CHAPTER

15

Systems Biology of Infectious Diseases and Vaccines

Helder I Nakaya[a,b,c]

[a]Department of Pathology, Emory University, Atlanta, GA, USA
[b]Vaccine Research Center, Emory University, Atlanta, GA, USA
[c]Department of Clinical Analyses and Toxicology, University of Sao Paulo, Sao Paulo, SP, Brazil

CONTENTS

Abstract

Pathogens and vaccines stimulate the immune system through an intricate and elaborate network of specialized cells and organs. Immune activation triggers several biological activities that require spatial and dynamic coordination. These activities range from recognition of pathogen-derived factors to cell intercommunication to cell differentiation. Systems biology approaches provide a comprehensive way of dissecting the complex interactions within these processes, and can lead to a better understanding of vaccine-induced immunity and disease pathogenesis. In recent years, systems biology has been successfully applied in analyzing the immune response to a wide range of vaccines and infectious agents. However, dealing with the large amount of data generated from high-throughput techniques and the inherent complexity of the immune system represent major computational and biological challenges. This chapter highlights the recent technological and methodological advances in the field and shows how systems biology can be applied to unraveling novel insights into the molecular mechanisms of immunity.

http://dx.doi.org/10.1016/B978-0-12-405926-9.00015-0

1 INTRODUCTION

The immune system is a complex defense system whose main function is to provide immunity, which is the body's protection from infection and disease. To achieve this, the immune system utilizes a vast number of cell types, each having a specialized role in recognizing and eliminating foreign pathogenic invaders. Although a great number of pathogens are naturally eliminated by the immune system, some evade the host defenses and proliferate, establishing infection that may result in hospitalization and death. In immunocompromised populations, non-pathogenic microorganisms can also result in severe infection. It is how the immune cells respond and interact with each other and the pathogens that determine the disease outcome. So far, very little is known about the dynamic network of cells and molecules behind such responses. Similarly, we have a limited knowledge of the molecular mechanisms that mediate vaccine-induced immunity. Despite their undisputed successes, most vaccines were designed empirically, with the gene networks and pathways associated to immunity remaining largely unknown (Nakaya et al. 2012).

Major technological breakthroughs of high-throughput technologies now allow scientists to study entire genomes, sets of transcripts (transcriptome), proteins (proteome), and metabolites (metabolome) of cells and tissues (Nakaya et al. 2012). As these technologies become more accessible and cost-efficient, their use has expanded, creating a unique opportunity for scientists to profile immune responses of large cohorts and in clinical trials (Chaussabel et al. 2010). Systems biology utilizes the enormous amount of data generated by these techniques to describe the complex interactions between all the parts of a biological system under different types of perturbations, and to construct mathematical models to predict the behavior of the system (Ideker et al. 2001; Kitano 2002). This powerful approach promises to revolutionize our understanding of biological systems.

The application of systems biology to vaccinology and infectious disease may potentially fill many fundamental gaps in basic immunology, as well as enable the rational design and testing of novel vaccines (Nakaya et al. 2012). Also, the exploration of the interconnected networks that control and drive the immune response to infection may shed some light on the understanding of infectious diseases (Casadevall and Pirofski 2003) and identify therapeutic targets and improve clinical outcomes (Aderem et al. 2011). However, the inherent complexity of immune system and the technical difficulties associated with the computational analyses of large datasets imposes great challenges for the field. In this chapter I describes how systems biology is re-shaping the study of immune responses and the many challenges associated with this exciting enterprise.

2 A BRIEF OVERVIEW OF THE IMMUNE RESPONSE

The capacity to respond to as well as to clear and/or control an invading pathogen is the primary goal of the immune system. This is achieved by two different but complementary arms: the innate immune system, which quickly (few hours to few days) recognizes and responds to pathogens, and the adaptive immune system, which is responsible for mounting a late response, but one that can persists for years (memory). Both arms form a remarkably versatile defense system (Kuby et al. 2002) that involves a myriad of cells and molecules that act together in a dynamic and complex fashion. This section provides a simplified description of the main effectors of immune responses.

The recognition of a pathogen by specialized cells can trigger a broad range of effector responses, each uniquely suited to eliminate a particular type of pathogen (Kuby et al. 2002). More importantly,

if a memory response is induced, a later exposure to the same pathogen will result in a rapid resolution by the immune system, eliminating the pathogen and preventing disease (Kuby et al. 2002). Vaccines attempt to mimic these naturally induced memory responses (or immunity) by exposing the immune system to altered forms of a pathogen or components of a pathogen which does not cause disease. Since adaptive immunity is contingent upon an effective innate immune response, a successful vaccine must be able to induce both arms of immune responses.

In addition to anatomic and humoral barriers, the innate immune system utilizes a great variety of specialized cells and complex biological processes. Neutrophils and macrophages are able to engulf and kill certain types of pathogens, whereas Natural Killer (NK) cells can non-specifically kill cells that have been infected by viruses. Dendritic cells can process and present antigens (i.e., foreign substances that evoke the production of antibodies) to cells of the adaptive immune system. All these processes involve an intricate number of chemical molecules (RNAs, proteins, metabolites) being produced, released, and detected by these innate cells. Among the soluble products of innate immunity, we can cite many cytokines and chemokines, as well as antimicrobial peptides and complement fragments.

Adaptive immunity protects the body against infection with antibodies and T cells. Antibodies (or immunoglobulins) are proteins produced and secreted by B cells, which can bind to antigens on invading pathogens. This interaction can directly inactivate the pathogen or activate a variety of inflammatory mediators that will destroy the pathogen (Rote 2009). T cells, during an immune response, can differentiate into several subpopulations of effector T cells that have an effect on many other cells (Rote 2009). Cytotoxic T cells (usually expressing the surface protein CD8) will attack and kill cells infected by viruses. T helper cells, which express the protein CD4 on their surface upon maturation, can stimulate the activities of other leukocytes through cell-to-cell contact or through the secretion of cytokines (Rote 2009). Thus, the appropriate interactions between T and B cells and the functions of both types of responses are critical for the success of an acquired immune response (Rote 2009).

The main effectors of vaccine response are becoming gradually unveiled in the last decades. Figure 15.1 summarizes the major events that may be involved in this response. Upon exposure to antigen, cells from the innate immune system, such as immature dendritic cells (DCs) and monocytes become activated and migrate to lymph nodes. During this migration, these antigen presenting cells process the antigen into small peptides and display them at the cell surface to B, CD4+ T, and CD8+ T cells. DC-activated CD4+ T cells can differentiate into mutually exclusive differentiation pathways (Figure 15.1). T helper type 1 (Th1) CD4+ T cells support CD8+ T cell activation and can also participates in the elimination of intracellular pathogens. T helper type 2 (Th2) CD4+ T cells (and Th1 cells) support B cell activation and differentiation into short-live antibody secreting plasma cells (plasmablasts), whereas follicular CD4+ helper T cells (Tfh) and follicular DCs (fDC) provide help to germinal center (GC) B cells. Vaccines may instead elicit regulatory CD4+ CD25+ T cells (Treg). DCs that capture antigen in the absence of "danger" signals remain immature in their migration to lymph nodes, inducing naïve T cells to differentiate into Tregs and not effector T cells. These Tregs can suppress Th1 and Th2 responses. CD8+ T cells activated by DCs become cytotoxic T cells which are capable of killing infected cells or pathogens. These cytotoxic T cells can differentiate into two types of memory T cells: effector memory T cells or central memory T cells. If the latter cells recognize antigens transported by activated DCs, they can quickly proliferate and differentiate into the former cells. Vaccine antigen-specific B cells when activated by DCs and CD4+ T cells rapidly differentiate into plasmablasts. If sufficient help from antigen-specific T cells is received, the B cells proliferate in the

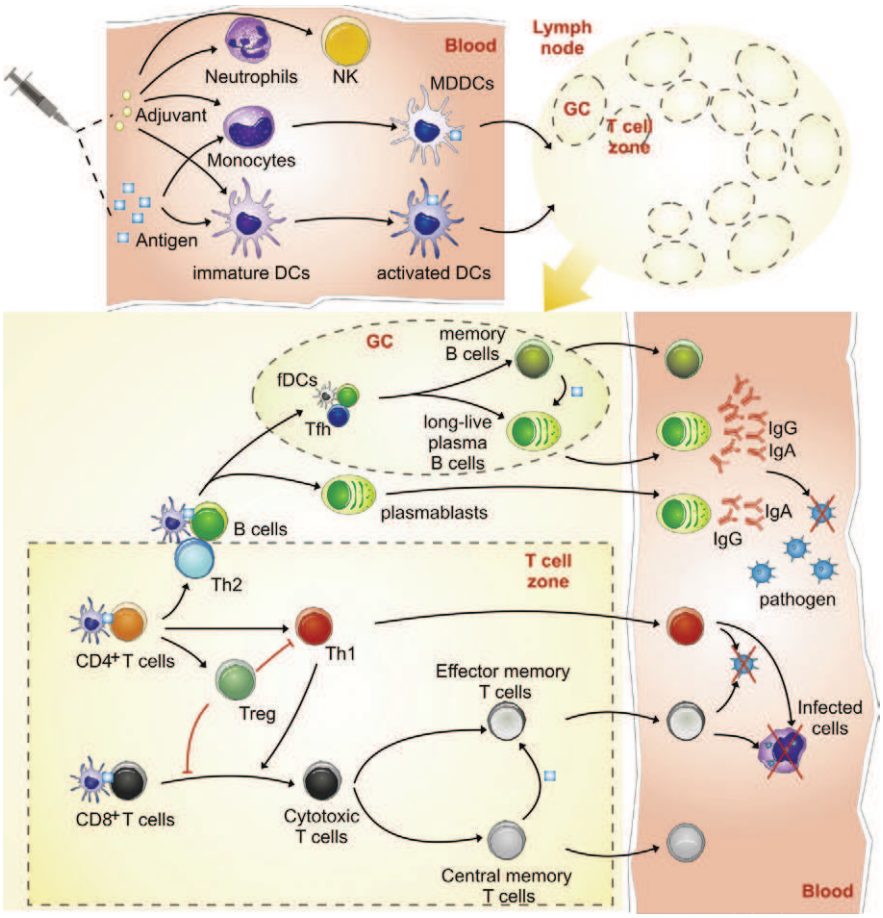

FIGURE 15.1 Main effectors of vaccine response. Vaccine-induced immune effectors include antibodies, produced by B lymphocytes, and cytotoxic CD8+ T lymphocytes. The generation and maintenance of specific B and CD8+ T cell responses is supported by CD4+ T helper lymphocytes. Cells from the innate immune system, such as dendritic cells and monocytes play a pivotal role in regulating the quantity, quality, and longevity of the adaptive immune response induced by the vaccine. DC, dendritic cell; fDC, follicular dendritic cells; Tfh, follicular CD4+ helper T cells; Treg, regulatory CD4+ CD25+ T cells; NK, natural killer cells; GC, germinal center.

germinal centers and can differentiate into memory or long-live antibody secreting plasma B cells. The secreted antibodies (mostly IgG and IgA) can neutralize the pathogen and protect the body from infection if their affinity (quality) and abundance (titers) are sufficiently high and sustained on mucosal surfaces (Figure 15.1). Finally, adjuvants play their role by stimulating different innate immune cells, improving the potency of the immune response to co-administered antigens (Kwissa et al. 2007).

Although the main cellular players of immune responses have been identified, the genetic and signaling programs that control their differentiation, activation status, proliferation, or death are poorly characterized. For instance, which genes should be activated in plasma B cells

to increase the quality and quantity of antibodies? What is the genetic program of activated DCs that will contribute to a strong CD8+ cytotoxic T cell response? Which genes determine the fate of CD4+ T cells after its contact with activated DCs? The road for answering these questions remains long but systems biology can increase our pace toward discovery.

