Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Gene Reports 5 (2016) 23–29

Contents lists available at ScienceDirect

Gene Reports

journal homepage: www.elsevier.com/locate/genrep

Systems biology approach: Panacea for unravelling host-virus interactions and dynamics of vaccine induced immune response

Sajad Ahmad Wani [a], Amit Ranjan Sahu [a], Shikha Saxena [a], Shahid Hussain [a], Aruna Pandey [a], Sonam Kanchan [a], Aditya Prasad Sahoo [a], Bina Mishra [b], Ashok Kumar Tiwari [c], Bishnu Prasad Mishra [a], Ravi Kumar Gandham [a,*], Raj Kumar Singh [a]

[a] Division of Veterinary Biotechnology, ICAR-Indian Veterinary Research Institute, Izatnagar 243122, India
[b] Division of Biological Products, ICAR-Indian Veterinary Research Institute, Izatnagar 243122, India
[c] Division of Biological Standardization, ICAR-Indian Veterinary Research Institute, Izatnagar 243122, India

ARTICLE INFO

Article history:
Received 4 April 2016
Received in revised form 24 June 2016
Accepted 1 August 2016
Available online 9 August 2016

Keywords:
Systems biology
Holistic
Reductionist
Genomics
Transcriptomics
Proteomics

ABSTRACT

Systems biology is an interdisciplinary research field in life sciences, which involves a comprehensive and quantitative analysis of the interactions between all of the components of biological systems over time. For the past 50 years the discipline of virology has overly focused on the pathogen itself. However, we now know that the host response is equally or more important in defining the eventual pathological outcome of infection. Systems biology has in recent years been increasingly recognised for its importance to infectious disease research. Host-virus interactions can be better understood by taking into account the dynamical molecular networks that constitute a biological system. To decipher the pathobiological mechanisms of any disease requires a deep knowledge of how multiple and concurrent signal-transduction pathways operate and are deregulated. Hence the intricacies of signalling pathways can be dissected only by system level approaches.

Contents

Abbreviation: DNA, deoxyribonucleic acid; RNA, ribonucleic acid; RNA-seq, ribonucleic acid sequencing; SNP, single nucleotide polymorphism; HPI, host pathogen interactions; miRNA, microRNA; PBMC, peripheral blood mononuclear cells.

* Corresponding author.
E-mail address: gandham71@gmail.com (R.K. Gandham).

http://dx.doi.org/10.1016/j.genrep.2016.08.002
2452-0144/

1. Introduction

"The whole is more than the sum of its parts."

[Aristotle.]

Systems biology is a newly advancing field that uses an interdisciplinary approach aimed at understanding and predicting the properties of a living system through systematic quantification of all its components with intensive mathematical and computational modelling. In systems biology each component of the system is measured using high-throughput 'omic' techniques and in theory examined from the cellular level to the whole organism and is thus a holistic rather than reductionist approach. This approach requires the capture and integration of measurements from as many hierarchical levels of information as possible. These can include DNA sequences, RNA and protein measurements, protein–protein and protein–DNA interactions, bio modules, signaling, gene regulatory networks, cells, organs, individuals, populations, and ecologies (Tan et al., 2007).Organisms function in an integrated manner, but biologists have historically studied organisms part by part and celebrated the modern ability to study them molecule by molecule, gene by gene. Biological systems have "emergent properties": that is, their sum is greater than their individual parts, and the biological outcomes of a system cannot be predicted by traditional reductionist methods that study only the individual components (Zak and Aderem, 2009). Regrettably, 'systems biology' is often used as a blanket term, mistakenly referenced in studies utilizing only high-throughput technologies or incorporated into titles to give extra weight or novelty. Integration of multiple high-throughput data types represents just a single dimension of the field necessary to elucidate host responses to infection (Xue and Miller-Jensen, 2012). In the case of systems virology, biological systems may range from virus-infected cells to tissues to whole organisms. Systems biology necessitates a global perspective when investigating infection and this approach to infectious disease research will furnish a greater understanding of the interplay between host and pathogen. Besides, mathematical modelling of interaction networks is essential for researchers to better relate changes at the molecular level to the global properties observed within a biological system during infection. Molecular studies of individual pathways have uncovered many viral host-protein targets; however, it is difficult to predict how viral perturbations will affect the signalling network as a whole (Xue and Miller-Jensen, 2012). Systems biology has the potential to discover novel pro-host therapeutic targets. By providing a more robust overview of the host cellular machinery and its response and interaction with a virus, these kinds of analyses offer inroads toward the development of innovative therapeutics that can act in concert with the host defence mechanism. Mathematical modelling together with documented virus-host interaction data can be used to predict key network components and/or connections (e.g. 'bottlenecks') that can then be assessed by introducing targeted perturbations and monitoring the effects of these changes on the network as a whole. Subsequent analysis of model-based predictions using siRNA knockdown studies or knock-out animal models relate model findings to infection phenotypes and disease outcome in these experimental systems, refining models and driving further predictions (Tisoncik and Katze, 2010). More recently, systems biology is being applied to vaccinology, with the goal of understanding the mechanisms by which vaccines stimulate protective immunity, and predicting the immunogenicity or efficacy of vaccines (Morrison and Katze, 2015; Law et al., 2013; Josset et al., 2013; Chakrabarti et al., 2010; De Backer et al., 2010). Several recent studies have attempted to model the dynamic host immune responses to different viruses (Hancioglu et al., 2007; Tchitchek et al., 2013; Morrison and Katze, 2015; Storey et al., 2005; Liquet et al., 2012; Wang et al., 2009; Law et al., 2013; Josset et al., 2013; Bonneau, 2008). Systems biology in a nutshell can be useful in understanding differences in resistance/susceptibility to a particular disease in different species, identifying disease markers, identifying early markers of infection, prediction of the vaccine efficacy, identifying potential antiviral targets and understanding molecular mechanisms of the host-virus interactions (Chakrabarti et al., 2010; Josset et al., 2013).

