Published in final edited form as:

Semin Immunol. 2013 October 31; 25(3): . doi:10.1016/j.smim.2012.11.003.

Human systems immunology: hypothesis-based modeling and unbiased data-driven approaches

Arnon Arazi^{a,b}, William F. Pendergraft III^{a,b,c,d}, Ruy M. Ribeiro^{e,f}, Alan S. Perelson^{f,*}, and Nir Hacohen^{a,b,c,*}

^aCenter for Immunology and Inflammatory Diseases, Massachusetts General Hospital, Charlestown, MA 02129, USA

^bBroad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA

^cDepartment of Medicine, Harvard Medical School, Boston, MA 02115, USA

^dDivision of Nephrology, Department of Medicine, Massachusetts General Hospital, Boston, MA 02114, USA

^fTheoretical Biology and Biophysics Group, MS-K710, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Abstract

Systems immunology is an emerging paradigm that aims at a more systematic and quantitative understanding of the immune system. Two major approaches have been utilized to date in this field: unbiased data-driven modeling to comprehensively identify molecular and cellular components of a system and their interactions; and hypothesis-based quantitative modeling to understand the operating principles of a system by extracting a minimal set of variables and rules underlying them. In this review, we describe applications of the two approaches to the study of viral infections and autoimmune diseases in humans, and discuss possible ways by which these two approaches can synergize when applied to human immunology.

Keywords

systems immunology; hypothesis-based modeling; unbiased data-driven approaches; dynamical modeling; viral infections; autoimmunity

1. Introduction

While the definition of systems biology continues to be debated [1-4], the methods and strategies associated with this field are being adopted in every area of biology and medicine. Systems biology encompasses at least two inter-related goals: first, a comprehensive

© 2012 Elsevier Ltd. All rights reserved.

*Corresponding authors: Massachusetts General Hospital, 149 13th St, Charlestown, MA 02129, USA, Tel.: +1 617-724-3768; fax: +1 617-726-5651.; Theoretical Biology and Biophysics Group, MS-K710, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, Tel: +1 505-667-6829; fax: +1 505-665-3493 arnonar@broadinstitute.org (A. Arazi), nhacohen@partners.org (N. Hacohen).

^cPresent address: Institute of Molecular Medicine, Faculty of Medicine, University of Lisbon, Portugal
wpendergraft@partners.org (W. Pendergraft), ruy@lanl.gov (R. Ribeiro), asp@lanl.gov (A. Perelson)

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

description of molecular and cellular components and their interactions; second, the development of mathematical models to help explain and predict the dynamics of a biological system. In this review, we describe two distinct approaches addressing these goals.

The first, more prevalent approach, which we term *data-driven modeling*, is guided by the idea that the model of a studied system should be derived directly from empirical measurements with minimal reliance on prior hypotheses, i.e. based on *unbiased* data collection [5, 6]. This approach utilizes relatively high-throughput measurement (microarrays, next-generation sequencing and protein mass spectrometry) and perturbation (RNAi, knockout) technologies to generate comprehensive inventories of the components of the system and their interactions. Computational and statistical tools are applied to integrate data originating from different assays and, in general, to infer the global and local organization of the system. While this approach is increasingly successful at identifying components and elaborating the structure of the interactions (i.e. networks) underlying biological systems, the high complexity of the resulting descriptions requires an unrealistic number of experiments and increase in computational power to build accurate and usable quantitative models [4, 7]. In human subjects, data-driven modeling has primarily been used to find biomarkers – providing clues to identify relevant biological processes together with novel diagnostic, prognostic or predictive markers. To date, genome-wide microarray profiling of transcript levels from peripheral blood leukocytes (which may or may not reflect processes at the relevant tissue) is the most utilized method [8]. Examples can be found in diverse areas of human immunology, including transplant rejection vs. tolerance [9], vaccine efficacy [10, 11] as well as infections (see Section 4) and autoimmunity (see Section 6).

In the second approach, referred to here as *hypothesis-based modeling*, presumed knowledge about the system is translated into formal mathematical models. For each hypothesis, the associated equations lead to specific predictions, which can then be compared to empirical data in order to identify the hypothesis that best matches the measurements. Quantitative approaches also allow inferring the drivers (e.g. concentrations of specific components, kinetics of particular reactions) of system behavior, enabling prediction-based design of perturbations to manipulate system behavior towards a desired end (such as effective therapy). While practical limitations necessitate the adoption of highly simplistic models, such models have proved useful in the study of human immunity and infection dynamics [12-14], generating new insights and driving further investigation of the underlying processes at the level of the complete organism.

Both approaches have been fruitfully applied to dissect the action of the immune system in the context of human disease; however, we argue that integration of these two approaches will be even more instrumental in advancing translational research. Here, we describe their application to two domains - viral infections and autoimmunity - focusing on a small number of examples. We then suggest different ways in which these approaches may be combined in the future. Since hypothesis-based modeling is less commonly used, we begin by briefly describing its main principles.

2. Hypothesis-based modeling

In the hypothesis-based modeling approach, one poses a question based on a specific hypothesis and then tries to develop a model (often quantitative) to help answer the question of interest. Such models typically simplify complex biological problems in order to reveal essential elements and make predictions about the experimental system.

In Figure 1, we present an example of building such a hypothesis-based model. HIV infection leads to immunodeficiency syndrome, i.e. AIDS, over a period of about 10 years,

suggesting the infection is very slow. Early trials of protease inhibitors to treat chronic HIV-1 infection however demonstrated that viral titers decreased substantially during the first two weeks of treatment (Fig. 1a) [15, 16]. This observation led to the hypothesis that HIV infection must be a very dynamic and rapid process even during the long asymptomatic phase (“clinical latency”). To investigate this hypothesis, a simple model of the infection process was developed (Fig. 1b) [13]. In this model, one assumes that target cells, i.e. CD4+ T cells, are infected by virus, generating infected cells, which in turn produce more virus. Infected cells and virus disappear by death or clearance by immune mechanisms. Each of these steps occurs at rates proportional to the concentration of the intervening players. It is clear that infection is a complex and, probably multistep, process, but in the model it is represented by a single step, akin to a chemical reaction (target cells + virus → infected cells). In the same way, clearance of free virus can represent multiple physiological processes, such as phagocytosis, antibody mediated-clearance, liver filtration, etc. but in the model it is represented by a single rate constant (c) that includes all those processes. All these simplifications are done on purpose to allow analysis and solution of the model, and to address only the specific question of HIV turnover during treatment. In the model, treatment by reverse transcriptase inhibitors can be represented by reducing/blocking infection and in the case of protease inhibitors by reducing/blocking infectious viral production [13]. The model in Fig. 1b can be represented mathematically as differential equations, which can be solved. The solution of the model predicts the decay observed (Fig. 1c) and allows the quantification of the rate of viral clearance and the infected cell turnover. Several analyses of data on treatment dynamics using models of this type showed that infected cells die with a half-life of about 1 day [17]. Since these are representative patients, it follows that this half-life must be typical through most of the infection and hence that infected cell turnover is massive and HIV infection is a very dynamic disease.