3 SYSTEMS IMMUNOLOGY TOOLS AND DATABASES

3.1 The arsenal of modern immunologists

Conventional immunology and reductionist methods of cellular and molecular biology have identified most of the known components of the immune system (Ricciardi-Castagnoli and Granucci 2002). Methods such as ELISA (Enzyme-linked immunosorbent assay), flow-cytometry, and real-time quantitative PCR are still invaluable in describing and analyzing specific cell subsets, proteins, or genes during an immune response. However, they have limited power when it comes to analyzing several features of a system in parallel, and focus instead on parts of the immune system rather than the whole (Nakaya et al. 2012). A more holistic biological approach is required in order to reassemble these parts and their connections and dynamics during an immune response.

With the revolution in nanotechnology, robotics, optics, and electronics, immunologists have now an arsenal of high-throughput technologies at their disposal. A wide range of molecular and cellular profiling assays, such as next-generation sequencing, CyTOF® and Luminex®, are easily accessible to laboratories investigating human health and disease. Global changes in metabolites or protein levels are assessed by mass spectrometry and nuclear magnetic resonance. Even the phosphorylation state of intracellular proteins, which is an important regulator of signal transduction pathways, can also be measured through phospho-specific flow cytometry. This method named phospho flow has been successfully applied to the characterization of signaling pathways after a microbial challenge and after antigenic stimulation (Hotson et al. 2009; Krutzik et al. 2005).

Among the so-called 'omics' technologies, DNA microarrays are still the preferable high-throughput screening tool for most systems immunologists (Haining and Wherry 2010). Representing thousands of distinct genes or genomic regions, microarrays can be used to profile a wide range of cellular and molecular processes: gene transcription, alternative splicing, DNA methylation, single-nucleotide polymorphisms (SNPs), proteins binding to DNA– and RNA–elements (ChIP-on-chip), microRNA and noncoding RNA transcription and genomic copy number variations. In addition to its flexibility, the availability of many easy-to-use analytical software and programs, allied to its relative low cost, DNA microarrays are being vastly applied in immunology, from measuring the global transcriptional changes of dendritic cell during differentiation (Le Naour et al. 2001) to assessing the global changes in STAT-target genes upon interferon treatments (Hartman et al. 2005). Many more examples can be found investigating the global expression changes of genes and microRNAs during infection (Jenner and Young 2005; Zeiner et al. 2010) or after vaccination (Gaucher et al. 2008; Nakaya et al. 2011; Querec et al. 2009).

The analysis of genomes (DNA-seq) and DNA elements (ChIP-seq), as well as transcriptomes (RNA-seq) and gene regulatory networks can also be achieved through next-generation sequencing technologies (Zak and Aderem 2009). This powerful approach generates millions of sequence reads per run and is able to measure the transcriptome of cells with high genome

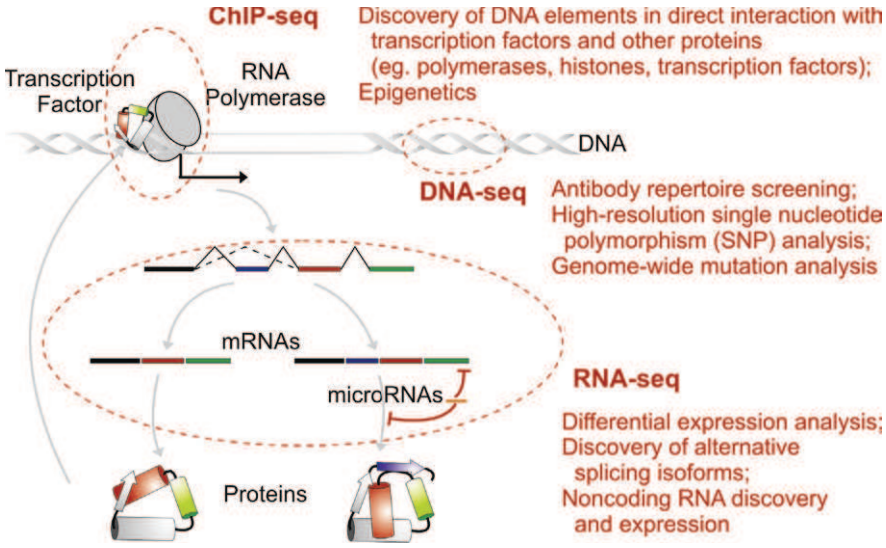

FIGURE 15.2 Applications of next-generation sequencing technology to study immune responses. Next-generation sequencing techniques can be used to analyze genomes or DNA (DNA-seq), DNA elements (ChIP-seq), and transcriptomes (RNA-seq) in the context of immune regulation.

coverage. As the cost of next-generation sequencing goes down, it will potentially replace hybridization-based technologies such as DNA microarrays (Nakaya et al. 2012). Next-generation sequencing has many potential applications for studying immune responses (Figure 15.2). For example, RNA-seq experiments revealed that RIG-I, a key player in the recognition of RNA viruses, preferentially associates with shorter viral RNA molecules in infected cells (Baum et al. 2010), or that long noncoding RNAs are differentially induced during the course of viral (Peng et al. 2010). A variation of RNA-seq method, called genome-wide nuclear 'run-on' analysis (GRO-seq), and chromatin immunoprecipitation combined with next-generation sequencing (ChIP-seq) were used to determine the impact of the Influenza virus NS1 gene on the dynamics of host gene transcription, demonstrating how the H3N2 influenza A virus interferes with antiviral gene expression by exploiting the very basic principles of the epigenetic control of gene regulation (Marazzi et al. 2012). And finally, DNA-seq can be used to characterize natural antibody repertoires induced by infection or vaccination, and may significantly improve antibody discovery (Fischer 2011).

At the cellular level, two-photon microscopy is the method that best captures the dynamic nature of how the cells behave in their native environment (Bullen et al. 2009). This technology allows the real–time visualization of immune cells during differentiation, migration, death, and interactions with other cells and microbes, and has shed important light on T cell development, antigen recognition, cell trafficking, and effector functions (Mirsky et al. 2011). The integration of cellular behavior data with the molecular data generated by microarrays and high-throughput sequencing technologies may have a huge impact on identifying the functional regulatory circuits associated to maturation, survival, and pathogen-recognition of immune cells.

A novel single cell detection technology, named cytometry by time-of-flight, or CyTOF®, promises to revolutionize the field. This technique is much more powerful than the Fluorescence Activated Cell Sorting (FACS) tool because it allows simultaneous measurement

of up to 100 surface markers and signaling proteins of immune cells (Bendall et al. 2011). It employs antibodies labeled with transition element isotopes and a time-of-flight mass spectrometer. By analyzing the expression of many more surface markers and cytokines, Newell et al were able to observe an intriguingly diverging phenotypes in $CD8^+$ T cell populations (Newell et al. 2012).

Genome-wide RNA interference-based screening represents one of the most powerful approaches of functional genomics. It generally utilizes small interfering RNA (siRNA) libraries constructed in 96/384-well plates where, on each well, a single gene should be silenced. Cells are then exposed to the same conditions (e.g. virus infection) and distributed on the plates. The effect of knocking down the expression of these genes on a given cellular process is evaluated by a subsequent screening readout (Houzet and Jeang 2011). This method was successfully applied to identify cellular factors implicated in HIV replication (Konig et al. 2008; Brass et al. 2008; Zhou et al. 2008), as well as influenza virus replication (Konig et al. 2010; Karlas et al. 2010).

New high-throughput multiplex immunoassays, such as antibody microarrays, are available for diagnostic discovery, biomarker-assisted drug development, or "simply" quantitative proteomics studies (Kingsmore 2006). In addition to common proteomics and metabolomics technologies that rely on quantitative mass spectrometry, antibody microarrays consist of ordered, immobilized capture antibodies paired with mixtures of detector antibody that allows highly parallel measurement of protein levels (Kingsmore and Patel 2003). This sensitive protein microarray technique that incorporated rolling-circle amplification technology was used to elucidate the cytokine response signature of systemic adverse events following smallpox immunization (McKinney et al. 2006).

3.2 Databases

Fortunately, the enormous amount of information derived from both reductionist and holistic approaches is being systematically deposited into public databases. More than simply data repositories, most of the databases are excellent tools for data mining and functional analysis. Certain databases contain a broad range of immune-related data, from gene expression data of macrophages to 3D protein structures of influenza viruses; some of them are summarized in Table 15.1.

Equally broad are the many applications of such databases to systems biology analyses. When available, high-quality data can be downloaded from the databases and used on a particular analysis or meta-analysis. For example, the comprehensive information on gene functions, manually curated pathways and transcription factor and microRNA regulation can be extremely useful for interpreting the findings of large–scale analyses. Additionally, gene expression and co-expression data of specific cell types under different conditions or the experimentally validated interactome (protein–protein, protein–DNA, microRNA–RNA) data can be a valuable resource for network analysis and pathway enrichment analysis. Also, many of these databases include data of animal models and *in vitro* experiments facilitating the design and translation to functional validation experiments.

The InnateDB (http://www.innatedb.com) represents a state-of-the-art platform that can facilitate systems-level analyses of mammalian innate immunity. It contains publicly available data of genes, proteins, networks, and signaling pathways involved in the innate immune response of mammals to microbial infection (Lynn et al. 2008). To date, the InnateDB team has

TABLE 15.1 Immunological databases and tools.

Name	Purpose	Data type	Reference or website
Macrophages.com	Macrophage biology	*GE; P; NET*	http://www.macrophages.com
DC-ATLAS	Dendritic cell biology	*GE; NET*	http://www.dc-atlas.net/
Immunome	Genes and proteins of the human immune system	*PD; GF*	http://bioinf.uta.fi/Immunome
Innate Immune Database (IIDB)	Genes and proteins of the mouse immune system	*GF; NET; GR*	http://www.innateimmunity-systemsbiology.org/
Immunological Genome Project	Gene expression data of well-defined cell types	*GE; NET; GR*	http://www.immgen.org/
Influenza Research Database	Resource for the influenza virus research community	*S; PD*	http://www.fludb.org
SystemsInfluenza.org	Immune system response and interaction with the influenza virus	*GE; I*	http://www.systemsinfluenza.org/
InnateDB	Innate immune response to infection	*GF; NET; I*	http://www.innatedb.com
Systems Virology Center	Molecular processes related to progression of infectious diseases	*GE; P*	https://www.systemsvirology.org
TB Database	Integrated platform for Tuberculosis research	*GF; GE; NET; P*	http://www.tbdb.org/
miR2Disease	miRNA deregulation in various human diseases	*miR*	http://www.mir2disease.org/
ImmunoDeficiency Resource	Compendium of information on the immunodeficiencies	*GF*	http://bioinf.uta.fi/idr/
IMGT	Sequence data of immune-related genes	*PD; S*	http://www.imgt.org/
Interferome	Interferon regulated genes	*GE; GF; GR; S*	http://www.interferome.org/
LymphTF DB	B and T Lymphocyte Transcription Factor Activity in Mouse	*GR*	http://www.iupui.edu/~tfinterx/
ImmPort	Data warehouse of immunological data	*GE; GF;*	https://immport.niaid.nih.gov
PHISTO	Comprehensive pathogen-human protein-protein interaction	*I*	http://www.phisto.org

GE = Gene expression data; P = Protein level data; NET = Pathways, gene networks and co-expressed genes data; PD = Protein domains and 3D structures; GF = Gene annotation or function; GR = Gene regulation data; S = sequence data; I = Interactome data; miR = micro RNA data.

manually curated more than 18 thousands molecular interactions related to innate immunity, and more recently to allergy and asthma. The platform can also be used as a tool for Gene Ontology analysis, network visualization, and Transcription Factor Binding Site analysis.

Another benchmark platform for systems immunology analyses was developed by the Immunological Genome Project team (http://www.immgen.org/). The database contains a large collection of gene expression profiles of well-defined cell types in different states of differentiation, maturation, and activation responses. This huge amount of information (publicly available for downloading) supports the computational reconstruction of the genetic circuitries underlying several immune-related processes and can be used to reveal the gene signatures of specific immune cell populations. The platform also contains user-friendly tools for generating interactive heatmaps of gene families, displaying gene modules of coregulated genes, visualization of RNA-seq data of two cell lineages, and for revealing the differences in gene expression between different cell populations.