1.1. Cybernetics and systems biology modeling

Systems biology finds its genesis in the quantitative modeling of enzyme kinetics, the mathematical modeling of population dynamics, the simulations developed to study neurophysiology, control theory and cybernetics. Norbert Wiener in his book "Cybernetics: or Control and Communication in the Animal and the Machine" defined cybernetics in 1948 as the scientific study of control and communication in the animal and the machine. Cybernetics is the study of human/machine interaction guided by the principle that numerous different types of systems can be studied according to principles of feedback, control, and communications. Mechanical analogies and cybernetic systems dominated the principles guiding systems biology modeling and simulation. Gene regulatory networks, metabolic networks, and signal transduction networks are also part of systems biology discourse and work with mechanical systems analogies. This mechanistic analogy is further combined with control theory in this form of systems biology. Control theory is the mathematical study of how to manipulate the parameters affecting the behavior of a system to produce the desired or optimal outcome. Heinz von Foerster, the founder of Biological Computer laboratory in 1958, is attributed as the originator of second order cybernetics. He is involved in studying similarities in cybernetic systems in biology and electronics and is widely known for his doomsday equitation, which predicted future population growth.

1.2. Reductionism and Holism

Organisms function in an integrated manner. But biologists have historically studied organisms part by part and celebrated the modern ability to study them molecule by molecule, gene by gene. Systems biology is "a new science, a critical science of the future that seeks to understand the integration of the pieces to form biological systems".

[David Baltimore, Nobel Laureate.]

A fundamental tenet of systems biology is that cellular and organismal constituents are interconnected, so that their structure and dynamics must be examined in intact cells and organisms rather than as isolated parts. The last decade has witnessed a backlash against the reductionism of molecular biology. The philosophical antecedents of holism can be traced back to Aristotle, who is said to have pithily observed, "the whole is more than the sum of its parts." The holistic host-directed approach stands in contrast to the more traditional reductionist approaches that focus on a pre-determined small set of molecules (genes, proteins or metabolites). Although often criticized for not being hypothesis-driven, systems-level (or discovery-based) analyses are instead increasingly being acknowledged as potent hypothesis generators. Moreover, for dynamical systems such as those involved in the host response to viral infection, systems-level analyses are considered the only way to understand emergent properties; that is, properties or biological outcomes that cannot be predicted by an understanding of the individual parts of a system alone, but rather only become apparent with knowledge of the specific organization and interactions between components. Because of this, systems virology is an essential and synergistic complement to traditional virology approaches (Law et al., 2013). Methodological reductionism and holism are not truly opposed to each other (De Backer et al., 2010). Each approach has its limitations. Reductionism may prevent scientists from recognizing important relationships between components or organisms in their natural settings, appreciating the evolutionary origins of processes and organisms, grasping probabilistic relationships underlying complicated and seemingly chaotic events, or perceiving heterogeneity

and emergent multilevel properties of complex systems. Holism, on the other hand, is inherently more challenging due to the complexity of living organisms in their environment. Fundamental principles may be difficult to discern within complex systems due to confounding factors like redundancy and pleiotropy. Signal may be swamped by noise. The technology is seductive, but more data do not necessarily translate into more understanding. It is not yet certain whether current approaches to holism, such as systems biology, are adequate to cope with the challenges posed by emergent properties of complex biological systems (Fang and Casadevall, 2011).

1.3. Static studies and dynamic studies

It is possible to divide systems biology studies into two categories: static studies, which take a "snapshot" of a biological network under a single condition, or limited set of conditions; and dynamic studies, which measure time-dependent changes in the network following treatment with environmental stimuli or other biological cues. Both approaches have the potential to significantly increase our understanding of the complex mechanisms involved in viral infection (Xue and Miller-Jensen, 2012). However application of dynamic systems biology approaches to virus-host signaling interactions, in which multiple signals in the network are measured simultaneously over time, may provide a better understanding of how a virus hijacks the host protein signaling network and wires signaling in favor of virus survival and replication. Magnitude and velocity represent emergent properties of the system that are best captured using an expanded time-course. The Magnitude Coefficient (MC) quantifies the magnitude effect as the transcriptomic distance from one biological condition to the condition. The Velocity Coefficient (VC) quantifies the velocity effect as the speed of the transcriptomic host response moving from one time point to the next in the succession of infection. Dynamic studies revealed that it is magnitude and velocity of the host response kinetics, rather than specific sets of up- and down- regulated genes in mutant and wild type H5N1 infected mice (Tchitchek et al., 2013). The analysis of the temporal change of gene expression should help in understanding the complex mechanisms of gene regulation. Gene set analysis methods, which consider predefined groups of genes in the analysis of genomic data, have been successfully applied for analyzing gene expression data in cross-sectional studies. The time-course gene set analysis (TcGSA) is an extension of gene set analysis to longitudinal data (Hejblum et al., 2015). TcGSA is a hypothesis driven method that identifies a priori defined gene sets with significant expression variations over time, taking into account the potential heterogeneity of expression within gene sets. In order to analyze such longitudinal high-dimensional data, several approaches have been suggested including a gene-by-gene statistical analysis (Storey et al., 2005) dimension reduction methods (Liquet et al., 2012) or gene set analysis (Wang et al., 2009).