Note that initial models of HIV infection ignored explicit mention of immune system processes as these were subsumed into the model parameters. This was done because the experimental data on viral load provided no information about the immune mechanisms underlying virion and infected cell clearance and because this simple model was adequate to describe the data. Later studies in which T cell and antibody responses were also measured allowed consideration of immune response dynamics [18, 19].

Studies involving hypothesis-based quantitative models tend to abstract many of the complexities of the system, in part to uncover general operating principles rather than the details of implementation, i.e., to “see the forest from the trees”. The objective is often to make mechanistic connections among assayed variables, such as between virus concentration and CD4+ T cell levels. Interestingly, a hypothesis-based model that does not describe the data well can provide much insight, because it indicates that the original hypothesis is at least incomplete. Altogether these models are useful for interpreting data, defining what data to collect (e.g. quantities to assay), generating new hypotheses, estimation of dynamic parameters and for gaining a better understanding of the underlying biological processes [20]. Once these goals are obtained more detailed mechanistic models can then be developed, especially if the barriers to obtaining high-resolution quantitative data can be overcome.

3. Hypothesis-based models of viral infections

3.1 Disease pathology

HIV infects CD4+ T cells and leads to their depletion over time, generating a state of immunodeficiency, which invariably when left untreated results in death. To understand the processes behind this depletion, a labeling technique for lymphocytes that could be safely used in humans was developed using deuterated glucose (and later heavy water, $^2\text{H}_2\text{O}$) [21].

However, interpretation of these data required mechanistic models of label incorporation and loss [22-24]. These models, applied to time series data, clearly indicated an increased turnover of both CD4+ and CD8+ T cells during HIV infection [22, 23]. These models also suggested that the increased turnover of CD4+ and CD8+ T cells involved different processes; CD4+ T cells proliferated and died faster, whereas individual CD8+ T cells did not proliferate more rapidly but HIV infection led to an increased fraction of short-lived activated proliferating CD8+ T cells [23]. Antiretroviral treatment led to a clear decrease in the turnover of both CD4+ and CD8+ T cells [22]. These models were then extended to analyze different subpopulations of T cells (e.g., naïve and memory) both in healthy and HIV-infected people [25-27]. One conclusion from these modeling analyses is that to properly evaluate the turnover of T cell subpopulations one needs to distinguish kinetic heterogeneity, i.e. subpopulations that have different turnover rates, from temporal heterogeneity where a subpopulation of cells, say memory cells, have different turnover rates at different times, i.e. when resting vs when activated (see for example [28]).

Another proposed mechanism for the depletion of CD4+ T cells was the failure of the thymus in generating new T cells, due to its natural involution with age and infection by HIV [29]. Again, mechanistic models were needed to interpret the results of experiments trying to measure thymic output using T cell receptor excision circles (TREC). These models showed that the increased T cell proliferation in HIV-infected subjects could explain the experimental observation of changes in the number of TREC per cell better than changes in thymic output [30]. However, these same models also suggested that measuring the dynamics of TREC/ml could shed light on the absolute output of the thymus [31, 32]. TREC are stable [29] and one can envisage that better assays of thymic output would be achieved with specific transient markers for cells of recent thymic origin [33].

3.2 Treatment outcome and optimization

One of the areas of medicine in which mechanistic models have had the most success is in the understanding of the dynamics of virus infection and the effects of treatment. When potent antiretroviral drugs against HIV were first put into patients, a rapid decrease in viral load was observed [15, 16]. Models of the processes of cell infection and viral production when used to interpret clinical data showed that the virus is cleared extremely rapidly and that infected cells while producing virus have an extremely short half-life of ~1 day [17, 34, 35]. These discoveries had immediate impact on our understanding of HIV pathogenesis and treatment. The rapid loss of infected cells gave insight into the processes leading to CD4+ T cell depletion, and the rapid clearance of virus implied that there was rapid production so as to maintain the approximately constant level of virus observed in HIV-infected subjects. This rapid production, estimated at over 10^{10} virions/day, underlies the ability of HIV to quickly mutate away from immune responses and develop resistance to drug treatment [36], a finding that argued for the need of triple combination therapy [13, 37]. After the initial fast decay, HIV viral load decreases more slowly, suggesting that there are other sources of virus besides productively infected CD4+ T cells. These sources were investigated with a variety of models allowing for long-lived infected cells, such as macrophages or infected resting T cells [38], release of virions attached to follicular dendritic cells [39], drug sanctuaries [40, 41], and release of virus from activated latently infected cells [38, 42, 43]. Interestingly, new techniques based on next-generation sequencing may allow discrimination of some of these alternatives, if these compartments harbor specific HIV sequences.

Treatment-based eradication of HIV has not been achieved so far due to the existence of HIV reservoirs, such as latently infected cells. Modeling the decay of the latent reservoir has shown it has a very long half-life which may take as long as 70 years to fully decay [44, 45]. Modeling predicted that this long half-life may be due to replenishment by homeostatic proliferation [46, 47], a prediction later confirmed by experiment [48].

The mechanistic models of viral dynamics developed for HIV have been adapted to study other viral infections, such as hepatitis C virus (HCV) [20], hepatitis B virus (HBV) [49] and influenza A infection [50]. The work on HCV was an important achievement, since the first partially successful treatment of HCV was with interferon- α (IFN α), which is a pan-immune modulator, but the actual mechanism of action was unknown and suggested to be interference with infection by placing cells into an antiviral state. Neumann et al. [51] showed via modeling that the most plausible mechanism involved impairment of viral production by already infected cells, a finding later confirmed *in vitro* [52, 53]. They further showed that the effectiveness of therapy in blocking viral production could be estimated from the observed HCV RNA decline under therapy. The other drug now used in combination with IFN α is ribavirin, a non-specific purine analog precursor, with a still unknown and highly debated mode of action [54]. One possibility is that ribavirin has a mutagenic effect [55], and indeed, modeling has shown that this hypothesis is capable of reconciling a set of disparate clinical results [56].

4. Data-driven models of viral infections

In this section we focus on data-driven studies of *in vivo* viral infections (in contrast to analogous approaches performed *in vitro* [57]). Most of these studies identify a molecular marker or set of markers that associate with particular disease outcomes, thus generating candidate diagnostic, prognostic and predictive markers, as well as novel hypotheses for further testing.

4.1 Infection classification

Systemic profiling has been shown to identify signatures associated with different types of infections. For example, Ramilo et al. [58] demonstrated that transcription profiles of freshly isolated PBMCs can accurately discriminate between acute viral and bacterial respiratory infections, while Ura et al. [59] showed that miRNA expression patterns in liver tissues can distinguish between healthy, HBV-infected and HCV-infected individuals. Similarly, analysis of serum metabolite profiles identified biomarkers of HBV infection [60].