4 BLOOD TRANSCRIPTOMICS

Blood is the main channel for migration and trafficking of immune cells between lymphoid organs and other sites of the body, and therefore provides a comprehensive view of the status of the immune system in health and disease (Chaussabel et al. 2010). In addition, blood can be extracted, stored and processed easily, making it one of the most favorable tissues for studying human immunity at the systems-level. To date, hundreds of blood transcriptomic experiments were performed and many more are underway. However, due to its inherent heterogeneity, the analysis of gene expression data from blood is challenging.

A common dilemma when designing blood transcriptomic studies involves the choice of sample types to be screened (Nakaya and Pulendran 2012). Blood (e.g. whole blood or peripheral blood mononuclear cells, PBMCs) is composed of multiple cell lineages in different states of activation and differentiation (Pulendran et al. 2010). In response to a stimulus (e.g. infection or vaccination), different cell subsets can proliferate, migrate, and die in a dynamic and interdependent fashion. Therefore, transcriptional profiling of this mixed population tissue will capture the changes of both cellular and transcript abundance, making results difficult to interpret (Chaussabel et al. 2010). An alternative solution would be to measure the transcriptome of isolated cell subsets, which can provide valuable mechanistic insights of specific cells (Nakaya et al. 2011; Grigoryev et al. 2010). Notwithstanding, the high cost and great amount of work associated with isolating and screening cell subsets make this approach unpractical and prohibitive to most clinical studies (Nakaya and Pulendran 2012).

Deconvolution methods can offer a good solution to deal with the cell-type composition problem of blood samples. They can be applied to accurately quantify the constituents of blood samples (Abbas et al. 2009) or, when cell subset frequency is known, to identify subset-specific differential expression (Shen-Orr et al. 2010). In order to quantify the proportions of cells in blood, the expression signatures of each immune cell subset must be known. Assuming that the expression of genes in a given sample can be modeled as a linear combination of the expression of those genes in each of the cells in the sample, linear equations can then be used to fit the fractions of cell subsets to the whole blood sample's expression (Lu et al. 2003). Gene signatures of immune cell subsets were collected or analyzed by many groups and can be used to deconvolute blood transcriptomic data (Abbas et al. 2005; Palmer et al. 2006; Miller et al. 2012; Shaffer et al. 2001). The second deconvolution method requires known or estimated frequencies of cell types in order to perform cell type–specific differential expression analysis

(Shen-Orr et al. 2010). Two caveats of these methods are that (1) the immune cell types are ambiguously defined, and (2) flow cytometry data and microarray expression were shown to be poorly correlated (Hyatt et al. 2006).

Other methods developed to analyze blood microarray data compare blood signatures to previously defined gene sets or modules of particular a cell type (Nakaya et al. 2011; Berry et al. 2010; Chaussabel et al. 2008). Transcriptional modules defined by genes coordinately expressed in multiple blood transcriptomic studies were successfully used to classify disease activity (Berry et al. 2010; Chaussabel et al. 2008). Using an alternative approach, our group first identified the gene signatures of the main peripheral blood mononuclear cells (i.e. monocytes, myeloid dendritic cells, plasmacytoid dendritic cells, B cells, T cells, and NK cells) and then compared these immune signatures to the genes identified in our blood transcriptomics analysis in order to find significant enrichment of specific cell types (Nakaya et al. 2011). Another computational method, developed by Bolen et al. (2011) attempts to predict the most likely cellular source for a predefined gene expression signature using only the transcriptional profiling data from total PBMCs.

5 SYSTEMS BIOLOGY OF INFECTIOUS DISEASES

Investigating the complex interactions and changes that occur during infection or exposure to a pathogen can help us answer many questions about microbes and the immune system. What are the factors and mechanisms that protect cells from being infected or that prevent a virus from efficiently replicating inside a cell? Which signaling pathways need to be activated or repressed to mount an effective T cell response? How do pathogens successfully avoid host recognition or their destruction by host immune cells? Why are some individuals susceptible to certain pathogens while others are not? The list of questions is long and many others will arise once we start unraveling the mechanisms of infectious diseases.

Systems-biology provides the most straightforward approach for global modeling of microbial pathogenesis. Figure 15.3 shows examples of experimental designs and strategies. In general, high-throughput technologies are used to measure the behavior of all the components of immune system cells or tissues during an infection. Computational and mathematical models are then constructed, validated, and analyzed using this data, and key networks of interactions are identified. At last, this approach can suggest novel functional properties and be able to predict the most informative sets of future experiments (Young et al. 2008).

Investigating the molecular interactions between a pathogen and the host can also be performed at systems-level. To be able to sense viruses, bacteria, parasites, or fungi, the cells of the innate immune system express the so-called pattern recognition receptors (Pulendran et al. 2010). In an elegant example of network modeling performed by Amit et al. (2009), the authors reconstructed, in an unbiased way, the transcriptional networks triggered by five pathogen-derived components, including the adjuvant CpG, a synthetic single-stranded DNA that binds to toll-like receptor 9. They also constructed a model that associated regulators with their targets by systematically silencing the expression of most of the candidate regulators in combination with gene expression profiling (Amit et al. 2009). Using a perfect example of systems-biology approach (i.e. the systematic perturbation of the system, the measurements of its effect on global

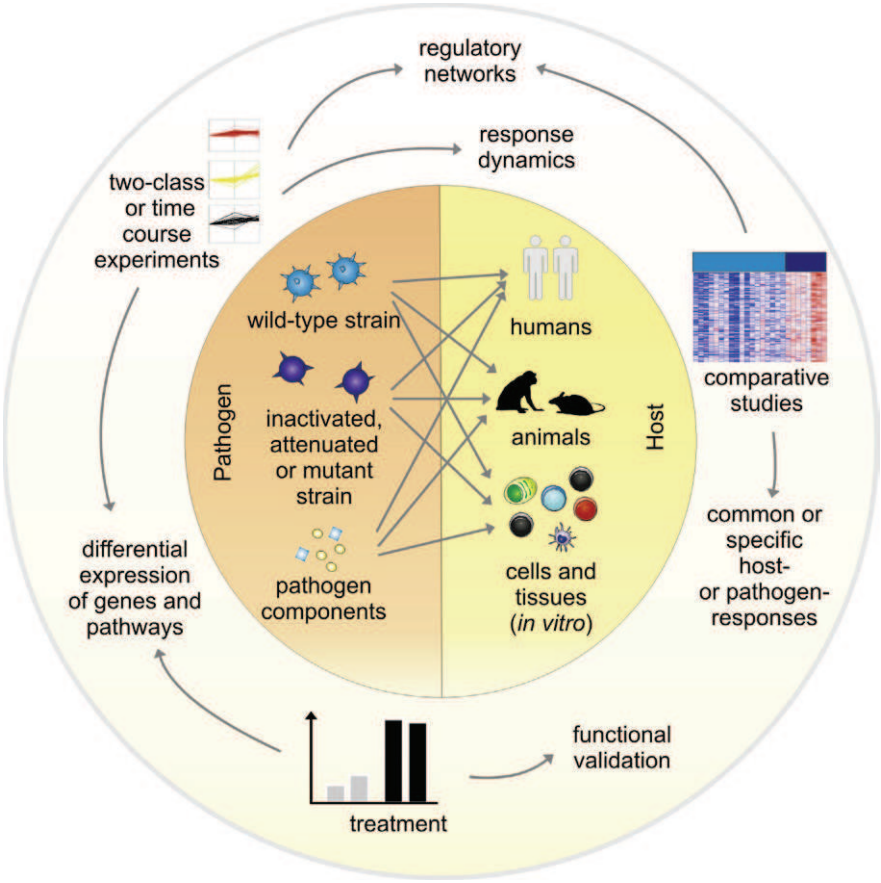

FIGURE 15.3 Systems-biology approaches applied to investigate host-pathogen interactions. Different host-pathogen models of infection or vaccination are shown in the central circle. The various types of analyses that can be exploited to holistically assess and predict these interactions are shown in the outer circle.

scale, and the mathematical modeling to explain or predict the data), Amit et al. found that different regulatory arms control two key programs related to inflammatory response and antiviral response (Amit et al. 2009).

Modeling infections at the population-level poses a great challenge to epidemiologists, immunologists, and computational scientists. The underlying assumption of most studies of host–pathogen and host-host interactions is that the human population can be treated as a homogenous group. However, models generated under such assumption tend to provide a poor fit to observed data and many times fail to make correct predictions (Young et al. 2008). Therefore, it is critical that individual variation be taken into account when modeling host responses.

As we will see in the following sections, the appropriate mathematical or computational model can vary greatly. Modeling methods, such as ordinary and partial differential

equations (Kell 2006; Fell 1992), stochastic schemes (Arkin et al. 1998), petri nets (Goss and Peccoud 1998), and Boolean logic (Watterson and Ghazal 2010) can be potentially applied to understand the regulatory circuitries involved in response to infections. The choice of the method relies, in part, on the type of data and on the dynamics that are under consideration (Young et al. 2008).

5.1 Viral infections

According to the World Health Organization (WHO), seasonal influenza epidemics occur yearly, resulting in 250,000 to 500,000 annual deaths worldwide. Due to this great public health threat, influenza virus infection is being thoroughly investigated through systems biology approaches. In a study by Huang et al. (2011), the authors analyzed the dynamics of host molecular responses that differentiate symptomatic and asymptomatic influenza infection. They challenged 17 healthy adults with live influenza virus and measured the blood transcriptomic changes that occur at 16 timepoints over 132 hours. While symptomatic hosts invoked multiple antiviral and inflammatory responses related to virus-induced oxidative stress, asymptomatic subjects exhibited elevated expression of genes that function in antioxidant responses and cell-mediated responses (Huang et al. 2011). A standard linear mixed model was used to estimate the correlation between the temporal expression patterns of clusters of genes (clustered by Self-Organizing Map) and the disease symptom scores (defined by a standardized symptom scoring system) (Huang et al. 2011). Unsupervised analyses (i.e. that did not require prior class information) were performed using Bayesian Linear Unmixing, revealing that the gene signature of symptomatic infection was strong enough to separate subjects into these two clinical phenotypes (in this case, "symptomatic" and "asymptomatic") (Huang et al. 2011).

However, host responses triggered by influenza infection will be contingent upon the cell type that is being infected and, of course, by the strain of the virus. One of these strains (the "1918 influenza virus strain") killed an estimated 50 million people during the "Spanish influenza" pandemics of 1918 and was recently reconstructed by modern reverse genetics techniques (Billharz et al. 2009). Billharz et al. performed systems-level analyses of human lung epithelial cells infected *in vitro* with 1918 virus to define the global host transcriptional response to this influenza strain (Billharz et al. 2009). Differentially expressed genes were identified by creating ratio profiles by combining replicates and applying error weighting (consisting in adjusting for additive and multiplicative noise) (Billharz et al. 2009). For that, they utilized the Rosetta Resolver system, a multi-user system for storage and analysis of gene expression microarray data (http://www.lgtc.nl/rosetta/). This approach revealed that the protein NS1 of the 1918 virus must block host interferon and lipid metabolism signaling pathways, both of which important for host antiviral response (Billharz et al. 2009).

Systems biology of influenza infection was also studied in animal models. In order to determine the factors responsible for the increased virulence of the 1918 influenza strain, Kash et al. performed a comprehensive genomic analysis of mice infected with the reconstructed 1918 influenza virus (Kash et al. 2006). Among their findings, they discovered that the increased and accelerated activation of host immune response genes caused by influenza infection were associated with severe pulmonary pathology (Kash et al. 2006). The pathogenicity of 1918 virus was also assessed in primates by gene expression microarrays (Kobasa et al. 2007). In their work,

they showed the severity and outcome of infection by the 1918 virus can be associated to a dysregulation of the antiviral response genes, such as *DDX58* and *IFIH1* (Kobasa et al. 2007).