1.4. Omics data as the basis for systems biology

Systems approaches can be "top-down," starting from "-omics" data and seeking to derive underlying explanatory principles, or "bottom-up," starting with molecular properties and deriving models that can subsequently be tested and validated (Bruggeman and Westerhoff, 2007). The first approach begins with data collection and a description of phenomena, while the latter is more mechanism based, but both produce models of system behaviour in response to perturbation that can be tested experimentally. The first necessary step for a systems biology approach is the gathering of large amounts of data that should be as comprehensive as possible (Fig. 1). Systems biology utilizes and integrates the large amount of data generated by high throughput techniques in order to describe the complex interactions between all parts of a biological system, with the ultimate goal of predicting the behaviour of the system. The genomic information from the host and the pathogen represents the basis for all further molecular analyses and bioinformatic investigations of host pathogen interaction (HPI) systems. Thus, genome sequencing is fundamental. It helps to improve diagnosis, typing of pathogen, virulence and antibiotic resistance detection, and development of new vaccines and culture media. Single nucleotide polymorphism (SNP) typing is important for both identification and characterisation of variants of pathogens (strains, clinical isolates) as well as to study the susceptibility of individuals for certain infections. In contrast to the static information from the genome, the transcriptome reflects the dynamics of HPI systems that results in temporal profiles of gene expression with changes in the scale of minutes and hours.

In the context of systems biology, transcriptomics generates critically analytical data, due to the high precision and relative ease of data generation. Analysis of the transcriptome is presently one of the few "omics" technologies that can be easily performed, and that records all changes, for all annotated, transcribed regions (Oberhardt et al., 2009). This technology allows investigators to conduct studies on virus-host interactions which were previously not possible or affordable, such as the identification of important alternate splice isoforms, miRNA discovery and profiling, and expression profiling in organisms for which a complete genomic sequence has not been determined. RNAseq, a new method for whole transcriptome analysis based on next-generation sequencing technology, offers a much greater dynamic range than microarrays, and therefore a better platform to quantify low-abundance transcripts. In the post-genomic era, genes and the corresponding proteins are studied thoroughly, allowing the identification of intra- and interspecies protein interaction networks. In the last few years, several advanced proteomic techniques have been established providing individual proteome charts of both pathogens and hosts; Epstein- Barr virus (Schellenberger et al., 2010; Thiele and Palsson, 2010), Hepatitis C virus (Lewis et al., 2009; Tripathi et al., 2010; Sonnenschein et al., 2012; Durmus et al., 2015), Influenza A virus (Guirimand et al., 2015), Human Immunodeficiency Virus (Calderone et al., 2015; Qiu et al., 2011), Dengue virus (Murphy, 1985), Measles virus (Lo et al., 2008), Human Respiratory Syncytial Virus (Boedigheimer and Ferbas, 2008).

1.5. Inference based and knowledge based network approaches

Two major network approaches have emerged to extract biological insight from this omics ocean: one is inference based and the other, knowledge based. Both approaches use an interconnected network of biological molecules to interpret omics data; however, there are crucial differences in how the networks are constructed and in the biological questions that can be studied.

1.6. Inference based

Inference-based approaches are based on statistical correlations or information theory, employing statistical methodologies to construct network models from correlation or recurring patterns in omics data (Bonneau, 2008; De Smet and Marchal, 2010). Statistical inference methods benefit from incorporation of all data in an omics set to guide hypothesis development related to unknown interactions. However, these methods are complicated by the fact that the component measurements are not independent and that they do not account for biochemical and genetic causality. A major shortcoming of inference-based methods is that they typically solve underdetermined problems, thus they are not guaranteed to provide a unique solution (De Smet and Marchal, 2010).

1.7. Knowledge based

Knowledge-based approach, which is also referred to as reconstruction based approach is essentially a two-dimensional genome annotation effort (Palsson, 2004) that constructs networks from biochemical and genetic data (Feist et al., 2009; Hyduke and Palsson, 2010). Network