4.2 Disease pathology

Clinical manifestations and progression of virus-induced pathology have been demonstrated to correlate with molecular patterns observed in both the infected tissue and peripheral blood, lending mechanistic insights and potentially facilitating easier diagnosis and prognosis. For instance, proteome profiling of serum samples identified predictors of fibrosis stage in HCV-infected individuals [61, 62], and an analysis of liver biopsies taken from HCV patients pointed to mitochondrial processes and the response to oxidative stress as key pathways whose dysregulation correlates with fibrosis progression [63]. In HIV infection, microarray analysis of peripheral CD4 $^{+}$ T cells revealed different gene expression patterns in viremic and aviremic individuals [64], with higher expression of genes related to RNA processing and protein trafficking and other processes in viremic patients. Furthermore, miRNA profiles of PBMCs were sufficient to accurately discriminate between 4 different classes of HIV-infected individuals, defined by low or high levels of CD4 $^{+}$ T cell counts and viral load [65].

4.3 Immune response

Unbiased profiling tools have generated mechanistic insights into the interactions between the virus and the host immune system, especially when serial measurements were made in the infected tissue. Bigger et al. [66] studied the dynamics of acute-resolving HCV infection in chimpanzees, analyzing serial liver biopsies using microarrays. They found a correlation between the biphasic decline in viral load and the expression of different sets of genes,

demonstrating that interferon-stimulated genes (ISGs) were upregulated early in infection and returned to baseline at the end of the rapid, first-phase decrease in viremia. Kobasa et al. [67] investigated the mechanisms underlying the increased virulence of the highly lethal 1918 influenza virus, analyzing global gene expression in serial bronchi samples from macaques infected with either this strain or a conventional human influenza virus. Animals infected with the 1918 strain displayed less dynamic gene expression changes, especially in the ISGs that were upregulated early in the self-resolving conventional infection. Furthermore, the expression of key cytokines and chemokines was delayed, indicating a dysregulated antiviral response. Applying proteomics tools, Brown et al. [68] studied serial lung samples from macaques infected with various influenza strains, including a highly virulent strain of avian influenza, and identified a cluster of proteins (involved in inflammation and metabolism) that showed early upregulation in response to the virulent strain only. Mechanistic insights can be gained also from the analysis of PBMCs [69, 70] in which distinct gene expression signatures in these cells were associated with symptomatic and asymptomatic infections by respiratory viruses, such that early and persistent upregulation of pattern recognition receptors (PRRs) is seen only in symptomatic hosts while effective cell-mediated responses appear in asymptomatic individuals.

4.4 Treatment outcome

Guadalupe et al. [71] have compared gene expression between jejunal samples from HIV-infected patients before and after treatment. Their results suggested that HAART-induced restoration of CD4+ T cells in the GALT is due to cell trafficking and not local proliferation, as genes associated with the former rather than the latter processes were upregulated in post-treatment samples. For HCV, genome-wide association studies pointed overwhelmingly to alleles of IL28B/IFN determining both spontaneous and treatment-induced resolution of infection [72, 73]. In a complementary study, proteome profiling of pretreatment sera identified unrelated biomarkers predictive of sustained virologic response to treatment in patients not carrying this favorable IL28B allele [74].

5. Hypothesis-based models of autoimmunity

There are relatively few examples of studies applying dynamical modeling conjointly with data collection in the investigation of autoimmune diseases, when compared to the study of viral infections. Below we describe some of the main achievements in this field.

5.1 Physiological processes in disease

In a seminal work with implications for diabetes (whether caused by autoimmunity or not), Bergman et al. [75] presented a “minimal model” describing the dynamics of plasma insulin and glucose, following intravenous administration of glucose into human subjects (as part of a standard glucose tolerance test). Using this model, the authors were able to estimate two physiological parameters, namely the responsiveness of pancreatic β -cells to glucose and an insulin sensitivity index, and quantified the contribution of these two parameters to impaired glucose tolerance, distinguishing between different subsets of patients. Subsequent work refined this model [76] and in addition incorporated into it the dynamics of β -cells [77, 78]. The Bergman minimal model and its extensions were applied to the characterization of glucose dynamics in various physiological and clinical conditions [79-84]. For example, when used in the study of Type I diabetes (T1D), the model demonstrated that, while insulin sensitivity is lower than normal in newly diagnosed T1D patients but returns to normal values in disease remission, glucose effectiveness (a term representing the capacity of glucose to promote its own uptake, at basal insulin levels) remains abnormally low in remission [83]. Future studies incorporating immunological measures into this model should help test hypotheses regarding how immunity contributes to β -cell destruction and disease.

The clearance of circulating immune complexes, a process thought to play a major role in systemic lupus erythematosus (SLE) pathogenesis, was modeled in several studies [85, 86]. The model was used to infer the rates of different steps in this process, based on short-term serial measurements taken from SLE patients following the injection of labeled autologous erythrocytes. The authors predicted a set of correlations between the inferred rates, the levels of serological pathogenic factors (anti-dsDNA antibodies, immune complexes) and disease activity, and verified their existence using the collected data.

5.2 Disease dynamics

Autoimmune diseases often manifest in waves of disease exacerbations and relapses. In one approach [96], a dynamical model was used to explain the cyclic fluctuations observed in circulating T cells in NOD mice, an animal model of T1D; the authors suggested that this may arise from a positive relation between the rate of T cell activation and the level of self-antigen on the one hand, and a negative relation between this level and the rate of immune cells differentiation on the other. In another study [97], urine biomarkers were incorporated into a mechanistic model of lupus nephritis to explain and predict the occurrence of flares in this disease.

5.3 Treatment outcome and optimization

The effect of T cell vaccination (i.e., adoptive transfer) on autoimmunity, in particular in the context of experimental autoimmune encephalomyelitis – an animal model of multiple sclerosis (MS) – was studied in a series of papers from the 1980s and 1990s [87-90]. These studies mainly tried to explain how the injection of a certain concentration of autoreactive T cells can induce autoimmunity, while a lower concentration results in vaccination, protecting mice from further challenges with lethal doses of T cells. In one study [88], a set of minimal phenomenological models consistent with this observation was identified; furthermore, this work predicted that the administration of extremely large doses of autoreactive T cells can cure disease – a counterintuitive result validated experimentally in a subsequent paper [91]. Later studies used direct biological knowledge (rather than phenomenological models) to explain the same findings with simple yet realistic models [89, 90].

A number of works from the Edelstein-Keshet group in close collaboration with experimentalists have addressed the effects of peptide therapy in T1D, while taking into consideration the complexities stemming from the polyclonal T cell response typical to this disease [92-94]. The authors were able to explain why this kind of therapy is more effective as it leads not only to the deletion of high-affinity clones but also to the expansion of low-avidity T cells, and in what scenarios it loses its efficacy – results that can be used to optimize treatment.