Infection by the human immunodeficiency virus (HIV) has also been heavily studied by systems biology approaches. A common theme for several of these studies was toward the understanding of HIV disease progression, utilizing for that cohorts of rare individuals who could naturally control HIV infection. These so-called elite controllers, despite detectable viral loads could prevent the progression of the disease for long periods of time even in the absence of treatment (Peretz et al. 2012). Distinct transcriptional profiles were identified in T cell subsets and intestinal mucosal tissue of elite controllers compared to disease progressors (the main findings of these studies were extensively reviewed by Peretz et al. (2012) and Fonseca et al. (2011)).

Genome-scale RNAi screens were also conducted by independent groups to discover the cellular factors associated with HIV replication (Konig et al. 2008; Brass et al. 2008; Zhou et al. 2008). Although many of the cellular factors identified by each group were previously implicated in HIV replication, a surprisingly low overlap was found between any pair of screens (Bushman et al. 2009). Nevertheless, when host factors that participate in similar cellular processes were grouped, the overlap between pair of screens increased (Bushman et al. 2009). This indicates that functional analysis approaches can provide a less biased and less noisy representation of global changes than traditional single gene analysis. Additionally, network analyses were performed by Bushman et al. (2009) and Konig et al. (2008) in order to reveal how the host cellular factors interact with each other and with the HIV-encoded proteins. The interaction networks were elucidated based upon protein-protein binding data derived from databases containing experimental and literature-curated data (e.g. Bind, MINT, Reactome, Y2H, NIAD, etc.) (Konig et al. 2008; Bushman et al. 2009). To further identify densely connected sub-networks within this interactome map, MCODE (Bader and Hogue 2003), a graph theoretic clustering algorithm, was utilized.

Genomic studies of lesser known viruses can indicate potential application of alternative drugs and biologic therapies. Transcriptomic analysis of mice infected with Chikungunya virus, a mosquito-borne alphavirus that causes a chronic debilitating polyarthralgia/polyarthritis in humans, was recently performed by our group (Nakaya et al. 2012). Our goal was to determine whether the inflammatory gene expression signature of Chikungunya arthritis showed any similarities with human rheumatoid arthritis (RA) or with a mouse model of RA. The rationale behind this was to assess whether drugs developed for rheumatoid arthritis might be useful in the treatment of alphavirus-induced arthritis, for which current treatments are often inadequate. First, genes significantly induced or repressed post-infection were identified by a statistical framework described by Ling et al. (2010). Their method relies on an iterative procedure to perform robust estimation of the null hypothesis (assuming that the gene expression background difference is normally distributed), and allowing the identification of differentially expressed genes as outliers (Ling et al. 2010). Next, gene set enrichment analysis (Subramanian et al. 2005) was used to show a highly significant overlap in the differentially expressed genes in the Chikungunya arthritis model and in RA, suggesting that RA, a chronic autoimmune arthritis, and Chikungunya disease, indeed share multiple inflammatory processes (Nakaya et al. 2012).

Finally, a systematic comparison of published transcriptional-profiling data from 32 studies that involved 77 interactions between different host and pathogens (mostly viruses) were

performed by (Jenner and Young 2005). This meta-analysis yielded not only novel and unique insights about microbial pathogenesis, but also revealed a common host-transcriptional-response, comprised of genes that seem to be similarly induced in many different cell types exposed to several pathogen species (Jenner and Young 2005). As expected, this common host response was enriched for genes associated with inflammatory processes, interferon response and genes that are known to activate or limit the immune responses (Jenner and Young 2005). However, the host–pathogen interactions described there, combined with further interactome data can be extremely useful to reconstruct the regulatory networks that underlie the transcriptional response to infection (Jenner and Young 2005).

5.2 Bacterial infections

Each pathogenic bacterium interacts with the host in a unique and specific way. Ramilo et al. addressed this issue by analyzing the gene expression profiles of subjects with acute infections caused by four common pathogens: *Escherichia coli*, *Staphylococcus aureus*, *Streptococcus pneumonia*, and influenza A virus (Ramilo et al. 2007). Applying nonparametric univariate tests, such as Mann-Whitney U or Fisher exact test, they first ranked genes on the basis of their ability to discriminate between groups of patients infected with the same pathogen, and then used a K-Nearest Neighbors method to determine the discriminatory power of the top ranked (i.e. classifier) genes (Ramilo et al. 2007). Using this strategy, they were able to identify discriminative blood transcriptional signatures, and demonstrate the potential clinical application of gene signatures in the diagnosis of infectious diseases (Ramilo et al. 2007).

Banchereau et al. also characterized the whole blood transcriptome of patients infected with *S. aureus* (Banchereau et al. 2012). In this work, however, they examined how the diverse clinical manifestations of the disease impact the blood gene signatures. To examine the heterogeneity in the blood signatures and then identify the factors associated with such variation, the authors calculated a "Molecular Distance To Health" score (MDTH) for each individual patient (Banchereau et al. 2012). The MDTH was introduced by Pankla et al. (2009) while analyzing the blood transcriptomics of patients with septicemic melioidosis (caused by the gram-negative bacillus *Burkholderia pseudomallei*). Figure 15.4 summarizes how the MDTH can be calculated. Using the MDTH score and blood transcription modules, Banchereau et al provided a new understanding of the relation/interaction between host response and clinical disease manifestations to *S. aureus* (Banchereau et al. 2012).

Of the clinically relevant bacterial infections, tuberculosis (TB), caused by *Mycobacterium tuberculosis* is by far among the top ones. The WHO estimated that, in 2011, 1.4 million people died from TB and 8.7 million fell ill. A much greater number of people are infected with *M. tuberculosis* but remain asymptomatic (latent TB). Since current tests fail to identify subjects with latent TB that will develop the disease, Berry et al tried to address if blood transcript signatures could perform this task (Berry et al. 2010). In their work, they defined a distinct 393-transcript signature in patients with active TB that was then applied to predict if latent TB patients, from two independent cohorts would develop active disease (Berry et al. 2010). The k-nearest neighbor class prediction method they used gave sensitivity and specificity rates higher than 60% and 90%, respectively for both test and validation cohorts (Berry et al. 2010). Additionally, they utilized blood co-transcriptional gene modules (Chaussabel et al. 2008) to

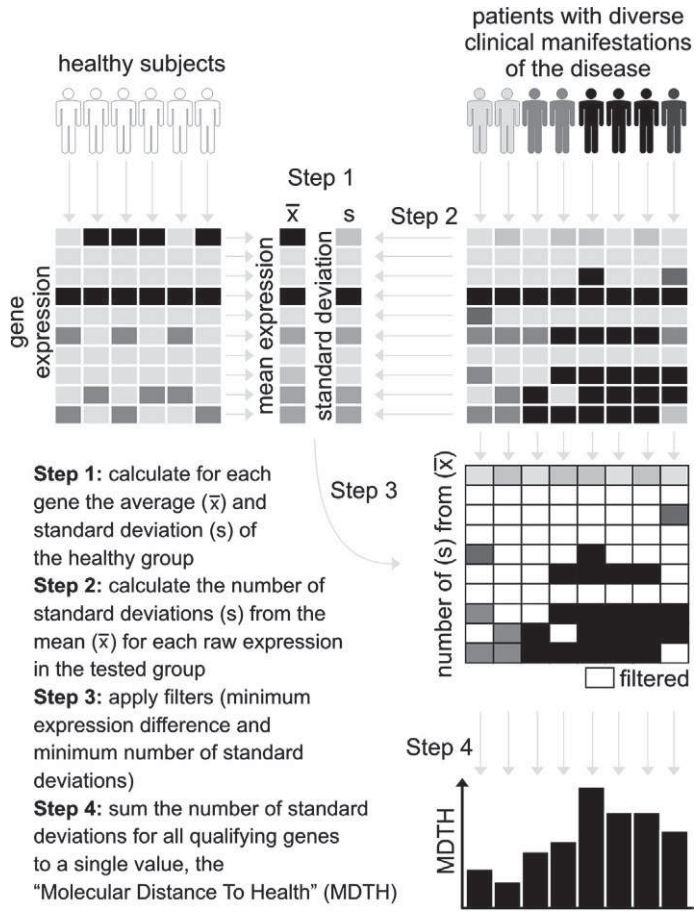

FIGURE 15.4 **Molecular distance to health (MDTH)**. This score measures the global transcriptional perturbation in each patient compared to the median of healthy controls.

identify functional components of the transcriptional host response during active TB (Berry et al. 2010). Combining the modular TB signature with flow cytometry data, they revealed that the TB signature detected a neutrophil-driven interferon-inducible gene profile and that it also includes changes in cellular composition (Berry et al. 2010).

Additionally, Magombedze and Mulder investigated the TB latency dynamics through a meta-analysis with publicly available gene expression data from time course experiments (Magombedze and Mulder 2012). They utilized reverse network engineering techniques to predict gene dependencies and regulatory interactions, and constructed a mathematical model for the inferred gene regulatory networks. Their systems biology approach predicted interesting key TB genes, such as *Rv1370c* and *Rv3131* in the latency/dormancy program that could be targeted as potential latency drug candidates (Magombedze and Mulder 2012). Hecker et al. (2009) and Marbach et al. (2010) nicely reviewed reverse engineering

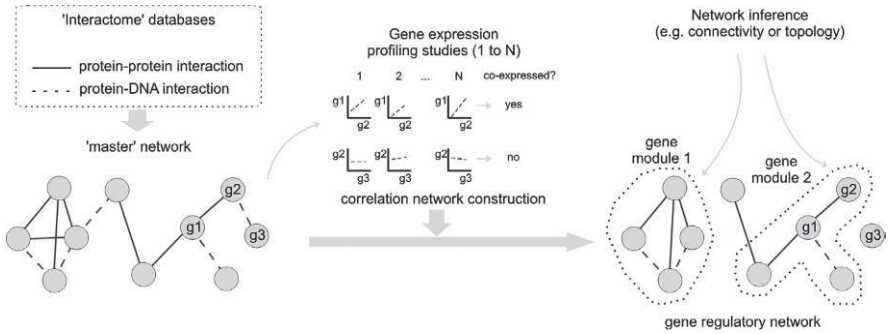

FIGURE 15.5 Example of gene regulatory network reconstruction. Interactome data (e.g. protein-protein, protein-DNA) is used to construct a 'master' network which delineates all the connections that exist between components. These connections are then confirmed on the basis of correlation network analysis from different gene-expression profiling studies, where if the correlation coefficient between two genes is above a given threshold they are predicted to interact. The network can be decomposed into network sub-modules by topology or connectivity.

techniques that can be used to predict unknown interaction networks based on high-throughput gene expression data. Figure 15.5 illustrates a simple example of network reconstruction.

In order to provide insights about different infection strategies of pathogens, Tekir et al. analyzed > 23,000 pathogen–human protein interactions, downloaded from PHISTO search tool (http://www.phisto.org) (Durmus Tekir et al. 2012). In their comprehensive analysis, they observed a higher connectivity of human proteins that are targeted by bacteria and viruses compared to those that are not (Durmus Tekir et al. 2012). Also, while bacteria seem to interact with human proteins associated with immunological defense mechanisms, viruses seem to interact with human cellular processes related to transcriptional machinery (Durmus Tekir et al. 2012).

5.3 Infection caused by other organisms

Malaria is one of the greatest public health problems worldwide and has been extensively studied by systems biology approaches. The transcriptional activity of Plasmodium genome, the parasite protozoan that causes malaria, was assessed during human blood stage by next-generation sequencing (Otto et al. 2010), tiling microrrays (Broadbent et al. 2011), and custom-made Affymetrix chips (Daily et al. 2007). Daily et al. (2007) compared the expression profiles of parasites derived directly from venous blood samples of patients to the ones obtained *in vitro* by other groups. Using a non-negative matrix factorization algorithm, they clustered the *in vivo* expression profiles into three distinct groups, and discovered that only one of them was similar to early ring-stage *in vitro* profiles (Daily et al. 2007). The other two clusters were not similar to those of either early or late *in vitro* stages, and therefore could represent novel transcriptional states (Daily et al. 2007). Their work has shown that the molecular signatures of the parasite *in vivo* can be different from the ones found in *in vitro* models, and that this can influence the development of novel drug targets or alternative therapies for malaria (Daily et al. 2007).