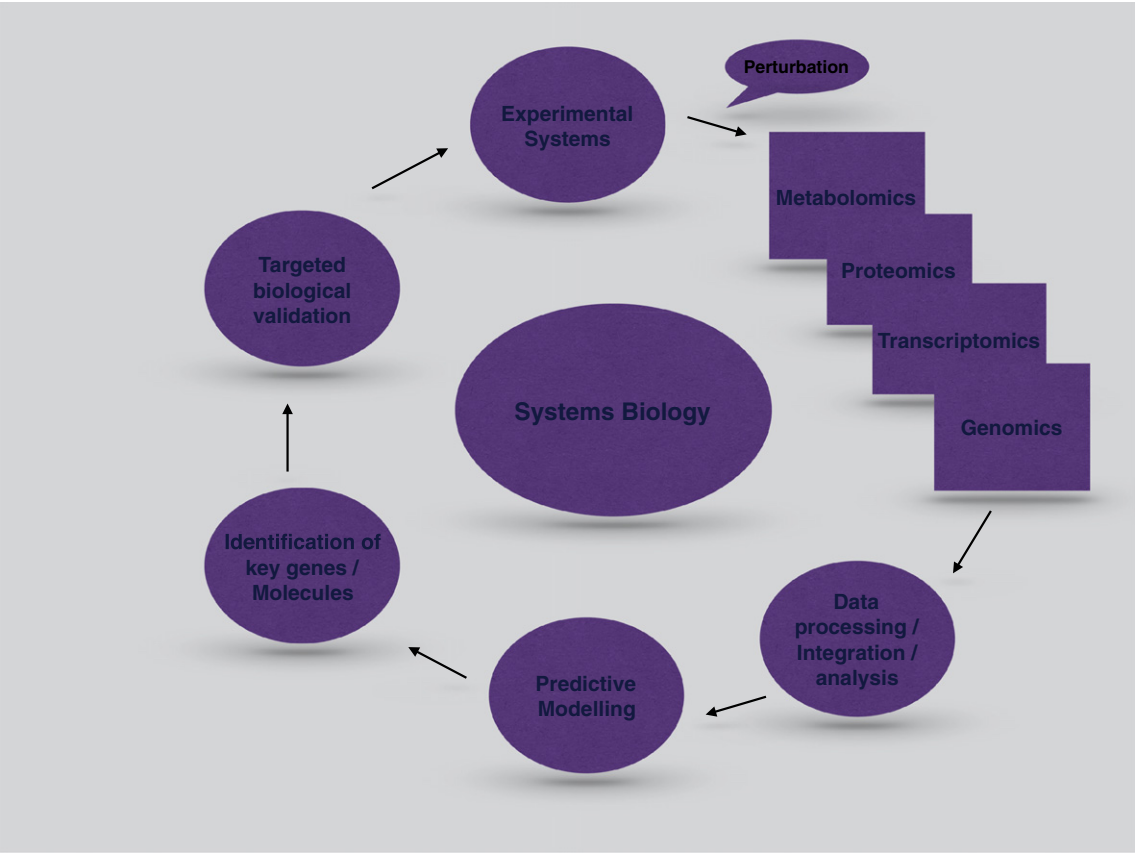

Fig. 1. Systems biology overview.

reconstruction employs established biochemical, genetic, and genomic data (Oberhardt et al., 2009; Schellenberger et al., 2010) to assemble a knowledge base of an organism's molecular components and interactions (Thiele and Palsson, 2010). As knowledge bases are constructed from biological information, they provide a biological context for omics analysis (Lewis et al., 2009). The major shortcoming of the knowledge base approach is that they do not, currently, account for the activities of all genes in a genome, thereby limiting the ability to discover novel relationships important to pathogenesis.

1.8. Modelling virus-host networks

Biological network models have been widely used to improve our understanding of infectious diseases (Mulder et al., 2014). To better understand diseases, systems biology involving iterative cycles, in which model organisms with different levels of complexity are perturbed and then measured using combinations of high-throughput technologies. After mining the multidimensional data, predictive computational models are developed, evaluated and then refined based on the model predictions with new iterations of manipulation of the systems. Virus-host interactions are being catalogued at an increasing rate using protein interaction assays and small interfering RNA screens for host factors necessary for infection. These interactions can be viewed as a network, where genes or proteins are nodes, and edges correspond to associations between them. In order to better understand the molecular network structure and dynamics leading to immune responses, redundancy in these networks, and the relative importance of highly connected ("hubs") as well as sparse regions of the networks ("bottlenecks"), is required to quantify robustness at the level of individual network components (Korth et al., 2013). In addition, computational research will have to focus on developing better approaches for the inference of network dynamics from network topology similar to what has recently been achieved for metabolic networks (Sonnenschein et al., 2012). In addition, it will be critical to develop computational methods and visualisation techniques capable of integrating diverse types of data, in a quantitative manner, and displaying them in a meaningful and understandable fashion. Computational biology has become quite adept at generating giant "hairball" networks, but such visualisations provide too little in the way of helping us understand the biology of the system or in providing targets for future experimentation. Moreover, it is difficult to know whether targeting an individual component of a system—and knowing which one to pick—will impact disease outcome, as there are many examples of compensatory pathways and responses. Nevertheless, progress is being made in this regard, including the development of methods for identifying important genes or proteins (network hubs or bottlenecks) on the basis of topological analysis of protein-protein interaction or inferred networks (McDermott et al., 2009; Diamond et al., 2010) and targets identified by this approach have been successfully validated through gene dysregulation or knockdown (Rasmussen et al., 2011).

At the computational side, systems biology of host pathogen interaction comprises (Durmus et al., 2015) of modelling of molecular mechanisms of infections, viz.

- Modelling of non-protective and protective immune defences against pathogens to generate information for possible immune therapy approaches,
- Modelling of HPI dynamics and identification of biomarkers for diagnosis and for individualised therapy of infections,
- Identifying essential virulence determinants and host factors, and thereby predicting potential drug targets
- Understanding of HPI, in particular the immune system and the immune evasion of the pathogens, as the result of evolutionary long-term adaptation and selection.

1.9. Web-based databases for virus host interactions

Database	Pathogen	Host	Reference
ViRBase	Only viruses	All hosts	Li et al. (2015)
VirHostNet	Only viruses	Animal, human, plant	Guirimand et al. (2015)
VirusMentha	Only viruses	All hosts	Calderone et al. (2015)
HCVPro	Only HCV	Only human	Kwofie et al. (2011)
HIV-1 Human	Only viruses	All hosts	Li et al. (2015)

1.10. Systems approach: dissecting of vaccine-induced immune response

Systems vaccinology has recently emerged from systems biology as an interdisciplinary field that combines systems-wide measurements, networks and predictive modelling in the context of dynamics of vaccine induced immune response. It aggregates the key properties of systems biology which are: perturbation of the system (vaccine administration), monitoring responses at the systems level (transcriptomics, proteomics, and metabolomics), data integration, network modelling and development of predictive rules that describe the system's response to individual perturbations (prediction of vaccine responses). Therefore, its application does not rely simply on data collection from high-throughput techniques, but also on the integration of different types of data in order to generate hypotheses and new insights that may explain the mechanism of vaccines. Systems vaccinology can be divided into 4 essential steps:

1. Measurements of the innate and adaptive responses to vaccination,
2. Determination of vaccine efficacy
3. Systems-level data integration leading to the identification of biomarkers and mechanistic insights
4. Perturbation of the vaccine response in an appropriate experimental system.