Meyer-Hermann et al. [95] constructed a model incorporating interactions between components of the immune and neuro-endocrine systems under circadian control, and applied it to rheumatoid arthritis (RA). The model was fitted to time series data of healthy individuals, and then modified slightly to explain the kinetics observed in RA patients. The authors were able to reproduce with this model the experimental findings showing that optimal inhibition of TNF levels in RA patients is obtained when glucocorticoid treatment is given between midnight and 2am.

6. Data-driven models of autoimmunity

6.1 Disease-specific signatures

As methods for unbiased profiling have improved, so has the ability of researchers to identify more appropriate patients and healthy controls to use as comparisons in order to

reveal more refined signatures. The levels of certain transcripts in peripheral leukocytes from autoimmune patients (e.g., RA or SLE) can be distinguished from levels in unaffected relatives [98] and even between a patient with MS and an unaffected twin [99]. In addition to microarrays, flow cytometric analyses can also define signatures, especially phospho-specific flow cytometry and mass cytometry; for example, a study of intracellular signaling in peripheral leukocytes of 48 humans with RA led to identification of a unique signature of p-AKT and p-p38, which may help in RA diagnosis [100]. More generally, a database of healthy and immune disease blood cell transcript signatures has been collected over several years, and should be quite useful in classifying and understanding autoimmune disease [101].

6.2 Disease Pathology and Progression

Analysis of gene expression in MS plaques revealed involvement of certain pro-inflammatory cytokines, including IL-6, IL-17 and IFN . Furthermore, comparison of samples from acute and chronic plaques showed that IL-17 is mostly upregulated in the latter [102]. Proteomic analysis of similar lesions identified upregulation of protein C inhibitor; interestingly, *in vivo* administration of activated protein C in a murine model of MS led to a decrease in the pathogenic Th17 response and disease [103]. By comparing transcriptomes between samples from MS patients in relapse or remission, dysregulation was observed in apoptosis and inflammation genes during relapse, suggesting novel candidate therapeutic targets [104]. In SLE, genome-wide measurements of transcript levels in blood revealed a type I interferon signature [101, 105], and a granulopoiesis signature that holds in several humoral autoimmune diseases [106]. This led to subsequent studies demonstrating that SLE-derived neutrophils die more frequently and appear to release DNA that can trigger TLR9 activation in pDCs. More recently, cross-species transcriptional networks in this disease were generated by comparison of microarray expression patterns of kidney tissue from three different murine lupus models and from patients with SLE, leading to the identification of multiple common pathways, including mononuclear phagocyte activation as a key component of lupus nephritis in both species, suggesting a path to therapeutic targets [107]. Finally, one study rapidly impacted patient care: microarray studies of JIA blood pointed to interleukin-1 (IL-1) and led to a successful trial of IL-1 antagonists to reduce disease in children [101].

6.3 Antigenic targets associated with disease

Many years of efforts have been placed into autoantigen and autoantibody discovery. Recent advances in unbiased approaches promise to further accelerate this field. For example, the use of mass spectrometry to identify antibody-bound proteins led to the finding of Kir4.1 as a target in ~50% of MS patients [108], and a novel expression library - representing the human proteome using synthetic peptides - was used to comprehensively search for targets of paraneoplastic neurological syndromes [109].

7. How can we combine hypothesis-based and data-driven modeling?

To a large extent, the two approaches discussed above complement each other and are expected to be synergistic if combined. In this section, we discuss a number of proposals for their integration.

7.1 Can hypothesis-based modeling guide design of unbiased data collection experiments?

Ideally, one may wish to use large-scale profiling tools to generate comprehensive descriptions of the immune response under different conditions, identifying and mapping many components and interactions. However, the activity of the immune system spans

several spatial and temporal scales, and therefore this will require frequent monitoring of an exceedingly large number of molecules, cell types, tissues and organs. Given the high costs entailed, such an enterprise will not be possible in the foreseeable future. On the other hand, results from hypothesis-based modeling can be used to narrow the search space of unbiased experiments.

When hypothesis-based modeling proves fruitful in identifying the main events governing and driving complex processes, it may be used to choose specific subsystems whose profiling at certain time intervals is most informative. For example, mathematical modeling of viral dynamics pointed to the loss rate of infected cells as a central parameter determining success or failure of treatment of HCV infection [51, 110]. This rate, which represents the sum of diverse processes, including cell death and the loss of intracellular viral replicative intermediates, is determined in part by the action of the different arms of immune system and varies substantially among infected individuals; however, despite its importance, the exact molecular and cellular mechanisms determining it are far from being clearly understood. It may therefore be of use to apply system-wide measurement tools to investigate this question, for example by studying dynamic global gene transcription in infected or immune cells, and by correlating identified signatures with the cell death rate estimated for each patient.

Hypothesis-based models may be useful even when falling short of explaining empirical data or supplying conclusive results. In particular, such cases may suggest specific applications of large-scale profiling tools to derive better models.

When studying the implications of several alternative hypotheses, hypothesis-based models may lead to the conclusion that all are capable of explaining existing data, thus failing to identify the most plausible explanation. In this case, however, the analysis may still suggest specific experiments to decide between the different alternatives. For example, as described in Section 3, various potential sources of virions were suggested to exist in HIV infection in addition to short-lived productively infected CD4⁺ T cells, including macrophages, infected resting T cells, follicular dendritic cells, drug sanctuaries, and activated latently infected cells. While modeling of these different alternatives did not generate conclusive results about the roles of each of these potential reservoirs [38-43], it nonetheless indicated that applying deep-sequencing to each cellular compartment may help resolve this question, assuming that each compartment harbors specific HIV sequences.

An initial failure to fit a hypothesis-based model to experimental data may lead to the suggestion of additional hypotheses, under which the predictions of the extended model agree with the available measurements. Such additional hypotheses should, however, be validated experimentally, and to this end large-scale profiling may prove to be instrumental. For example, when modeling primary HBV infection, Ciupe et al. [111] concluded that the data would be best explained if over time a fraction of hepatocytes could exist in a state refractory to infection. One can assume that such refractory cells would have a different gene expression profile from cells permissive to infection, and analysis of these profiles at the single cell level with high throughput assays could inform the development of mechanistic models of infection.

Rather than modifying underlying hypotheses, an alternative approach to improve an initially unsatisfactory fit to data is to allow model parameters to change over time. This approach, which was recently demonstrated in models of SLE [97], acute HCV infections in chimpanzees [112] and influenza A virus infection [113], may be particularly reasonable in diseases featuring periodic exacerbations (flares), as changes in parameter values may correspond to and point to unknown disease processes contributing to disease exacerbation.