Most host-pathogen studies have focused on monitoring the changes that occur either on the host side or on the pathogen side. Tierney et al approached this differently (Tierney et al. 2012). The authors quantified the gene expression dynamics of both *Candida albicans* (fungal pathogen) and *Mus musculus* (host) during phagocytosis by dendritic cells using RNA-seq (Tierney et al. 2012). In order to generate the inferred regulatory network, they combined the expression kinetics of candidate genes (prioritized with a specific set of criteria) and the putative regulatory interactions obtained from different data sources (and scored based on the "confidence" of the prior knowledge source) (Tierney et al. 2012). As a result, they were able to identify novel interspecies host-pathogen interactions, and validate experimentally some of them (Tierney et al. 2012).

6 SYSTEMS VACCINOLOGY

The emerging field of systems vaccinology represents a stepping stone for understanding how vaccine-induced immunity works. This would be of paramount importance in the rational design of future vaccines (Pulendran et al. 2010), breaking up several centuries of empirically designed vaccines. The major goals are to unravel the molecular mechanisms of vaccines and to identify signatures that correlate or predict vaccine immunogenicity and/or disease protection (Nakaya et al. 2012). These can only be achieved through the use of high-throughput technologies to monitor all the parts of immune system and with computational modeling to analyze the complex interactions between these parts (Nakaya and Pulendran 2012).

Recently, systems vaccinology has been successfully applied to study empirically designed vaccines and is beginning to yield mechanistic insights about immune regulation (Nakaya and Pulendran 2012). However, the successful integration of systems approaches with vaccine development still faces several challenges and potential pitfalls. The biological complexity underlying the immune system itself, and the intrinsic noise, bias, and errors of data generated by high-throughput techniques cause major challenges (Nakaya and Pulendran 2012). Others factors, such as prior immunological history and ethics, represent different challenges that are inherent (and sometimes unique) to this field of research.

6.1 Yellow fever vaccine

The Yellow Fever YF-17D vaccine is a live attenuated virus vaccine, and is one of the most successful vaccines ever developed. A single immunization confers protection in nearly 90% of vaccinees and results in strong T cell response (both cytototoxic T cells and helper T cells) and in robust neutralizing antibody response that can persist for nearly 40 years (Pulendran et al. 2010). Therefore, it is no surprise that YF-17D is among the first vaccines to be studied by systems vaccinology (Gaucher et al. 2008; Querec et al. 2009). In order to obtain biological insights about the mechanism of action of this vaccine, our group (Querec et al. 2009) and others (Gaucher et al. 2008) independently analyzed the gene expression profiles induced

in the blood of individuals a few days after YF-17D vaccination. These analyses revealed innate gene signatures composed of type I interferon, inflammasome, and complement genes (Gaucher et al. 2008; Querec et al. 2009). We also identified two key transcription factors that mediate type I interferon responses, *IRF7* and *STAT1* as being up-regulated post-YF-17D vaccination (Querec et al. 2009). Additional analyses revealed statistically over-represented frequencies of IRF7 transcription factor binding sites in the promoters region of differentially expressed genes, indicating its importance and central role in YF-17D vaccine response (Querec et al. 2009).

The second goal of our study with YF-17D vaccine was to identify early innate signatures that predict adaptive immune outcomes (i.e. neutralizing antibody and CD8+ T cell responses). However, for the YF-17D vaccine, these outcomes were not clearly defined. For instance, what is the minimum level of neutralizing antibody titers induced by the vaccine that is necessary to confer protection against subsequent exposure to the wild-type yellow fever virus? In addition, how much of an increase in cell frequency is considered a "sufficient" CD8+ T cell response? Thus, before proceeding with the predictive analyses, we should first deal with the lack of such well-defined "correlates of protection" for YF-17D vaccine (i.e. the level of a given immune parameter that needs to be reached or exceeded after vaccination in order to assume protection). We addressed this issue by performing unsupervised principal component analyses (using selected genes), which segregated the subjects into two distinct subgroups (Querec et al. 2009). A cutoff of 3% for CD8+ T cell responses and a neutralizing antibody titer cutoff of 170 were used to define the "correlates of immunogenicity" for YF-17D vaccine (Querec et al. 2009).

The challenge was to identify gene signatures that could correctly classify a second and independent YF-17D vaccine trial, which was collected more than 1 year after the first trial. Using the trial 1 as a training set and the trial 2 as a validation set, we tested two different classification methods: the classification to the nearest centroid, or ClaNC (Dabney 2005) and the discriminant analysis via mixed integer programming, or DAMIP (Lee 2007). While both methods correctly classified >80% of the vaccinees according to their CD8+ T cell response, DAMIP program achieved this precision using gene sets comprised of very few genes (called as 'classification rules') (Querec et al. 2009). DAMIP also correctly predicts the magnitude of B cell neutralizing antibody responses with very high accuracy (Querec et al. 2009). These provided a proof-of-concept demonstration that systems biology approaches can indeed be used to predict the immunogenicity of vaccines (Nakaya and Pulendran 2012). It also demonstrated that the DAMIP model is a very powerful supervised-learning classification approach (Lee 2007), that generates classification rules with high prediction accuracy even among small training sets (Brooks and Lee 2010).

Potentially novel mechanistic insights of vaccine action emerged from our study with YF-17D vaccine (Querec et al. 2009). The gene *EIF2AK4* (also known as *GCN2*) was frequently found in the DAMIP classification rules that predict CD8+ T cell responses. When the cell is under certain types of cellular stresses (e.g. amino acid starvation), it phosphorylates GCN2, triggering a signaling pathway that results in the shutdown of housekeeping mRNA genes and the formation of stress granules (Kedersha et al. 2007). Mechanistic studies using *GCN2* knockout mice are being conducted by our group to investigate the role of GCN2 in regulating CD8+ T cell responses (Nair et al., manuscript in preparation).

6.2 Influenza vaccines

Investigating the molecular mechanism of action of influenza vaccines proved to be even more challenging than the YF-17D study. In the United States, there are two types of licensed vaccines for seasonal influenza: trivalent inactivated influenza vaccine (TIV), given by intramuscular injection; and live attenuated influenza vaccine (LAIV), administered intranasally. Each vaccine contains three strains of influenza viruses which are selected each year on the basis of the results of the Global Influenza Surveillance and Response System (more information can be found at http://www.influenzacentre.org). Therefore, depending on the antigenic match between the vaccine and the circulating influenza strains, the vaccine efficacy can significantly vary from one year to the next (Nakaya et al. 2011). Additionally, unlike the subjects used in the YF-17D study, most individuals vaccinated against seasonal influenza have already been exposed to the virus through previous infections or vaccinations (Nakaya et al. 2011). It is unclear how the preexisting levels of antibody in these non-naïve subjects would impact the search for predictors of vaccine immunogenicity.

To determine on a genome-wide scale the expression changes induced by TIV or LAIV vaccination, we did microarray analysis using peripheral blood mononuclear cells (PBMCs) collected from 56 vaccinees on days 0, 3, and 7 after vaccination. Pathway analyses revealed that the expression of genes from inflammasome and antimicrobial pathways was similarly altered by both influenza vaccines (Nakaya et al. 2011). LAIV induced the expression of several interferon-related genes, similar to that observed with the other live attenuated virus vaccine YF-17D. TIV, however, induced a signature that was characteristic of plasma B cell response (Nakaya et al. 2011).

The DAMIP method was used to identify innate gene signatures that predict the antibody responses induced one month after vaccination (Nakaya et al. 2011). To train the program to establish an unbiased estimate of correct classification, we utilized the expression profiles of subjects vaccinated with TIV during the 2008-2009 influenza season (Figure 15.6). A second trial (the test set) consisting of gene-expression profiles of TIV vaccinees from 2007-2008 influenza season was used to evaluate the predictive accuracy of the classification rules identified in the first trial. With this approach, DAMIP model generated signatures with a tenfold cross-validation accuracy of > 90% in the training trial and a "blind prediction" accuracy also > 90% for the testing trial (Nakaya et al. 2011). To validate the utility of these gene signatures in predicting the magnitude of antibody response, we included a third trial comprised of TIV vaccinees from the 2009-2010 influenza season. Again, DAMIP was able to generate signatures that classify with 85% accuracy the vaccinees with high (i.e. four-fold or higher increase in antibody titers 30 days after vaccination) or low (less than four-fold) antibody responses (Nakaya et al. 2011).

To demonstrate that systems vaccinology can be used to generate new hypotheses, we selected one gene from our classification rules, *CAMK4*, for functional confirmation experiments. Using *in vitro* experiments and mice deficient in CAMK4, we demonstrated an unappreciated role of CAMK4 in B-cell responses (Nakaya et al. 2011). Thus, by using a knockout mouse model to validate the signatures identified from vaccinated humans, our study was the first to complete the iterative systems vaccinology cycle (Zak and Aderem 2012).

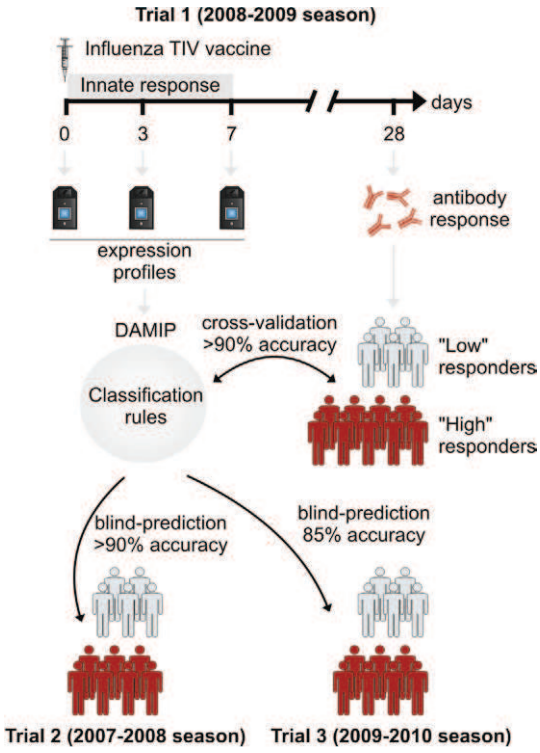

FIGURE 15.6 Signatures that predict the antibody response induced by TIV influenza vaccine. The top part shows the experimental design used in our previous work (Nakaya et al. 2011). The DAMIP method was used to identify innate gene signatures, called 'classification rules' that could predict the antibody responses induced 28 days after vaccination. DAMIP was performed using the 2008–2009 trial as the training set and the 2007–2008 and 2009–2010 trials as the validation sets. The accuracy represents the number of subjects correctly classified as "low" or "high" antibody responders.

Analyses of the global transcriptional response to influenza vaccination were also performed by others (Bucasas et al. 2011; Zhu et al. 2010). Zhu et al. conducted a study with 85 children 12–35 months of age vaccinated either with TIV or LAIV, and found that among the genes induced by the vaccines, several of them were type 1 interferon genes (Zhu et al. 2010). Bucasa et al. also identified interferon-related genes being induced 24 hours post-TIV immunization (Bucasas et al. 2011). Their findings suggest that even an inactivated virus vaccine (TIV) can elicit a type of immune response that is usually activated by innate cells during viral infections (Bucasas et al. 2011).