Molecular patterns or signatures of genes in the blood after vaccination might predict, latter, development of protective immune responses, representing a strategy to prospectively determine vaccine efficacy. Blood cells provide a snapshot of many lineages and differentiation states within the immune system including the sites of vaccination. Predictive correlates that can be identified prior to vaccination are emerging in systems vaccinology studies. The first systems biological studies to dissect human vaccine-induced responses utilised the yellow fever vaccine, YF-17D (Morrison and Katze, 2015) (Xue and Miller-Jensen, 2012). Both CD8 + T cell and B cell signatures identified in microarray profiles were correlated with protective cell-mediated and antibody responses, thus providing predictive signatures. A gene that is most frequently found in the predictive signatures of the CD8 + T cell response against YF-17D is EIF2AK4. In the case of antibody responses, TNFRSF17, a key gene in the signatures that predict the magnitude of the neutralising antibody response, encodes for a protein, (also known as BCMA), which is a member of the BAFF/BLyS family of receptors, known to promote B cell survival and enhance responses to BCR and TLR signalling. Systems biology studies with influenza vaccines identified modules of genes that were positively correlated with protective immune responses. Interferon responsive genes that were up-regulated at early time points after TIV vaccination positively correlated with robust haemagglutinin inhibition (HAI) titres (Law et al., 2013) (Josset et al., 2013) (Chakrabarti et al., 2010). Transcription factor XBP-1, found to be up regulated, which is necessary for the terminal differentiation of antibody-forming plasma cells, in RNA from sorted B cells, but not from PBMC, after TIV vaccination (Josset et al., 2013; Reimold et al., 2001). Several recent studies have attempted to model the dynamic host immune responses to different viruses (Hancioglu et al., 2007; Tchitchek et al., 2013; Morrison and Katze, 2015; Storey et al., 2005; Liquet et al., 2012; Wang et al., 2009; Law et al., 2013; Josset et al., 2013; Bonneau, 2008). Hyperspectral cytometry is an emerging technology for single-cell analysis that combines ultrafast optical spectroscopy and flow cytometry. Analysis of data generation from next generation cell cytometry experiments is a challenging task given large throughput data generated from analyzing up to 30 or more single cell parameters limiting the technical approaches that rely on manual data gating. To categorise cells into individual types the concept of gating is used (Qiu et al., 2011). A gate is a region, defined in a biaxial plot of two measurements, which is used to select cells with a desired phenotype for downstream analysis. Gates are either manually drawn using software such as FlowJo and FlowCore or automatically defined by clustering algorithms (Murphy, 1985; Lo et al., 2008; Boedigheimer and Ferbas, 2008; Chan et al., 2008). Manual gating being highly subjective relies on the person's knowledge of experiment, the experimental interpretation and heavily relies on expert's knowledge. On the other hand, automatic gating is based on simple logical presumption that cells belonging to one cluster are supposed to be more like than those of other cluster. To apply the concept of automatic gating tools have been developed for analysis of data generation from cytometry of cells and to analyze data for identification of cell phenotypes from cell populations. ACCENSE otherwise known as Automatic Classification of Cellular Expression by Nonlinear Stochastic Embedding is used in exploratory analysis of high dimensional single cell data generated from mass cytometry experiments like CyTOF™ and Fluidgm Corp. ACCENSE classifies cells automatically while retaining single cell resolution and thus circumvents any need for manual gating. ACCEDES has been used in analysis of data of CD8 + T cells derived from SPF mice and categories cells into individual phenotypic subpopulations (Shekhar et al., 2014). Similarly, the SPADE (spanning-tree progression analysis of density-normalized events) algorithm analyse multidimensional flow cytometric data generated from mouse and human bone marrow cells (Qiu et al., 2011). SPADE organizes data into hierarchy of similar phenotypes, hence facilitating analysis of cellular heterogeneity, identification of cell types and comparison of functional markers in response to perturbations. SPADE is essentially quite complementary to the existing approaches in analysis of cytometric data by enabling multiple cell types to be visualized in a tree like format whereby behavior of protein markers across cell types is seen allowing researchers to find their known cell types and discover the unknowns. viSINE is a visualization tool based on t-Distributed Stochastic Neighbor Embedding (t-SINE) has been used to map high dimensional cytometry data onto 2D while also conserving the high dimension structure (Amir el et al., 2013).