This can be done systematically by searching for the most plausible (e.g. simplest) parameter changes yielding a good fit, as well as the timing of these changes. To identify key cellular and molecular events underlying such parameter changes, one can then apply large-scale profiling tools at the time points corresponding to these changes, focusing on specific compartments hinted by this analysis to be most relevant. For instance, in SLE modeling one may write an equation for the dynamics of circulating DNA-containing immune complexes (ICs), representing two processes – their clearance, and their formation following the binding of circulating nucleosomes and autoantibodies specific to them. Such a model can be fitted to time-series data of IC measurements, with parallel time series data of circulating nucleosomes and anti-nucleosome antibodies serving as inputs. If the fit fails, one may attempt fitting again while allowing one of the parameters – the affinity of the antibodies or the clearance rate of ICs – to change with time. If changes in the antibody affinity are found to produce a good fit, one may then proceed to profile relevant cell populations – e.g., B cells or follicular helper T cells – at time points just before or after the change. If changes in the clearance rate are deemed a more plausible cause, one may do the same for cells participating in the clearance of ICs, e.g. macrophages. These measurements would then lead to more mechanistic hypotheses to explain what drives these cell populations to behave differently in each individual.

7.2 Can unbiased data collection be used to generate hypotheses for modeling?

In the above examples, hypothesis-based modeling was used as the starting point, and its results defined the subsystem to be studied using data-driven modeling. However, the opposite may be performed as well: in particular, data-driven modeling may be used initially to define or refine the set of hypotheses represented in the model, which can then be explored mathematically. For example, genome-wide association studies identified IL28B as a major factor affecting the responsiveness to IFN treatment in HCV-infected patients [72] as well the rate of spontaneous clearance of infection [73]: this observation can be represented by modifying existing hypothesis-based models such that the rates of viral decline induced by IFN treatment or by endogenous IFN depend on IL28B genotype. As another example, given the complex nature of SLE, it is unclear what processes should be incorporated in a minimal model for the study of this disease. As a possible solution, the modeler may rely on the findings of unbiased profiling studies performed in SLE patients, which consistently point to a major role played by at least three components of the immune system - the type I IFN response, neutrophils and plasma cells [105, 114] – as a starting point for such a model.

7.3 Conclusion

We conclude that quantitative models can guide large-scale data collection, and that in turn, unbiased datasets may create, refine or elaborate hypotheses used in quantitative models. We believe that an increasing interplay between these two complementary approaches is likely to be central and synergistic in the future of biology and medicine.

Acknowledgments

Portions of this work were done under the auspices of the U. S. Department of Energy under contract DE-AC52-06NA25396 (AP) and supported by NIH grants AI028433 (AP), OD011095 (AP), P20-RR018754 (AP), New Innovator Award DP2 OD002230 (NH), P50 HG006193 (NH), GM093080 (NH), and NIH contract HHSN272201000055C (AP). RMR received partial funding from the Foundation for Science and Technology of Portugal (PCOFUND-GA-2009-246542). WFP is supported by a research fellowship from the American Society of Nephrology.

References

- [1]. Kirschner MW. The meaning of systems biology. *Cell*. 2005; 121:503–504. [PubMed: 15907462]
- [2]. Benoist C, Germain RN, Mathis D. A plaidoyer for 'systems immunology'. *Immunol Rev*. 2006; 210:229–234. [PubMed: 16623774]
- [3]. Zak DE, Aderem A. Systems biology of innate immunity. *Immunol Rev*. 2009; 227:264–282. [PubMed: 19120490]
- [4]. Germain RN, Meier-Schellersheim M, Nita-Lazar A, Fraser ID. Systems biology in immunology: a computational modeling perspective. *Annu Rev Immunol*. 2011; 29:527–585. [PubMed: 21219182]
- [5]. Amit I, Regev A, Hacohen N. Strategies to discover regulatory circuits of the mammalian immune system. *Nat Rev Immunol*. 2011; 11:873–880. [PubMed: 22094988]
- [6]. Chuang HY, Hofree M, Ideker T. A decade of systems biology. *Annu Rev Cell Dev Biol*. 2010; 26:721–744. [PubMed: 20604711]
- [7]. Kholodenko B, Yaffe MB, Kolch W. Computational approaches for analyzing information flow in biological networks. *Sci Signal*. 2012; 5:re1. [PubMed: 22510471]
- [8]. Chaussabel D, Pascual V, Banchereau J. Assessing the human immune system through blood transcriptomics. *BMC biology*. 2010; 8:84. [PubMed: 20619006]
- [9]. Roedder S, Vitalone M, Khatri P, Sarwal MM. Biomarkers in solid organ transplantation: establishing personalized transplantation medicine. *Genome Med*. 2011; 3:37. [PubMed: 21658299]
- [10]. Nakaya HI, et al. Systems biology of vaccination for seasonal influenza in humans. *Nat Immunol*. 2011; 12:786–795. [PubMed: 21743478]
- [11]. Nakaya HI, Li S, Pulendran B. Systems vaccinology: learning to compute the behavior of vaccine induced immunity. *Wiley Interdiscip Rev Syst Biol Med*. 2012; 4:193–205. [PubMed: 22012654]
- [12]. Perelson AS, Weisbuch G. Immunology for physicists. *Review of Modern Physics*. 1997; 69:1219–1268.
- [13]. Perelson AS. Modelling viral and immune system dynamics. *Nat Rev Immunol*. 2002; 2:28–36. [PubMed: 11905835]
- [14]. Antia R, Ganusov VV, Ahmed R. The role of models in understanding CD8+ T-cell memory. *Nat Rev Immunol*. 2005; 5:101–111. [PubMed: 15662368]
- [15]. Wei X, et al. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature*. 1995; 373:117–122. [PubMed: 7529365]
- [16]. Ho DD, et al. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*. 1995; 373:123–126. [PubMed: 7816094]
- [17]. Markowitz M, et al. A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and T-cell decay in vivo. *J Virol*. 2003; 77:5037–5038. [PubMed: 12663814]
- [18]. Ganusov VV, et al. Fitness costs and diversity of the cytotoxic T lymphocyte (CTL) response determine the rate of CTL escape during acute and chronic phases of HIV infection. *J Virol*. 2011; 85:10518–10528. [PubMed: 21835793]
- [19]. Tomaras GD, et al. Initial B-cell responses to transmitted human immunodeficiency virus type 1: virion-binding immunoglobulin M (IgM) and IgG antibodies followed by plasma anti-gp41 antibodies with ineffective control of initial viremia. *J Virol*. 2008; 82:12449–12463. [PubMed: 18842730]
- [20]. Layden TJ, Layden JE, Ribeiro RM, Perelson AS. Mathematical modeling of viral kinetics: a tool to understand and optimize therapy. *Clin Liver Dis*. 2003; 7:163–178. [PubMed: 12691465]
- [21]. Hellerstein M, et al. Directly measured kinetics of circulating T lymphocytes in normal and HIV-1-infected humans. *Nat Med*. 1999; 5:83–89. [PubMed: 9883844]
- [22]. Mohri H, et al. Increased turnover of T lymphocytes in HIV-1 infection and its reduction by antiretroviral therapy. *J Exp Med*. 2001; 194:1277–1287. [PubMed: 11696593]