6.3 Other vaccines

Global gene expression profiling was used to investigate the molecular mechanisms of RTS,S candidate malaria vaccine (Vahey et al. 2010). This vaccine which is administered

in 3 doses at 1-month interval showed that the risk of vaccinated children experiencing clinical malaria is reduced by 56% (Agnandji et al. 2011). The work performed by Vahey et al. assessed the patterns of expressed genes in the PBMCs of 39 volunteers on the day of the third vaccination followed by 24 hours, 72 hours, 2 weeks after vaccination, and on day 5 after challenge (determined by experimental challenge with mosquito-borne *Plasmodium falciparum* malaria) (Vahey et al. 2010). The authors perform class prediction analyses using the "prediction analysis of microarrays with R" tool (PAM-R), which uses a "nearest shrunken centroids" method to identify subsets of genes that best characterize each class (Tibshirani et al. 2002). Using the expression data from day 5 after challenge, PAM-R identified a 393-gene signature that correctly classified all vaccine recipients into protected, delayed in the onset of parasitemia, and not delayed in the onset of parasitemia (Vahey et al. 2010). Next, using gene set enrichment analysis (GSEA, http://www.broadinstitute.org), the authors identified a 32-gene set related to proteasome degradation pathway that was up-regulated in persons in the protected group at 2 weeks after the third vaccination but before challenge (Vahey et al. 2010). The involvement of these genes with the efficient processing of the MHC peptides suggests a potential role of the vaccine in conferring major histocompatibility complex class 1–mediated protection (Vahey et al. 2010).

7 CHALLENGES AND LIMITATIONS

The application of systems biology approaches to any field, especially that of immunology, faces many challenges and potential limitations. These challenges, which have been extensively reviewed elsewhere (Nakaya et al. 2012; Nakaya and Pulendran 2012; Pulendran et al. 2010; Young et al. 2008), are in part related to confounding factors generally found in studying immune responses. These factors can be divided into different levels, from working with highly heterogeneous human cohorts to reconstructing networks using specific experimental data (Figure 15.7). This section will not focus on the inherent problems associated with high-throughput technologies, such as errors, noise, and biases that, if not corrected, may affect downstream analyses (Nakaya and Pulendran 2012). Nor will it focus on the intrinsic variation introduced by the stochasticity of many biological processes, such as transcription initiation, translation, and posttranscriptional regulation mechanisms inside cells. Instead, this section will address some of the challenges associated with analyzing the system behavior of immune responses.

One of the first and most critical steps is the experimental design itself. In general, due to the high costs of performing high-throughput experiments, systems biology studies are comprised of relatively small cohorts and limited number of samples (and time points). In addition to minimal sample sizes for achieving statistical power, the challenge consists of selecting volunteers that will be part of the cohort. The human population, and therefore their immune responses, is incredibly heterogeneous. Of the many factors that could impact someone's immunological outcome, we can cite gender, ethnicity, age, stress status, vaccination history, prior and current disease conditions, genetic background, and baseline levels of immune parameters (Figure 15.7). Therefore, systems biology studies should be performed on cohorts that are relatively uniform with respect to such variables (Nakaya and Pulendran 2012).

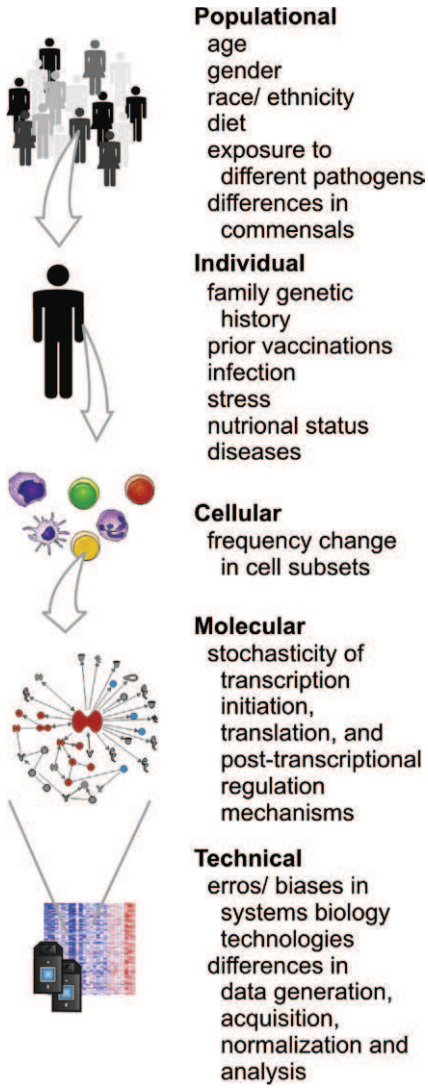

FIGURE 15.7 Potential confounding factors when investigating immune responses at systems-level.

Conversely, if the goal of the study is to examine immune responses that differ with regard to specific age groups (e.g. young adults versus elderly versus infants), or geographical or ethnic origins, or disease states (e.g. healthy subjects versus those with autoimmune diseases or HIV), then the design must carefully control all the other variables (Nakaya and Pulendran 2012).

The choice of time points at which to track the kinects of immune responses can be a very difficult task. Innate responses can be detected a few hours after encountering a pathogen and may be comprised of waves of genes expressed in a tightly regulated time-frame (Granucci et al.

2001). Similarly, during adaptive immune responses, the levels and frequencies of genes and cell types constantly and specifically change over the course of a few days to years (Perelson 2002). Therefore, the choice of time points for an experimental design may considerably impact the applicability of mathematical models created to describe the dynamics of immune responses.

Finally, a better understanding of the functional principles and dynamics of cellular systems can only be achieved with the integration of multiple layers of information, derived. from distinct 'omics' analyses (genomics, transcriptomics, proteomics, interactomics etc.) (Nakaya et al. 2012). However, the integration of genome-scale data types is not a trivial task. Many software packages were developed for this purpose, and some of them were summarized by (Joyce and Palsson 2006). The integration is generally approached by (1) delineating all the connections that exist between cellular components and constructing a 'master' network scaffold; (2) decomposing the network scaffold into network sub-modules; and (3) developing models to simulate and predict the network behavior under specific conditions (Joyce and Palsson 2006). After the many technical, computational, and biological issues are overcome, multi'omics' analyses can provide a bigger picture of the intricate and complex mechanisms of immune responses (Nakaya et al. 2012).

8 CONCLUSIONS

This chapter showed how systems biology can be used to understand and predict the complex mechanisms underlying immune responses. The development of testable predictive models of infectious diseases can potentially lead to improvements in diagnosis and disease therapy. When applied to vaccinology, it could enable the rational design of vaccines or to improvements on existing ones. However, the next big step lies not in the acquisition of additional data but in how to transform it into meaningful knowledge.

With the continuous development of computational and informatics tools to integrate and to analyze 'omics' data, and combined with the large amount of information shared in public databases, systems biology analysis is becoming increasingly accessible to immunologists with no computational training. This may have a profound impact on data interpretation and on the design of innovative experiments. More importantly, it will reduce the gap between discovery-based science and translational science.

References

Abbas, A. R., Baldwin, D., Ma, Y., Ouyang, W., Gurney, A., Martin, F., Fong, S., Campagne, M. V., Godowski, P., Williams, P. M., Chan, A. C., and Clark, H. F. (2005). Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun*. 6:319–331.

Abbas, A. R., Wolslegel, K., Seshasayee, D., Modrusan, Z., and Clark, H. F. (2009). Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. *PLoS One* 4:e6098.

Aderem, A., Adkins, J. N., Ansong, C., Galagan, J., Kaiser, S., Korth, M. J., Law, G. L., McDermott, J. G., Proll, S. C., Rosenberger, C., Schoolnik, G., and Katze, M. G. (2011). A Systems Biology Approach to Infectious Disease Research: Innovating the Pathogen-Host Research Paradigm. *Mbio* 2.

Agnandji, S. T., Lell, B., Soulanoudjingar, S. S., Fernandes, J. F., Abossolo, B. P., Conzelmann, C., Methogo, B. G. N. O., Doucka, Y., Flamen, A., Mordmuller, B., Issifou, S., Kremsner, P. G., Sacarlal, J., Aide, P., Lanaspa, M., Aponte, J.

J., Nhamuave, A., Quelhas, D., Bassat, Q., Mandjate, S., Macete, E., Alonso, P., Abdulla, S., Salim, N., Juma, O., Shomari, M., Shubis, K., Machera, F., Hamad, A. S., Minja, R., Mpina, M., Mtoro, A., Sykes, A., Ahmed, S., Urassa, A. M., Ali, A. M., Mwangoka, G., Tanner, M., Tinto, H., D'Alessandro, U., Sorgho, H., Valea, I., Tahita, M. C., Kabore, W., Ouedraogo, S., Sandrine, Y., Guiguemde, R. T., Ouedraogo, J. B., Hamel, M. J., Kariuki, S., Odero, C., Oneko, M., Otieno, K., Awino, N., Omoto, J., Williamson, J., Muturi-Kioi, V., Laserson, K. F., Slutsker, L., Otieno, W., Otieno, L., Nekoye, O., Gondi, S., Otieno, A., Ogutu, B., Wasuna, R., Owira, V., Jones, D., Onyango, A. A., Njuguna, P., Chilengi, R., Akoo, P., Kerubo, C., Gitaka, J., Maingi, C., Lang, T., Olotu, A., Tsofa, B., Bejon, P., Peshu, N., Marsh, K., Owusu-Agyei, S., Asante, K. P., Osei-Kwakye, K., Boahen, O., Ayamba, S., Kayan, K., Owusu-Ofori, R., Dosoo, D., Asante, I., Adjei, G., Kwara, E., Chandramohan, D., Greenwood, B., Lusingu, J., Gesase, S., Malabeja, A., Abdul, O., Kilavo, H., Mahende, C., Liheluka, E., Lemnge, M., Theander, T., Drakeley, C., Ansong, D., Agbenyega, T., Adjei, S., Boateng, H. O., Rettig, T., Bawa, J., Sylverken, J., Sambian, D., Agyekum, A., Owusu, L., Martinson, F., Hoffman, I., Mvalo, T., Kamthunzi, P., Nkomo, R., Msika, A., Jumbe, A., Chome, N., Nyakuipa, D., Chintedza, J., Ballou, W. R., Bruls, M., Cohen, J., Guerra, Y., Jongert, E., Lapierre, D., Leach, A., Lievens, M., Ofori-Anyinam, O., Vekemans, J., Carter, T., Leboulleux, D., Loucq, C., Radford, A., Savarese, B., Schellenberg, D., Sillman, M., Vansadia, P., and Partnership, S. C. T. (2011). First results of phase 3 trial of RTS, S/AS01 malaria vaccine in African children. *N. Engl. J. Med.* 365:1863–1875.

Amit, I., Garber, M., Chevrier, N., Leite, A. P., Donner, Y., Eisenhaure, T., Guttman, M., Grenier, J. K., Li, W., Zuk, O., Schubert, L. A., Birditt, B., Shay, T., Goren, A., Zhang, X., Smith, Z., Deering, R., McDonald, R. C., Cabili, M., Bernstein, B. E., Rinn, J. L., Meissner, A., Root, D. E., Hacohen, N., and Regev, A. (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science* 326:257–263.

Arkin, A., Ross, J., and McAdams, H. H. (1998). Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* 149:1633–1648.

Bader, G. D., and Hogue, C. W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *Bmc Bioinformatics* 4.

Banchereau, R., Jordan-Villegas, A., Ardura, M., Mejias, A., Baldwin, N., Xu, H., Saye, E., Rossello-Urgell, J., Nguyen, P., Blankenship, D., Creech, C. B., Pascual, V., Banchereau, J., Chaussabel, D., and Ramilo, O. (2012). Host immune transcriptional profiles reflect the variability in clinical disease manifestations in patients with *Staphylococcus aureus* infections. *PLoS One* 7:e34390.

Baum, A., Sachidanandam, R., and Garcia-Sastre, A. (2010). Preference of RIG-I for short viral RNA molecules in infected cells revealed by next-generation sequencing. *Proc. Natl. Acad. Sci. USA* 107:16303–16308.

Bendall, S. C., Simonds, E. F., Qiu, P., Amir el, A. D., Krutzik, P. O., Finck, R., Bruggner, R. V., Melamed, R., Trejo, A., Ornatsky, O. I., Balderas, R. S., Plevritis, S. K., Sachs, K., Pe'er, D., Tanner, S. D., and Nolan, G. P. (2011). Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* 332: 687–696.