1.11. Systems approach-elucidating host-virus interactions

Trying to understand the countless and complex pathogen–host interactions and intra- and inter-cellular signalling events that occur during the course of infectious disease is indeed a formidable task. Systems biology approach revealed that the tissue tropism of West Nile Virus is regulated by antiviral genes and innate immune systems cellular processes. The actions of the RIG-I like receptor (RLR) and type I interferon (IFN) signalling pathways are essential for a protective innate immune response against the emerging flavivirus West Nile virus (WNV). In mice lacking RLR or IFN signalling pathways, WNV exhibits enhanced tissue tropism, indicating that specific host factors of innate immune defence restrict WNV infection and dissemination in peripheral tissues (Bonneau, 2008). A similar approach identified different regulatory networks targeted by Kaposi's sarcoma associated herpes virus (KSHV miR-K12-11) in B cells and endothelial cells. Using a systems biology approach, it was inferred that miR-K12-11 (an ortholog of the human tumour gene hsa-miR-155.) establishes its (gene regulatory networks) GRN by both repressing master TFs and influencing signalling pathways, to counter the host anti-viral response and to promote proliferation and survival of infected cells. The targeted GRNs were more reproducible and informative than target gene identification, and this approach could be applied to other regulatory factors of interest. This systems

approach revealed that miR-K12-11 opposes host defences and contributes to the proliferation and survival of KSHV infected cells by influencing key elements in cellular GRNs like TFs and signalling proteins (Yang et al., 2014). Aevermann et al., 2014 analysed 47 transcriptomic and proteomic datasets from 30 studies generated by SysBio centers and made them available in public domain. These studies characterised the host response to infection by members of the *Orthomyxoviridae* and *Coronaviridae* virus families, including pandemic influenza A H1N1 virus, highly pathogenic H5N1 avian influenza (HPAI) virus, severe acute respiratory syndrome coronavirus (SARS-CoV), and Middle East respiratory syndrome coronavirus (MERS-CoV) (Storey et al., 2005). This collection represented the first coordinated effort to create a systems level description of host-pathogen interactions using multiple viral strains, host models, and '-omics' technologies. Systems biology approaches has been used by Katze laboratory in depth, and their research is focused to define and model virus-host interactions, innate immune signaling, and the varied strategies used by viruses; (Wang et al., 2009; Law et al., 2013; Josset et al., 2013; Cheung et al., 2012; Diamond et al., 2012; Navare et al., 2012; Aderem et al., 2011; Li et al., 2015; Murali et al., 2011; Rasmussen et al., 2011).

1.12. Challenges and barriers

Critics of systems biology call it of being too slow, too expensive and a fishing expedition (Cassman, 2005). Major challenges faced by system biologists' are, computation, data, money, engagement and dedication. The results of systems analysis need to be validated by experiments generating functional data such as protein techniques, gene perturbation, or deficient animal models. Furthermore, it is necessary to have a knowledge base that describes all known interactions at the molecular and cellular level, and that is constantly updated and validated by the scientific community. Finally, systems biology requires multidisciplinary and close collaborative experts including biologists, vaccinologists, immunologists, systems bioinformatics, computational specialists, and clinicians. Nevertheless, the approach is expensive, and with ever-tightening budgets, more money for systems biology means less money elsewhere. Moreover, because the approach has been extensively hyped as being revolutionary, expectations have been set high, and many are understandably disappointed with the pace of progress. Although there is still a long way to go, exciting times in this emerging field are lying ahead.

2. Conclusion

Perturbation analysis is a cornerstone of systems biology. All perturbations result in systems responses, which can be recorded at the level of transcriptome, proteome and metabolome. Viral infection is a systems-level perturbation, and therefore systems biology approaches are naturally suited for studying the complexity of viral host interactions. Systems biology approaches rely on multivariate, context-dependent measurements and computational analysis to elucidate how viral infection alters host cell signalling at a network level. It is necessary to identify the host factors that are required to successfully fight an infection or that cause adverse responses. For this, a highly integrated research strategy is needed to understand all aspects of the complex interplay between the host and invading pathogen. This systematic approach has to go way beyond in vitro cell culture systems and needs to address all aspects of host–virus interactions at the molecular, cellular, organ, and organism level. The challenge now is the need to integrate not only multiple levels of biological data from an individual experiment but also data from different groups for the same assay, and to translate high-throughput data into digested results that can be easily interpreted by a broader audience, including clinicians, governmental regulators and other scientists. This will require close collaborations among virologists, pathologists, clinicians, biologists, statisticians and bioinformaticians, and collaborative consortia and large-scale networked science with partnerships between industry and academia to ensure high quality of the samples, data generation, processing and analysis, as well as ease of data accessibility and interpretation.

> *"Models of the cellular networks—imperfect as they may be— offer a route forward. The union of biology and mathematics may be a shotgun wedding, but may be once its offspring walk and talk, initial misgivings will fade"*
>
> [Colin Macilwain.]

References

Aderem, A., Adkins, J.N., Ansong, C., Galagan, J., Kaiser, S., et al., 2011. A systems biology approach to infectious disease research: innovating the pathogen-host research paradigm. MBio 2 (e00325-00310).

Amir el, A.D., Davis, K.L., Tadmor, M.D., Simonds, E.F., Levine, J.H., et al., 2013. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat. Biotechnol. 31, 545–552.

Boedigheimer, M.J., Ferbas, J., 2008. Mixture modeling approach to flow cytometry data. Cytometry A 73, 421–429.

Bonneau, R., 2008. Learning biological networks: from modules to dynamics. Nat. Chem. Biol. 4, 658–664.

Bruggeman, F.J., Westerhoff, H.V., 2007. The nature of systems biology. Trends Microbiol. 15, 45–50.

Calderone, A., Licata, L., Cesareni, G., 2015. VirusMentha: a new resource for virus-host protein interactions. Nucleic Acids Res. 43, D588–D592.

Cassman, M., 2005. Barriers to progress in systems biology. Nature 438, 1079.