- [23]. Ribeiro RM, Mohri H, Ho DD, Perelson AS. In vivo dynamics of T cell activation, proliferation, and death in HIV-1 infection: why are CD4+ but not CD8+ T cells depleted? *Proc Natl Acad Sci U S A*. 2002; 99:15572–15577. [PubMed: 12434018]
- [24]. Asquith B, Debaq C, Macallan DC, Willems L, Bangham CR. Lymphocyte kinetics: the interpretation of labelling data. *Trends Immunol*. 2002; 23:596–601. [PubMed: 12464572]
- [25]. Hellerstein MK, et al. Subpopulations of long-lived and short-lived T cells in advanced HIV-1 infection. *J Clin Invest*. 2003; 112:956–966. [PubMed: 12975480]
- [26]. Vrisekoop N, et al. Sparse production but preferential incorporation of recently produced naive T cells in the human peripheral pool. *Proc Natl Acad Sci U S A*. 2008; 105:6115–6120. [PubMed: 18420820]
- [27]. Ganusov VV, Borghans JA, De Boer RJ. Explicit kinetic heterogeneity: mathematical models for interpretation of deuterium labeling of heterogeneous cell populations. *PLoS Comput Biol*. 2010; 6:e1000666. [PubMed: 20140186]
- [28]. De Boer RJ, Perelson AS, Ribeiro RM. Modelling deuterium labelling of lymphocytes with temporal and/or kinetic heterogeneity. *J R Soc Interface*. 2012; 9:2191–2200. [PubMed: 22513720]
- [29]. Douek DC, et al. Changes in thymic function with age and during the treatment of HIV infection. *Nature*. 1998; 396:690–695. [PubMed: 9872319]
- [30]. Hazenberg MD, et al. Increased cell division but not thymic dysfunction rapidly affects the T-cell receptor excision circle content of the naive T cell population in HIV-1 infection. *Nat Med*. 2000; 6:1036–1042. [PubMed: 10973325]
- [31]. Lewin SR, et al. Dynamics of T cells and TCR excision circles differ after treatment of acute and chronic HIV infection. *J Immunol*. 2002; 169:4657–4666. [PubMed: 12370406]
- [32]. Arron ST, et al. Impact of thymectomy on the peripheral T cell pool in rhesus macaques before and after infection with simian immunodeficiency virus. *Eur J Immunol*. 2005; 35:46–55. [PubMed: 15593297]
- [33]. Kong F, Chen CH, Cooper MD. Thymic function can be accurately monitored by the level of recent T cell emigrants in the circulation. *Immunity*. 1998; 8:97–104. [PubMed: 9462515]
- [34]. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*. 1996; 271:1582–1586. [PubMed: 8599114]
- [35]. Ramratnam B, et al. Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume plasma apheresis. *Lancet*. 1999; 354:1782–1785. [PubMed: 10577640]
- [36]. Coffin JM. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science*. 1995; 267:483–489. [PubMed: 7824947]
- [37]. Perelson AS, Essunger P, Ho DD. Dynamics of HIV-1 and CD4+ lymphocytes in vivo. *AIDS*. 1997; 11(Suppl A):S17–24. [PubMed: 9451962]
- [38]. Perelson AS, et al. Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature*. 1997; 387:188–191. [PubMed: 9144290]
- [39]. Hlavacek WS, Stilianakis NI, Notermans DW, Danner SA, Perelson AS. Influence of follicular dendritic cells on decay of HIV during antiretroviral therapy. *Proc Natl Acad Sci U S A*. 2000; 97:10966–10971. [PubMed: 10995459]
- [40]. Kepler TB, Perelson AS. Drug concentration heterogeneity facilitates the evolution of drug resistance. *Proc Natl Acad Sci U S A*. 1998; 95:11514–11519. [PubMed: 9751697]
- [41]. Callaway DS, Perelson AS. HIV-1 infection and low steady state viral loads. *Bull Math Biol*. 2002; 64:29–64. [PubMed: 11868336]
- [42]. Rong L, Perelson AS. Modeling latently infected cell activation: viral and latent reservoir persistence, and viral blips in HIV-infected patients on potent therapy. *PLoS Comput Biol*. 2009; 5:e1000533. [PubMed: 19834532]
- [43]. Rong L, Perelson AS. Modeling HIV persistence, the latent reservoir, and viral blips. *J Theor Biol*. 2009; 260:308–331. [PubMed: 19539630]
- [44]. Finzi D, et al. Latent infection of CD4+ T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat Med*. 1999; 5:512–517. [PubMed: 10229227]