Berry, M. P. R., Graham, C. M., McNab, F. W., Xu, Z. H., Bloch, S. A. A., Oni, T., Wilkinson, K. A., Banchereau, R., Skinner, J., Wilkinson, R. J., Quinn, C., Blankenship, D., Dhawan, R., Cush, J. J., Mejias, A., Ramilo, O., Kon, O. M., Pascual, V., Banchereau, J., Chaussabel, D., and O'Garra, A. (2010). An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 466:973–977.

Billharz, R., Zeng, H., Proll, S. C., Korth, M. J., Lederer, S., Albrecht, R., Goodman, A. G., Rosenzweig, E., Tumpey, T. M., Garcia-Sastre, A., and Katze, M. G. (2009). The NS1 Protein of the 1918 Pandemic Influenza Virus Blocks Host Interferon and Lipid Metabolism Pathways. *J. Virol.* 83:10557–10570.

Bolen, C. R., Uduman, M., and Kleinstein, S. H. (2011). Cell subset prediction for blood genomic studies. *Bmc Bioinformatics* 12.

Brass, A. L., Dykxhoorn, D. M., Benita, Y., Yan, N., Engelman, A., Xavier, R. J., Lieberman, J., and Elledge, S. J. (2008). Identification of host proteins required for HIV infection through a functional genomic screen. *Science* 319:921–926.

Broadbent, K. M., Park, D., Wolf, A. R., Van Tyne, D., Sims, J. S., Ribacke, U., Volkman, S., Duraisingh, M., Wirth, D., Sabeti, P. C., and Rinn, J. L. (2011). A global transcriptional analysis of *Plasmodium falciparum* malaria reveals a novel family of telomere-associated lncRNAs. *Genome Biol.* 12:R56.

Brooks, J. P., and Lee, E. K. (2010). Analysis of the consistency of a mixed integer programming-based multi-category constrained discriminant model. *Ann. Operations Res.* 174:147–168.

Bucasas, K. L., Franco, L. M., Shaw, C. A., Bray, M. S., Wells, J. M., Nino, D., Arden, N., Quarles, J. M., Couch, R. B., and Belmont, J. W. (2011). Early Patterns of Gene Expression Correlate With the Humoral Immune Response to Influenza Vaccination in Humans. *J. Infect. Dis.* 203:921–929.

Bullen, A., Friedman, R. S., and Krummel, M. F. (2009). Two-Photon Imaging of the Immune System: A Custom Technology Platform for High-Speed, Multicolor Tissue Imaging of Immune Responses. *Visualizing Immun.* 334:1–29.

Bushman, F. D., Malani, N., Fernandes, J., D'Orso, I., Cagney, G., Diamond, T. L., Zhou, H. L., Hazuda, D. J., Espeseth, A. S., Konig, R., Bandyopadhyay, S., Ideker, T., Goff, S. P., Krogan, N. J., Frankel, A. D., Young, J. A. T., and Chanda, S. K. (2009). Host cell factors in HIV replication: Meta-analysis of genome-wide studies. *Plos Pathog.* 5.

Casadevall, A., and Pirofski, L. A. (2003). The damage-response framework of microbial pathogenesis. *Nat. Rev. Microbiol.* 1:17–24.

Chaussabel, D., Quinn, C., Shen, J., Patel, P., Glaser, C., Baldwin, N., Stichweh, D., Blankenship, D., Li, L., Munagala, I., Bennett, L., Allantaz, F., Mejias, A., Ardura, M., Kaizer, E., Monnet, L., Allman, W., Randall, H., Johnson, D., Lanier, A., Punaro, M., Wittkowski, K. M., White, P., Fay, J., Klintmalm, G., Ramilo, O., Palucka, A. K., Banchereau, J., and Pascual, V. (2008). A modular analysis framework for blood genomics studies: Application to systemic lupus erythematosus. *Immunity* 29:150–164.

Chaussabel, D., Pascual, V., and Banchereau, J. (2010). Assessing the human immune system through blood transcriptomics. *Bmc Biol.* 8.

Dabney, A. R. (2005). Classification of microarrays to nearest centroids. *Bioinformatics* 21:4148–4154.

Daily, J. P., Scanfeld, D., Pochet, N., Le Roch, K., Plouffe, D., Kamal, M., Sarr, O., Mboup, S., Ndir, O., Wypij, D., Levasseur, K., Thomas, E., Tamayo, P., Dong, C., Zhou, Y., Lander, E. S., Ndiaye, D., Wirth, D., Winzeler, E. A., Mesirov, J. P., and Regev, A. (2007). Distinct physiological states of *Plasmodium falciparum* in malaria-infected patients. *Nature* 450:1091–1095.

Durmus Tekir, S., Cakir, T., and Ulgen, K. O. (2012). Infection strategies of bacterial and viral pathogens through pathogen-human protein-protein interactions. *Front Microbiol.* 3:46.

Fell, D. A. (1992). Metabolic control analysis - a survey of its theoretical and experimental development. *Biochem. J.* 286:313–330.

Fischer, N. (2011). Sequencing antibody repertoires the next generation. *Mabs* 3:17–20.

Fonseca, S. G., Procopio, F. A., Goulet, J. P., Yassine-Diab, B., Ancuta, P., and Sekaly, R. P. (2011). Unique features of memory T cells in HIV elite controllers: a systems biology perspective. *Curr. Opin. HIV/AIDS* 6:188–196.

Gaucher, D., Therrien, R., Kettaf, N., Angermann, B. R., Boucher, G., Filali-Mouhim, A., Moser, J. M., Mehta, R. S., Drake, D. R., Castro, E., Akondy, R., Rinfret, A., Yassine-Diab, B., Said, E. A., Chouikh, Y., Cameron, M. J., Clum, R., Kelvin, D., Somogyi, R., Greller, L. D., Balderas, R. S., Wilkinson, P., Pantaleo, G., Tartaglia, J., Haddad, E. K., and Sekaly, R. P. (2008). Yellow fever vaccine induces integrated multilineage and polyfunctional immune responses. *J. Exp. Med.* 205:3119–3131.

Goss, P. J. E., and Peccoud, J. (1998). Quantitative modeling of stochastic systems in molecular biology by using stochastic petri nets. *Proc. Natl. Acad. Sci. USA* 95:6750–6755.

Granucci, F., Vizzardelli, C., Pavelka, N., Feau, S., Persico, M., Virzi, E., Rescigno, M., Moro, G., and Ricciardi-Castagnoli, P. (2001). Inducible IL-2 production by dendritic cells revealed by global gene expression analysis. *Nat. Immunol.* 2:882–888.

Grigoryev, Y. A., Kurian, S. M., Avnur, Z., Borie, D., Deng, J., Campbell, D., Sung, J., Nikolcheva, T., Quinn, A., Schulman, H., Peng, S. L., Schaffer, R., Fisher, J., Mondala, T., Head, S., Flechner, S. M., Kantor, A. B., Marsh, C., and Salomon, D. R. (2010). Deconvoluting post-transplant immunity: cell subset-specific mapping reveals pathways for activation and expansion of memory T, monocytes and B cells. *PLoS One* 5:e13358.

Haining, W. N., and Wherry, E. J. (2010). Integrating genomic signatures for immunologic discovery. *Immunity* 32:152–161.

Hartman, S. E., Bertone, P., Nath, A. K., Royce, T. E., Gerstein, M., Weissman, S., and Snyder, M. (2005). Global changes in STAT target selection and transcription regulation upon interferon treatments. *Genes Dev.* 19:2953–2968.

Hecker, M., Lambeck, S., Toepfer, S., van Someren, E., and Guthke, R. (2009). Gene regulatory network inference: Data integration in dynamic models-A review. *Biosystems* 96:86–103.

Hotson, A. N., Hardy, J. W., Hale, M. B., Contag, C. H., and Nolan, G. P. (2009). The T cell STAT signaling network is reprogrammed within hours of bacteremia via secondary signals. *J. Immunol.* 182:7558–7568.

Houzet, L., and Jeang, K. T. (2011). Genome-wide screening using RNA interference to study host factors in viral replication and pathogenesis. *Exp. Biol. Med.* 236:962–967.

Huang, Y. S., Zaas, A. K., Rao, A., Dobigeon, N., Woolf, P. J., Veldman, T., Oien, N. C., McClain, M. T., Varkey, J. B., Nicholson, B., Carin, L., Kingsmore, S., Woods, C. W., Ginsburg, G. S., and Hero, A. O. (2011). Temporal

dynamics of host molecular responses differentiate symptomatic and asymptomatic influenza a infection. *Plos Genet*. 7.

Hyatt, G., Melamed, R., Park, R., Seguritan, R., Laplace, C., Poirot, L., Zucchelli, S., Obst, R., Matos, M., Venanzi, E., Goldrath, A., Nguyen, L., Luckey, J., Yamagata, T., Herman, A., Jacobs, J., Mathis, D., and Benoist, C. (2006). Gene expression microarrays: Glimpses of the immunological genome. *Nat. Immunol*. 7:686–691.

Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: Systems biology. *Annu. Rev. Genomics Hum. Genet*. 2:343–372.

Jenner, R. G., and Young, R. A. (2005). Insights into host responses against pathogens from transcriptional profiling. *Nat. Rev. Microbiol*. 3:281–294.

Joyce, A. R., and Palsson, B. O. (2006). The model organism as a system: Integrating 'omics' data sets. *Nat. Rev. Mol. Cell Biol*. 7:198–210.

Karlas, A., Machuy, N., Shin, Y., Pleissner, K. P., Artarini, A., Heuer, D., Becker, D., Khalil, H., Ogilvie, L. A., Hess, S., Maurer, A. P., Muller, E., Wolff, T., Rudel, T., and Meyer, T. F. (2010). Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* 463:818–822.

Kash, J. C., Tumpey, T. M., Proll, S. C., Carter, V., Perwitasari, O., Thomas, M. J., Basler, C. F., Palese, P., Taubenberger, J. K., Garcia-Sastre, A., Swayne, D. E., and Katze, M. G. (2006). Genomic analysis of increased host immune and cell death responses induced by 1918 influenza virus. *Nature* 443:578–581.

Kedersha, N., and Anderson, P. (2007). Mammalian stress granules and processing bodies. *Transl. Initiation Cell Biol. High-Throughput Methods Chem-Based App*. 431:61.

Kell, D. B. (2006). Systems biology, metabolic modelling and metabolomics in drug discovery and development. *Drug Disco. Today* 11:1085–1092.

Kingsmore, S. F. (2006). Multiplexed protein measurement: Technologies and applications of protein and antibody arrays. *Nat. Rev. Drug Discov*. 5:310–320.

Kingsmore, S. F., and Patel, D. D. (2003). Multiplexed protein profiling on antibody-based microarrays by rolling circle amplification. *Curr. Opin. Biotechnol*. 14:74–81.

Kitano, H. (2002). Computational systems biology. *Nature* 420:206–210.

Kobasa, D., Jones, S. M., Shinya, K., Kash, J. C., Copps, J., Ebihara, H., Hatta, Y., Kim, J. H., Halfmann, P., Hatta, M., Feldmann, F., Alimonti, J. B., Fernando, L., Li, Y., Katze, M. G., Feldmann, H., and Kawaoka, Y. (2007). Aberrant innate immune response in lethal infection of macaques with the 1918 influenza virus. *Nature* 445:319–323.

Konig, R., Zhou, Y. Y., Elleder, D., Diamond, T. L., Bonamy, G. M. C., Irelan, J. T., Chiang, C. Y., Tu, B. P., De Jesus, P. D., Lilley, C. E., Seidel, S., Opaluch, A. M., Caldwell, J. S., Weitzman, M. D., Kuhen, K. L., Bandyopadhyay, S., Ideker, T., Orth, A. P., Miraglia, L. J., Bushman, F. D., Young, J. A., and Chanda, S. K. (2008). Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell* 135:49–60.