Chakrabarti, A.K., Vipat, V.C., Mukherjee, S., Singh, R., Pawar, S.D., et al., 2010. Host gene expression profiling in influenza A virus-infected lung epithelial (A549) cells: a comparative analysis between highly pathogenic and modified H5N1 viruses. Virol. J. 7, 219.

Chan, C., Feng, F., Ottinger, J., Foster, D., West, M., et al., 2008. Statistical mixture modeling for cell subtype identification in flow cytometry. Cytometry A 73, 693–701.

Cheung, C.Y., Chan, E.Y., Krasnoselsky, A., Purdy, D., Navare, A.T., Bryan, J.T., Leung, C.K., Hui, K.P., Peiris, J.S., Katze, M.G., 2012. H5N1 virus causes significant perturbations in host proteome very early in influenza virus-infected primary human monocyte-derived macrophages. J. Infect. Dis. 206, 640–645.

De Backer, P., De Waele, D., Van Speybroeck, L., 2010. Ins and outs of systems biology vis-a-vis molecular biology: continuation or clear cut? Acta Biotheor. 58, 15–49.

De Smet, R., Marchal, K., 2010. Advantages and limitations of current network inference methods. Nat. Rev. Microbiol. 8, 717–729.

Diamond, D.L., Syder, A.J., Jacobs, J.M., Sorensen, C.M., Walters, K.A., Proll, S.C., McDermott, J.E., Gritsenko, M.A., Zhang, Q., Zhao, R., Metz, T.O., Camp 2nd, D.G., Waters, K.M., Smith, R.D., Rice, C.M., Katze, M.G., 2010. Temporal proteome and lipidome profiles reveal hepatitis C virus-associated reprogramming of hepatocellular metabolism and bioenergetics. PLoS Pathog. 6, e1000719.

Diamond, D.L., Krasnoselsky, A.L., Burnum, K.E., Monroe, M.E., Webb-Robertson, B.J., McDermott, J.E., Yeh, M.M., Dzib, J.F., Susnow, N., Strom, S., Proll, S.C., Belisle, S.E., Purdy, D.E., Rasmussen, A.L., Walters, K.A., Jacobs, J.M., Gritsenko, M.A., Camp, D.G., Bhattacharya, R., Perkins, J.D., Carithers Jr., R.L., Liou, I.W., Larson, A.M., Benecke, A., Waters, K.M., Smith, R.D., Katze, M.G., 2012. Proteome and computational analyses reveal new insights into the mechanisms of hepatitis C virus-mediated liver disease posttransplantation. Hepatology 56, 28–38.

Durmus, S., Cakir, T., Ozgur, A., Guthke, R., 2015. A review on computational systems biology of pathogen-host interactions. Front. Microbiol. 6, 235.

Fang, F.C., Casadevall, A., 2011. Reductionistic and holistic science. Infect. Immun. 79, 1401–1404.

Feist, A.M., Herrgard, M.J., Thiele, I., Reed, J.L., Palsson, B.O., 2009. Reconstruction of biochemical networks in microorganisms. Nat. Rev. Microbiol. 7, 129–143.

Guirimand, T., Delmotte, S., Navratil, V., 2015. VirHostNet 2.0: surfing on the web of virus/host molecular interactions data. Nucleic Acids Res. 43, D583–D587.

Hancioglu, B., Swigon, D., Clermont, G., 2007. A dynamical model of human immune response to influenza A virus infection. J. Theor. Biol. 246, 70–86.

Hejblum, B.P., Skinner, J., Thiebaut, R., 2015. Time-course gene set analysis for longitudinal gene expression data. PLoS Comput. Biol. 11, e1004310.

Hyduke, D.R., Palsson, B.O., 2010. Towards genome-scale signalling network reconstructions. Nat. Rev. Genet. 11, 297–307.

Josset, L., Tisoncik-Go, J., Katze, M.G., 2013. Moving H5N1 studies into the era of systems biology. Virus Res. 178, 151–167.

Korth, M.J., Tchitchek, N., Benecke, A.G., Katze, M.G., 2013. Systems approaches to influenza-virus host interactions and the pathogenesis of highly virulent and pandemic viruses. Semin. Immunol. 25, 228–239.

Kwofie, S.K., Schaefer, U., Sundararajan, V.S., Bajic, V.B., Christoffels, A., 2011. HCVpro: hepatitis C virus protein interaction database. Infect. Genet. Evol. 11, 1971–1977.

Law, G.L., Korth, M.J., Benecke, A.G., Katze, M.G., 2013. Systems virology: host-directed approaches to viral pathogenesis and drug targeting. Nat. Rev. Microbiol. 11, 455–466.

Lewis, N.E., Cho, B.K., Knight, E.M., Palsson, B.O., 2009. Gene expression profiling and the use of genome-scale in silico models of *Escherichia coli* for analysis: providing context for content. J. Bacteriol. 191, 3437–3444.

Liquet, B., Le Cao, K.A., Hocini, H., Thiebaut, R., 2012. A novel approach for biomarker selection and the integration of repeated measures experiments from two assays. BMC Bioinf. 13, 325.

Li, Y., Wang, C., Miao, Z., Bi, X., Wu, D., Jin, N., Wang, L., Wu, H., Qian, K., Li, C., Zhang, T., Zhang, C., Yi, Y., Lai, H., Hu, Y., Cheng, L., Leung, K.S., Li, X., Zhang, F., Li, K., Li, X., Wang, D., 2015. ViRBase: a resource for virus-host ncRNA-associated interactions. Nucleic Acids Res. 43, D578–582.