- [45]. Ramratnam B, et al. The decay of the latent reservoir of replication-competent HIV-1 is inversely correlated with the extent of residual viral replication during prolonged anti-retroviral therapy. *Nat Med.* 2000; 6:82–85. [PubMed: 10613829]
- [46]. Kim H, Perelson AS. Viral and latent reservoir persistence in HIV-1-infected patients on therapy. *PLoS Comput Biol.* 2006; 2:e135. [PubMed: 17040122]
- [47]. Rong L, Perelson AS. Asymmetric division of activated latently infected cells may explain the decay kinetics of the HIV-1 latent reservoir and intermittent viral blips. *Math Biosci.* 2009; 217:77–87. [PubMed: 18977369]
- [48]. Chomont N, et al. HIV reservoir size and persistence are driven by T cell survival and homeostatic proliferation. *Nat Med.* 2009; 15:893–900. [PubMed: 19543283]
- [49]. Ribeiro RM, Lo A, Perelson AS. Dynamics of hepatitis B virus infection. *Microbes Infect.* 2002; 4:829–835. [PubMed: 12270730]
- [50]. Baccam P, Beauchemin C, Macken CA, Hayden FG, Perelson AS. Kinetics of influenza A virus infection in humans. *J Virol.* 2006; 80:7590–7599. [PubMed: 16840338]
- [51]. Neumann AU, et al. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science.* 1998; 282:103–107. [PubMed: 9756471]
- [52]. Chung RT, et al. Hepatitis C virus replication is directly inhibited by IFN-alpha in a full-length binary expression system. *Proc Natl Acad Sci U S A.* 2001; 98:9847–9852. [PubMed: 11493707]
- [53]. Castet V, et al. Alpha interferon inhibits hepatitis C virus replication in primary human hepatocytes infected in vitro. *J Virol.* 2002; 76:8189–8199. [PubMed: 12134024]
- [54]. Feld JJ, Hoofnagle JH. Mechanism of action of interferon and ribavirin in treatment of hepatitis C. *Nature.* 2005; 436:967–972. [PubMed: 16107837]
- [55]. Crotty S, Cameron CE, Andino R. RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc Natl Acad Sci U S A.* 2001; 98:6895–6900. [PubMed: 11371613]
- [56]. Dixit NM, Layden-Almer JE, Layden TJ, Perelson AS. Modelling how ribavirin improves interferon response rates in hepatitis C virus infection. *Nature.* 2004; 432:922–924. [PubMed: 15602565]
- [57]. Shapira SD, Hacohen N. Systems biology approaches to dissect mammalian innate immunity. *Curr Opin Immunol.* 2011; 23:71–77. [PubMed: 21111589]
- [58]. Ramilo O, et al. Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood.* 2007; 109:2066–2077. [PubMed: 17105821]
- [59]. Ura S, et al. Differential microRNA expression between hepatitis B and hepatitis C leading disease progression to hepatocellular carcinoma. *Hepatology.* 2009; 49:1098–1112. [PubMed: 19173277]
- [60]. Yang J, et al. High performance liquid chromatography-mass spectrometry for metabolomics: potential biomarkers for acute deterioration of liver function in chronic hepatitis B. *J Proteome Res.* 2006; 5:554–561. [PubMed: 16512670]
- [61]. White IR, et al. Serum proteomic analysis focused on fibrosis in patients with hepatitis C virus infection. *J Transl Med.* 2007; 5:33. [PubMed: 17625010]
- [62]. Cheung KJ, Tilleman K, Deforce D, Colle I, Van Vlierberghe H. The HCV serum proteome: a search for fibrosis protein markers. *J Viral Hepat.* 2009; 16:418–429. [PubMed: 19226329]
- [63]. Diamond DL, et al. Proteomic profiling of human liver biopsies: hepatitis C virus-induced fibrosis and mitochondrial dysfunction. *Hepatology.* 2007; 46:649–657. [PubMed: 17654742]
- [64]. Chun TW, et al. Gene expression and viral production in latently infected, resting CD4+ T cells in viremic versus aviremic HIV-infected individuals. *Proc Natl Acad Sci U S A.* 2003; 100:1908–1913. [PubMed: 12552096]
- [65]. Houzet L, et al. MicroRNA profile changes in human immunodeficiency virus type 1 (HIV-1) seropositive individuals. *Retrovirology.* 2008; 5:118. [PubMed: 19114009]
- [66]. Bigger CB, Brasky KM, Lanford RE. DNA microarray analysis of chimpanzee liver during acute resolving hepatitis C virus infection. *J Virol.* 2001; 75:7059–7066. [PubMed: 11435586]
- [67]. Kobasa D, et al. Aberrant innate immune response in lethal infection of macaques with the 1918 influenza virus. *Nature.* 2007; 445:319–323. [PubMed: 17230189]

- [68]. Brown JN, et al. Macaque proteome response to highly pathogenic avian influenza and 1918 reassortant influenza virus infections. *J Virol.* 2010; 84:12058–12068. [PubMed: 20844032]
- [69]. Zaas AK, et al. Gene expression signatures diagnose influenza and other symptomatic respiratory viral infections in humans. *Cell Host Microbe.* 2009; 6:207–217. [PubMed: 19664979]
- [70]. Huang Y, et al. Temporal dynamics of host molecular responses differentiate symptomatic and asymptomatic influenza a infection. *PLoS Genet.* 2011; 7:e1002234. [PubMed: 21901105]
- [71]. Guadalupe M, et al. Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy. *J Virol.* 2003; 77:11708–11717. [PubMed: 14557656]
- [72]. Ge D, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature.* 2009; 461:399–401. [PubMed: 19684573]
- [73]. Thomas DL, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature.* 2009; 461:798–801. [PubMed: 19759533]
- [74]. Patel K, et al. High predictive accuracy of an unbiased proteomic profile for sustained virologic response in chronic hepatitis C patients. *Hepatology.* 2011; 53:1809–1818. [PubMed: 21381069]
- [75]. Bergman RN, Phillips LS, Cobelli C. Physiologic evaluation of factors controlling glucose tolerance in man: measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose. *J Clin Invest.* 1981; 68:1456–1467. [PubMed: 7033284]
- [76]. Cobelli C, Caumo A, Omenetto M. Minimal model SG overestimation and SI underestimation: improved accuracy by a Bayesian two-compartment model. *Am J Physiol.* 1999; 277:E481–488. [PubMed: 10484360]
- [77]. Topp B, Promislow K, deVries G, Miura RM, Finegood DT. A model of beta-cell mass, insulin, and glucose kinetics: pathways to diabetes. *J Theor Biol.* 2000; 206:605–619. [PubMed: 11013117]
- [78]. Bonner-Weir S. beta-cell turnover: its assessment and implications. *Diabetes.* 2001; 50(Suppl 1):S20–24. [PubMed: 11272192]
- [79]. Chen M, Bergman RN, Pacini G, Porte D Jr. Pathogenesis of age-related glucose intolerance in man: insulin resistance and decreased beta-cell function. *The Journal of clinical endocrinology and metabolism.* 1985; 60:13–20. [PubMed: 3880560]
- [80]. Ward WK, et al. Insulin resistance and impaired insulin secretion in subjects with histories of gestational diabetes mellitus. *Diabetes.* 1985; 34:861–869. [PubMed: 3896896]
- [81]. Chang RJ, Geffner ME. Associated non-ovarian problems of polycystic ovarian disease: insulin resistance. *Clin Obstet Gynaecol.* 1985; 12:675–685. [PubMed: 3905162]
- [82]. Raghu P, et al. Reduced insulin sensitivity in nondiabetic, HLA-identical siblings of insulin-dependent diabetic subjects. *Diabetes.* 1985; 34:991–994. [PubMed: 4043555]
- [83]. Finegood DT, Hramiak IM, Dupre J. A modified protocol for estimation of insulin sensitivity with the minimal model of glucose kinetics in patients with insulin-dependent diabetes. *The Journal of clinical endocrinology and metabolism.* 1990; 70:1538–1549. [PubMed: 2189884]
- [84]. Rickels MR, Naji A, Teff KL. Insulin sensitivity, glucose effectiveness, and free fatty acid dynamics after human islet transplantation for type 1 diabetes. *The Journal of clinical endocrinology and metabolism.* 2006; 91:2138–2144. [PubMed: 16569741]
- [85]. Kimberly RP, Meryhew NL, Runquist OA. Mononuclear phagocyte function in SLE. I. Bipartite Fc- and complement-dependent dysfunction. *J Immunol.* 1986; 137:91–96. [PubMed: 3011912]
- [86]. Meryhew NL, Kimberly RP, Messner RP, Runquist OA. Mononuclear phagocyte system in SLE. II. A kinetic model of immune complex handling in systemic lupus erythematosus. *J Immunol.* 1986; 137:97–102. [PubMed: 3486919]
- [87]. Cohen IR, Atlan H. Network regulation of autoimmunity: an automation model. *J Autoimmun.* 1989; 2:613–625. [PubMed: 2572229]
- [88]. Segel LA, Jager E. Reverse engineering: a model for T-cell vaccination. *Bull Math Biol.* 1994; 56:687–721. [PubMed: 8054891]
- [89]. Borghans JA, De Boer RJ. A minimal model for T-cell vaccination. *Proc Biol Sci.* 1995; 259:173–178. [PubMed: 7732037]