Konig, R., Stertz, S., Zhou, Y., Inoue, A., Hoffmann, H. H., Bhattacharyya, S., Alamares, J. G., Tscherne, D. M., Ortigoza, M. B., Liang, Y. H., Gao, Q. S., Andrews, S. E., Bandyopadhyay, S., De Jesus, P., Tu, B. P., Pache, L., Shih, C., Orth, A., Bonamy, G., Miraglia, L., Ideker, T., Garcia-Sastre, A., Young, J. A. T., Palese, P., Shaw, M. L., and Chanda, S. K. (2010). Human host factors required for influenza virus replication. *Nature* 463:813–817.

Krutzik, P. O., Hale, M. B., and Nolan, G. P. (2005). Characterization of the murine immunological signaling network with phosphospecific flow cytometry. *J. Immunol*. 175:2366–2373.

Kuby, J., Kindt, T. J., Osborne, B. A., and Goldsby, R. A. (2002). *Immunology*. (5th ed.). W. H. Freeman.

Kwissa, M., Kasturi, S. P., and Pulendran, B. (2007). The science of adjuvants. *Expert Rev. Vaccines* 6:673–684.

Le Naour, F., Hohenkirk, L., Grolleau, A., Misek, D. E., Lescure, P., Geiger, J. D., Hanash, S., and Beretta, L. (2001). Profiling changes in gene expression during differentiation and maturation of monocyte-derived dendritic cells using both oligonucleotide microarrays and proteomics. *J. Biol. Chem*. 276:17920–17931.

Lee, E. K. (2007). Large-scale optimization-based classification models in medicine and biology. *Ann. Biomed. Eng*. 35:1095–1109.

Ling, Z. Q., Wang, Y., Mukaisho, K., Hattori, T., Tatsuta, T., Ge, M. H., Jin, L., Mao, W. M., and Sugihara, H. (2010). Novel statistical framework to identify differentially expressed genes allowing transcriptomic background differences. *Bioinformatics* 26:1431–1436.

Lu, P., Nakorchevskiy, A., and Marcotte, E. M. (2003). Expression deconvolution: A reinterpretation of DNA microarray data reveals dynamic changes in cell populations. *Proc. Natl. Acad. Sci. USA* 100:10370–10375.

Lynn, D. J., Winsor, G. L., Chan, C., Richard, N., Laird, M. R., Barsky, A., Gardy, J. L., Roche, F. M., Chan, T. H. W., Shah, N., Lo, R., Naseer, M., Que, J., Yau, M., Acab, M., Tulpan, D., Whiteside, M. D., Chikatamarla, A., Mah, B.,

Munzner, T., Hokamp, K., Hancock, R. E. W., and Brinkman, F. S. L. (2008). InnateDB: Facilitating systems-level analyses of the mammalian innate immune response. *Mol. Syst. Biol.* 4.

Magombedze, G., and Mulder, N. (2012). Understanding TB latency using computational and dynamic modelling procedures. *Infec. Genet. Evol.* 13:267–283.

Marazzi, I., Ho, J. S. Y., Kim, J., Manicassamy, B., Dewell, S., Albrecht, R. A., Seibert, C. W., Schaefer, U., Jeffrey, K. L., Prinjha, R. K., Lee, K., Garcia-Sastre, A., Roeder, R. G., and Tarakhovsky, A. (2012). Suppression of the antiviral response by an influenza histone mimic. *Nature* 483:428–433.

Marbach, D., Prill, R. J., Schaffter, T., Mattiussi, C., Floreano, D., and Stolovitzky, G. (2010). Revealing strengths and weaknesses of methods for gene network inference. *Proc. Natl. Acad. Sci. USA* 107:6286–6291.

McKinney, B. A., Reif, D. M., Rock, M. T., Edwards, K. M., Kingsmore, S. F., Moore, J. H., and Crowe, J. E. (2006). Cytokine expression patterns associated with systemic adverse events following smallpox immunization. *J. Infect. Dis.* 194:444–453.

Miller, J. C., Brown, B. D., Shay, T., Gautier, E. L., Jojic, V., Cohain, A., Pandey, G., Leboeuf, M., Elpek, K. G., Helft, J., Hashimoto, D., Chow, A., Price, J., Greter, M., Bogunovic, M., Bellemare-Pelletier, A., Frenette, P. S., Randolph, G. J., Turley, S. J., Merad, M., and Consortium, I. G. (2012). Deciphering the transcriptional network of the dendritic cell lineage. *Nat. Immunol.* 13:888–899.

Mirsky, H. P., Miller, M. J., Linderman, J. J., and Kirschner, D. E. (2011). Systems biology approaches for understanding cellular mechanisms of immunity in lymph nodes during infection. *J. Theor. Biol.* 287:160–170.

Nakaya, H. I., and Pulendran, B. (2012). Systems vaccinology: Its promise and challenge for HIV vaccine development. *Curr. Opin. HIV/AIDS* 7:24–31.

Nakaya, H. I., Wrammert, J., Lee, E. K., Racioppi, L., Marie-Kunze, S., Haining, W. N., Means, A. R., Kasturi, S. P., Khan, N., Li, G. M., McCausland, M., Kanchan, V., Kokko, K. E., Li, S., Elbein, R., Mehta, A. K., Aderem, A., Subbarao, K., Ahmed, R., and Pulendran, B. (2011). Systems biology of vaccination for seasonal influenza in humans. *Nat. Immunol.* 12:786–795.

Nakaya, H. I., Li, S., and Pulendran, B. (2012). Systems vaccinology: Learning to compute the behavior of vaccine induced immunity. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 4:193–205.

Nakaya, H. I., Gardner, J., Poo, Y. S., Major, L., Pulendran, B., and Suhrbier, A. (2012). Gene profiling of chikungunya virus arthritis in a mouse model reveals significant overlap with rheumatoid arthritis. *Arthritis Rheum.* 64:3553–3563.

Newell, E. W., Sigal, N., Bendall, S. C., Nolan, G. P., and Davis, M. M. (2012). Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8(+) T cell phenotypes. *Immunity* 36:142–152.

Otto, T. D., Wilinski, D., Assefa, S., Keane, T. M., Sarry, L. R., Bohme, U., Lemieux, J., Barrell, B., Pain, A., Berriman, M., Newbold, C., and Llinas, M. (2010). New insights into the blood-stage transcriptome of *Plasmodium falciparum* using RNA-seq. *Mol. Microbiol.* 76:12–24.

Palmer, C., Diehn, M., Alizadeh, A. A., and Brown, P. O. (2006). Cell-type specific gene expression profiles of leukocytes in human peripheral blood. *BMC Genomics* 7.

Pankla, R., Buddhisa, S., Berry, M., Blankenship, D. M., Bancroft, G. J., Banchereau, J., Lertmemongkolchai, G., Chaussabel, D., et al. (2009). Genomic transcriptional profiling identifies a candidate blood biomarker signature for the diagnosis of septicemic melioidosis. *Genome Biol.* 10:R127.

Peng, X. X., Gralinski, L., Armour, C. D., Ferris, M. T., Thomas, M. J., Proll, S., Bradel-Tretheway, B. G., Korth, M. J., Castle, J. C., Biery, M. C., Bouzek, H. K., Haynor, D. R., Frieman, M. B., Heise, M., Raymond, C. K., Baric, R. S., and Katze, M. G. (2010). Unique signatures of long noncoding RNA expression in response to virus infection and altered innate immune signaling. *Mbio* 1.

Perelson, A. S. (2002). Modelling viral and immune system dynamics. *Nat. Rev. Immunol.* 2:28–36.

Peretz, Y., Cameron, C., and Sekaly, R. P. (2012). Dissecting the HIV-specific immune response: A systems biology approach. *Curr. Opin. HIV/AIDS* 7:17–23.

Pulendran, B., Li, S., and Nakaya, H. I. (2010). Systems vaccinology. *Immunity* 33:516–529.

Querec, T. D., Akondy, R. S., Lee, E. K., Cao, W., Nakaya, H. I., Teuwen, D., Pirani, A., Gernert, K., Deng, J., Marzolf, B., Kennedy, K., Wu, H., Bennouna, S., Oluoch, H., Miller, J., Vencio, R. Z., Mulligan, M., Aderem, A., Ahmed, R., and Pulendran, B. (2009). Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nat. Immunol.* 10:116–125.

Ramilo, O., Allman, W., Chung, W., Mejias, A., Ardura, M., Glaser, C., Wittkowski, K. M., Piqueras, B., Banchereau, J., Palucka, A. K., and Chaussabel, D. (2007). Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood* 109:2066–2077.

Ricciardi-Castagnoli, P., and Granucci, F. (2002). Interpretation of the complexity of innate immune responses by functional genomics. *Nat. Rev. Immunol.* 2:881–888.

Rote, N. S. (2009). Pathophysiology: The Biologic Basis for Disease in Adults and Children. In (Kathryn L. McCance, and Sue E. Huethe, eds.), 6th ed. Mosby.

Shaffer, A. L., Rosenwald, A., Hurt, E. M., Giltnane, J. M., Lam, L. T., Pickeral, O. K., and Staudt, L. M. (2001). Signatures of the immune response. *Immunity* 15:375–385.

Shen-Orr, S. S., Tibshirani, R., Khatri, P., Bodian, D. L., Staedtler, F., Perry, N. M., Hastie, T., Sarwal, M. M., Davis, M. M., and Butte, A. J. (2010). Cell type-specific gene expression differences in complex tissues. *Nat. Methods* 7:287–289.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102:15545–15550.

Tibshirani, R., Hastie, T., Narasimhan, B., and Chu, G. (2002). Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc. Natl. Acad. Sci. USA* 99:6567–6572.

Tierney, L., Linde, J., Muller, S., Brunke, S., Molina, J. C., Hube, B., Schock, U., Guthke, R., and Kuchler, K. (2012). An interspecies regulatory network inferred from simultaneous RNA-seq of candida albicans invading innate immune cells. *Front Microbiol.* 3:85.

Vahey, M. T., Wang, Z. N., Kester, K. E., Cummings, J., Heppner, D. G., Nau, M. E., Ofori-Anyinam, O., Cohen, J., Coche, T., Ballou, W. R., and Ockenhouse, C. F. (2010). Expression of genes associated with immunoproteasome processing of major histocompatibility complex peptides is indicative of protection with adjuvanted RTS, S malaria vaccine. *J. Infect. Dis.* 201:580–589.

Watterson, S., and Ghazal, P. (2010). Use of logic theory in understanding regulatory pathway signaling in response to infection. *Future Microbiol.* 5:163–176.

Young, D., Stark, J., and Kirschner, D. (2008). Systems biology of persistent infection: tuberculosis as a case study. *Nat. Rev. Microbiol.* 6:520–528.

Zak, D. E., and Aderem, A. (2009). Systems biology of innate immunity. *Immunol. Rev.* 227:264–282.

Zak, D. E., and Aderem, A. (2012). Overcoming limitations in the systems vaccinology approach: a pathway for accelerated HIV vaccine development. *Curr. Opin. HIV/AIDS* 7:58–63.

Zeiner, G. M., Norman, K. L., Thomson, J. M., Hammond, S. M., and Boothroyd, J. C. (2010). Toxoplasma gondii infection specifically increases the levels of key host microRNAs. *PLoS One* 5:e8742.

Zhou, H. L., Xu, M., Huang, Q., Gates, A. T., Zhang, X. H. D., Castle, J. C., Stec, E., Ferrer, M., Strulovici, B., Hazuda, D. J., and Espeseth, A. S. (2008). Genome-Scale RNAi Screen for Host Factors Required for HIV Replication. *Cell Host Microbe* 4:495–504.

Zhu, W., Higgs, B. W., Morehouse, C., Streicher, K., Ambrose, C. S., Woo, J., Kemble, G. W., Jallal, B., and Yao, Y. H. (2010). A whole genome transcriptional analysis of the early immune response induced by live attenuated and inactivated influenza vaccines in young children. *Vaccine* 28:2865–2876.