. Rengarajan, J., Bloom, B. R. & Rubin, E. J. Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages. *Proc. Natl Acad. Sci. USA* **102**, 8327–8332 (2005).
81. Stewart, G. R., Patel, J., Robertson, B. D., Rae, A. & Young, D. B. Mycobacterial mutants with defective control of phagosomal acidification. *PLoS Pathog.* **1**, 269–278 (2005).
82. Beste, D. J. *et al.* GSMN–TB: a web-based genome-scale network model of *Mycobacterium tuberculosis* metabolism. *Genome Biol.* **8**, R89 (2007).
83. Jamshidi, N. & Palsson, B. O. Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the *in silico* strain *iNJ*661 and proposing alternative drug targets. *BMC Syst. Biol.* **1**, 26 (2007).
84. Hart, P. D., Armstrong, J. A., Brown, C. A. & Draper, P. Ultrastructural study of the behavior of macrophages toward parasitic mycobacteria. *Infect. Immun.* **5**, 803–807 (1972).
85. Armstrong, J. A. & Hart, P. D. Phagosome–lysosome interactions in cultured macrophages infected with virulent tubercle bacilli. Reversal of the usual nonfusion pattern and observations on bacterial survival. *J. Exp. Med.* **142**, 1–16 (1975).
86. Russell, D. G. *Mycobacterium tuberculosis*: here today, and here tomorrow. *Nature Rev. Mol. Cell Biol.* **2**, 569–577 (2001).
87. van der Wel, N. *et al. M. tuberculosis* and *M. leprae* translocate from the phagolysosome to the cytosol in myeloid cells. *Cell* **129**, 1287–1298 (2007).
88. Jordao, L., Bleck, C. K. E., Mayorga, L., Griffiths, G. & Anes, E. On the killing of mycobacteria by macrophages. *Cell. Microbiol.* **10**, 529–548 (2008).
89. Ehrt, S. *et al.* Reprogramming of the macrophage transcriptome in response to interferon-γ and *Mycobacterium tuberculosis*: signaling roles of nitric oxide synthase-2 and phagocyte oxidase. *J. Exp. Med.* **194**, 1123–1140 (2001).
90. Gilchrist, M. *et al.* Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. *Nature* **441**, 173–178 (2006).
One of the first successful attempts to apply systems-biology approaches to the analysis of the dynamic response of macrophages.
91. Roach, J. C. *et al.* Transcription factor expression in lipopolysaccharide-activated peripheral-blood-derived mononuclear cells. *Proc. Natl Acad. Sci. USA* **104**, 16245–16250 (2007).
92. Reed, M. B. *et al.* A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature* **431**, 84–87 (2004).
93. Newton, S. M. *et al.* A deletion defining a common Asian lineage of *Mycobacterium tuberculosis* associates with immune subversion. *Proc. Natl Acad. Sci. USA* **103**, 15594–15598 (2006).
94. Chang, S., Linderman, J. & Kirschner, D. Multiple mechanisms allow *Mycobacterium tuberculosis* to continuously inhibit MHC class II-mediated antigen presentation by macrophages. *Proc. Natl Acad. Sci. USA* **102**, 4530–4535 (2005).
95. Hund, T. J., Kucera, J. P., Otani, N. F. & Rudy, Y. Ionic charge conservation and long-term steady state in the Luo–Rudy dynamic cell model. *Biophys. J.* **81**, 3324–3331 (2001).
96. Alarcon, T., Byrne, H. M. & Maini, P. K. Towards whole-organ modelling of tumour growth. *Prog. Biophys. Mol. Biol.* **85**, 451–472 (2004).
The first published paper to propose a multi-scale approach to understanding a host biological process.
97. Comisar, W. A., Hsiong, S. X., Kong, H. J., Mooney, D. J. & Linderman, J. J. Multi-scale modeling to predict ligand presentation within RGD nanopatterned hydrogels. *Biomaterials* **27**, 2322–2329 (2006).
98. Kirschner, D. E., Chang, S. T., Riggs, T. W., Perry, N. & Linderman, J. J. Toward a multiscale model of antigen presentation in immunity. *Immunol. Rev.* **216**, 93–118 (2007).
99. Kirschner, D. in In silico *Immunology* (eds. Flower, D. & Timmis, J.) 289–312 (Springer, New York, 2006).
100. Flynn, J. L. & Chan, J. Immunology of tuberculosis. *Annu. Rev. Immunol.* **19**, 93–129 (2001).
101. Wigginton, J. E. & Kirschner, D. A model to predict cell-mediated immune regulatory mechanisms during human infection with *Mycobacterium tuberculosis*. *J. Immunol.* **166**, 1951–1967 (2001).
The first detailed mathematical model developed to study the host response to *M. tuberculosis*. Showed how virtual deletions and depletions can be used to perform studies that are not currently tractable in wet laboratories and predict mechanisms to explain experimental results.
102. Sud, D., Bigbee, C., Flynn, J. L. & Kirschner, D. E. Contribution of $CD8^+$ T cells to control of *Mycobacterium tuberculosis* infection. *J. Immunol.* **176**, 4296–4314 (2006).
103. Marino, S. *et al.* Differences in reactivation of tuberculosis induced from anti-TNF treatments are based on bioavailability in granulomatous tissue. *PloS Comput. Biol.* **3**, e194 (2007).
104. Segovia-Juarez, J. L., Ganguli, S. & Kirschner, D. Identifying control mechanisms of granuloma formation during *M. tuberculosis* infection using an agent-based model. *J. Theor. Biol.* **231**, 357–376 (2004).
The first spatial, stochastic framework for studying granuloma formation in TB. This model predicted previously unidentified features of the immune response that are key to containing infection.
105. Marino, S., Hogue, I. B., Ray, C. J. & Kirschner, D. E. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.* 20 Apr 2008 (doi:10.1016/j.jtbi.2008.04.011).

Acknowledgements

This Review was made possible by financial support from the UK Biotechnology and Biotechnology and Biological Sciences Research Council (BBSRC) via the Centre for Integrative Systems Biology at Imperial College (CISBIC), BB/C519670/1. Work described in this Review was supported, in part, by National Institutes of Health grants to D.K. (NIH R01 LM 009027 and NIH R01 HL 072682).

DATABASES

Entrez Genome Project: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj
Mycobacterium tuberculosis

FURTHER INFORMATION

CISBIC homepage: http://www.imperial.ac.uk/cisbic
Comprehensive Microbial Resource of the J. Craig Venter Institute: http://www.jcvi.org/cms/research/projects/cmr/
Pathogen website of the Wellcome Trust Sanger Institute: http://www.sanger.ac.uk/Projects/Pathogens/
Proteome 2D-PAGE Database: http://web.mpiib-berlin.mpg.de/cgi-bin/pdbs/2d-page/extern/index.cgi
TARGET: http://webhost.nts.jhu.edu/target/
TB Database: http://www.tbdb.org/
The Global Plan to Stop TB 2006–2015: http://www.stoptb.org/globalplan/
Time Lapse Simulations of Agent Based Models: http://malthus.micro.med.umich.edu/lab/movies/abm/
***web*TB:** http://www.webtb.org/

ALL LINKS ARE ACTIVE IN THE ONLINE PDF