- [90]. Borghans JA, De Boer RJ, Sercarz E, Kumar V. T cell vaccination in experimental autoimmune encephalomyelitis: a mathematical model. *J Immunol.* 1998; 161:1087–1093. [PubMed: 9686566]
- [91]. Segel LA, Jager E, Elias D, Cohen IR. A quantitative model of autoimmune disease and T-cell vaccination: does more mean less? *Immunol Today.* 1995; 16:80–84. [PubMed: 7888071]
- [92]. Maree AF, Kublik R, Finegood DT, Edelstein-Keshet L. Modelling the onset of Type 1 diabetes: can impaired macrophage phagocytosis make the difference between health and disease? *Philos Transact A Math Phys Eng Sci.* 2006; 364:1267–1282.
- [93]. Khadra A, Santamaria P, Edelstein-Keshet L. The role of low avidity T cells in the protection against type 1 diabetes: a modeling investigation. *J Theor Biol.* 2009; 256:126–141. [PubMed: 18950644]
- [94]. Khadra A, Tsai S, Santamaria P, Edelstein-Keshet L. On how monospecific memory-like autoregulatory CD8+ T cells can blunt diabetogenic autoimmunity: a computational approach. *J Immunol.* 2010; 185:5962–5972. [PubMed: 20962260]
- [95]. Meyer-Hermann M, Figge MT, Straub RH. Mathematical modeling of the circadian rhythm of key neuroendocrine-immune system players in rheumatoid arthritis: a systems biology approach. *Arthritis Rheum.* 2009; 60:2585–2594. [PubMed: 19714618]
- [96]. Mahaffy JM, Edelstein-Keshet L. Modeling cyclic waves of circulating T cells in autoimmune diabetes. *Siam J Appl Math.* 2007; 67:915–937.
- [97]. Budu-Grajeanu P, Schugart RC, Friedman A, Birmingham DJ, Rovin BH. Mathematical framework for human SLE Nephritis: disease dynamics and urine biomarkers. *Theor Biol Med Model.* 2010; 7:14. [PubMed: 20478032]
- [98]. Maas K, Chen H, Shyr Y, Olsen NJ, Aune T. Shared gene expression profiles in individuals with autoimmune disease and unaffected first-degree relatives of individuals with autoimmune disease. *Human molecular genetics.* 2005; 14:1305–1314. [PubMed: 15814587]
- [99]. Sarkijarvi S, et al. Gene expression profiles in Finnish twins with multiple sclerosis. *BMC medical genetics.* 2006; 7:11. [PubMed: 16504146]
- [100]. Galligan CL, et al. Multiparameter phospho-flow analysis of lymphocytes in early rheumatoid arthritis: implications for diagnosis and monitoring drug therapy. *PloS one.* 2009; 4:e6703. [PubMed: 19693272]
- [101]. Pascual V, Chaussabel D, Banchereau J. A genomic approach to human autoimmune diseases. *Annu Rev Immunol.* 2010; 28:535–571. [PubMed: 20192809]
- [102]. Lock C, et al. Gene-microarray analysis of multiple sclerosis lesions yields new targets validated in autoimmune encephalomyelitis. *Nat Med.* 2002; 8:500–508. [PubMed: 11984595]
- [103]. Han MH, et al. Proteomic analysis of active multiple sclerosis lesions reveals therapeutic targets. *Nature.* 2008; 451:1076–1081. [PubMed: 18278032]
- [104]. Arthur AT, et al. Genes implicated in multiple sclerosis pathogenesis from consilience of genotyping and expression profiles in relapse and remission. *BMC medical genetics.* 2008; 9:17. [PubMed: 18366677]
- [105]. Bennett L, et al. Interferon and granulopoiesis signatures in systemic lupus erythematosus blood. *J Exp Med.* 2003; 197:711–723. [PubMed: 12642603]
- [106]. Alcorta DA, et al. Leukocyte gene expression signatures in antineutrophil cytoplasmic autoantibody and lupus glomerulonephritis. *Kidney international.* 2007; 72:853–864. [PubMed: 17667990]
- [107]. Berthier CC, et al. Cross-species transcriptional network analysis defines shared inflammatory responses in murine and human lupus nephritis. *J Immunol.* 2012; 189:988–1001. [PubMed: 22723521]
- [108]. Srivastava R, et al. Potassium channel KIR4.1 as an immune target in multiple sclerosis. *N Engl J Med.* 2012; 367:115–123. [PubMed: 22784115]
- [109]. Larman HB, et al. Autoantigen discovery with a synthetic human peptidome. *Nature biotechnology.* 2011; 29:535–541.
- [110]. Guedj J, Perelson AS. Second-phase hepatitis C virus RNA decline during telaprevir-based therapy increases with drug effectiveness: implications for treatment duration. *Hepatology.* 2011; 53:1801–1808. [PubMed: 21384401]

- [111]. Ciupe SM, Ribeiro RM, Nelson PW, Dusheiko G, Perelson AS. The role of cells refractory to productive infection in acute hepatitis B viral dynamics. *Proc Natl Acad Sci U S A*. 2007; 104:5050–5055. [PubMed: 17360406]
- [112]. Dahari H, et al. Mathematical modeling of primary hepatitis C infection: noncytolytic clearance and early blockage of virion production. *Gastroenterology*. 2005; 128:1056–1066. [PubMed: 15825086]
- [113]. Pawelek KA, et al. Modeling within-host dynamics of influenza virus infection including immune responses. *PLoS Comput Biol*. 2012; 8:e1002588. [PubMed: 22761567]
- [114]. Chaussabel D, et al. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity*. 2008; 29:150–164. [PubMed: 18631455]

Highlights

- Human systems immunology is an emerging approach for the study of immune disorders.
- Two major approaches used are unbiased data-driven and hypothesis-based modeling.
- Unbiased data-driven approaches are useful in identifying components and their interactions.
- Hypothesis-based modeling derives rules that govern the behavior of a system.
- Synergistic interplay of both approaches will be essential for the development of systems immunology.

Figure 1.

Building a hypothesis-based model. (a) Observation of the decay of HIV viral load under therapy led to the hypothesis that turnover may be fast; (b) and a model was developed to investigate this hypothesis. In the model, lymphocytes (T) are infected by virus at rate k , generating infected cells (I), which produce virus (V) at rate p . Infected cells are lost at rate δ and infected cells are cleared at rate c . (c) This model gives a good description of the data and allows the estimation of the turnover rates by finding the best fit parameters. (d) The modeling results in turn give us biological insight into the dynamics of infection. Data adapted from a published study of HIV dynamics [37].