Lo, K., Brinkman, R.R., Gottardo, R., 2008. Automated gating of flow cytometry data via robust model-based clustering. Cytometry A 73, 321–332.

Morrison, J., Katze, M.G., 2015. Gene expression signatures as a therapeutic target for severe H7N9 influenza — what do we know so far? Expert Opin. Ther. Targets 19, 447–450.

McDermott, J.E., Taylor, R.C., Yoon, H., Heffron, F., 2009. Bottlenecks and hubs in inferred networks are important for virulence in Salmonella typhimurium. J. Comput. Biol. 16, 169–180.

Mulder, N.J., Akinola, R.O., Mazandu, G.K., Rapanoel, H., 2014. Using biological networks to improve our understanding of infectious diseases. Comput. Struct. Biotechnol. J. 11, 1–10.

Murphy, R.F., 1985. Automated identification of subpopulations in flow cytometric list mode data using cluster analysis. Cytometry 6, 302–309.

Murali, T.M., Dyer, M.D., Badger, D., Tyler, B.M., Katze, M.G., 2011. Network-based prediction and analysis of HIV dependency factors. PLoS Comput. Biol. 7, e1002164.

Oberhardt, M.A., Palsson, B.O., Papin, J.A., 2009. Applications of genome-scale metabolic reconstructions. Mol. Syst. Biol. 5, 320.

Navare, A.T., Sova, P., Purdy, D.E., Weiss, J.M., Wolf-Yadlin, A., Korth, M.J., Chang, S.T., Proll, S.C., Jahan, T.A., Krasnoselsky, A.L., Palermo, R.E., Katze, M.G., 2012. Quantitative proteomic analysis of HIV-1 infected CD4+ T cells reveals an early host response in important biological pathways: protein synthesis, cell proliferation, and T-cell activation. Virology 429, 37–46.

Palsson, B., 2004. Two-dimensional annotation of genomes. Nat. Biotechnol. 22, 1218–1219.

Qiu, P., Simonds, E.F., Bendall, S.C., Gibbs Jr., K.D., Bruggner, R.V., et al., 2011. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. Nat. Biotechnol. 29, 886–891.

Rasmussen, A.L., Diamond, D.L., McDermott, J.E., Gao, X., Metz, T.O., Matzke, M.M., Carter, V.S., Belisle, S.E., Korth, M.J., Waters, K.M., Smith, R.D., Katze, M.G., 2011. Systems virology identifies a mitochondrial fatty acid oxidation enzyme, dodecenoyl coenzyme A delta isomerase, required for hepatitis C virus replication and likely pathogenesis. J. Virol. 85, 11646–11654.

Reimold, A.M., Iwakoshi, N.N., Manis, J., Vallabhajosyula, P., Szomolanyi-Tsuda, E., Gravallese, E.M., Friend, D., Grusby, M.J., Alt, F., Glimcher, L.H., 2001. Plasma cell differentiation requires the transcription factor XBP-1. Nature 412, 300–307.

Schellenberger, J., Park, J.O., Conrad, T.M., Palsson, B.O., 2010. BiGG: a biochemical genetic and genomic knowledgebase of large scale metabolic reconstructions. BMC Bioinf. 11, 213.

Shekhar, K., Brodin, P., Davis, M.M., Chakraborty, A.K., 2014. Automatic classification of cellular expression by nonlinear stochastic embedding (ACCENSE). Proc. Natl. Acad. Sci. U. S. A. 111, 202–207.

Sonnenschein, N., Golib Dzib, J.F., Lesne, A., Eilebrecht, S., Boulkroun, S., et al., 2012. A network perspective on metabolic inconsistency. BMC Syst. Biol. 6, 41.

Storey, J.D., Xiao, W., Leek, J.T., Tompkins, R.G., Davis, R.W., 2005. Significance analysis of time course microarray experiments. Proc. Natl. Acad. Sci. U. S. A. 102, 12837–12842.

Tan, S.L., Ganji, G., Paeper, B., Proll, S., Katze, M.G., 2007. Systems biology and the host response to viral infection. Nat. Biotechnol. 25, 1383–1389.

Tchitchek, N., Eisfeld, A.J., Tisoncik-Go, J., Josset, L., Gralinski, L.E., et al., 2013. Specific mutations in H5N1 mainly impact the magnitude and velocity of the host response in mice. BMC Syst. Biol. 7, 69.

Thiele, I., Palsson, B.O., 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat. Protoc. 5, 93–121.

Tisoncik, J.R., Katze, M.G., 2010. What is systems biology? Future Microbiol 5, 139–141.

Tripathi, L.P., Kataoka, C., Taguwa, S., Moriishi, K., Mori, Y., et al., 2010. Network based analysis of hepatitis C virus core and NS4B protein interactions. Mol. BioSyst. 6, 2539–2553.

Wang, L., Chen, X., Wolfinger, R.D., Franklin, J.L., Coffey, R.J., et al., 2009. A unified mixed effects model for gene set analysis of time course microarray experiments. Stat. Appl. Genet. Mol. Biol. 8 (Article 47).

Xue, Q., Miller-Jensen, K., 2012. Systems biology of virus-host signaling network interactions. BMB Rep. 45, 213–220.

Zak, D.E., Aderem, A., 2009. System biology of innate immunity. Immunol. Rev. 227, 264–282.

Yang, Y., Boss, I.W., McIntyre, L.M., Renne, R., 2014. A systems biology approach identified different regulatory networks targeted by KSHV miR-K12-11 in B cells and endothelial cells. BMC Genomics 15, 668.