Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Virology

Systems approaches to coronavirus pathogenesis

Alexandra Schäfer[1], Ralph S Baric[1] and Martin T Ferris[2]

Coronaviruses comprise a large group of emergent human and animal pathogens, including the highly pathogenic SARS-CoV and MERS-CoV strains that cause significant morbidity and mortality in infected individuals, especially the elderly. As emergent viruses may cause episodic outbreaks of disease over time, human samples are limited. Systems biology and genetic technologies maximize opportunities for identifying critical host and viral genetic factors that regulate susceptibility and virus-induced disease severity. These approaches provide discovery platforms that highlight and allow targeted confirmation of critical targets for prophylactics and therapeutics, especially critical in an outbreak setting. Although poorly understood, it has long been recognized that host regulation of virus-associated disease severity is multigenic. The advent of systems genetic and biology resources provides new opportunities for deconvoluting the complex genetic interactions and expression networks that regulate pathogenic or protective host response patterns following virus infection. Using SARS-CoV as a model, dynamic transcriptional network changes and disease-associated phenotypes have been identified in different genetic backgrounds, leading to the promise of population-wide discovery of the underpinnings of Coronavirus pathogenesis.

Addresses
[1] Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, United States
[2] Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill, United States

Corresponding author: Ferris, Martin T (mtferris@unc.edu)

Current Opinion in Structural Biology 2014, **6**:61–69

This review comes from a themed issue on **Viral pathogenesis**

Edited by **Mark Heise**

http://dx.doi.org/10.1016/j.coviro.2014.04.007

Introduction

Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) emerged in Guangdong province, China, in 2002, causing a global epidemic that resulted in about 8000 reported cases and an overall mortality rate of ~10% [1]. The virus was initially present in horseshoe bat populations, and either evolved mutations that allowed transition to Palm Civets and Raccoon Dogs before emerging in human populations, or was directly transmitted from bats to humans and subsequently amplified through intermediate hosts [2–4]. From there, SARS-CoV rapidly spread across the globe, with focal outbreaks in China, Singapore, Vietnam, Taiwan and Canada [1]. More recently, the antigenically distinct Middle East Respiratory Syndrome (MERS-CoV) emerged in 2012 and is still currently circulating in animal and human populations in the Middle East, resulting in 184 cases and 80 deaths to date (http://www.promed.org). MERS-CoV most likely emerged from circulating bat strains and appears to also replicate efficiently in camels [5,6]. Both pathogens cause a respiratory disease, with many severely impacted individuals transitioning into an acute respiratory distress syndrome (ARDS) [7–10]. Although the SARS-CoV outbreak was controlled by epidemiological measures, the recent identification of SARS-like bat-CoVs that can recognize human angiotensin 1 converting enzyme 2 receptors and replicate efficiently in primate cells documents the inevitability of a SARS-CoV-like virus re-emergence event in the near future [11•]. Together, these data highlight prototypical outbreak concerns for the 21st century, where increased travel and community pressures on wildlife areas present numerous opportunities for novel viral disease emergence followed by rapid spread worldwide, sometimes within a matter of months [12–14]. Rapid response platforms are clearly needed to maximize public health preparedness against emerging viruses.

A fundamental problem in dealing with emerging infectious disease control is both the limited accessibility to and the limited number of biological samples associated with an expanding epidemic, confounding insights into susceptibility and mechanistic disease processes which are critical for rational antiviral and vaccine design strategies. In order to advance our understanding of those disease processes at work, novel approaches have been evolved that utilize newly developed state-of-the-art techniques and technologies. Systems biology [15] utilizes an integration of traditional pathogenesis approaches, as well as high-throughput molecular profiling, and computational modeling to identify key host genes and pathways involved in pathogenesis. In a related way [16], systems genetics integrates molecular profiling and pathogenesis readouts within genetically complex populations to identify genes and pathways that contribute to disease variation across genetically diverse populations. Integration of both platforms provides unparalleled power in identifying and studying host susceptibility networks that contribute to disease outcomes. The common feature of both discovery platforms is that they seek to understand viral disease as part of

complex, interacting systems with multiple genes and response pathways. While fundamentally different from standard reductionist strategies, these approaches still rely on standard genetic, molecular biology, biochemical and immunologic strategies to validate the role of targeted genes and networks in disease processes. Using these approaches, there is hope that model systems and platform approaches can be utilized to identify critical regulators of disease across genetically diverse human populations, and to transition these findings into prophylactic and therapeutic drugs.

Systems biology approaches

Over the past decade, a series of important technological advances, genome wide molecular screening platforms and computational strategies have emerged that provide new opportunities for rapid response against newly emerging viral disease threats, globally. The paradigm of these systems biology approaches [15,17] is that (Figure 1) a model system or systems (e.g. tissue culture model, *in vivo* animal model, or even human challenge model and vaccine studies) are perturbed, in our case by viral challenge, preferably resulting in a spectra of disease severities (e.g., lethal vs sub-lethal) to maximize contrast for downstream data mining and modeling. Over a time course, multiple global measures of the system's performance are taken in response to infection, including high-throughput molecular measures (transcriptome, proteome, metabolome, etc.), as well as a variety of virologic, immunologic and pathologic measures (e.g. weight loss, respiratory function, inflammatory response, mortality and histopathological damage). A variety of computation methodologies ([18,19,20••,21•] and reviewed more fully in [22]) and network approaches are then used to *de novo* identify regulatory networks, with these networks and their kinetic responses then being correlated to different disease outcomes in the system. Following these initial descriptions, there are a series of continuing cycles of testing and perturbations (host gene knockout, virus mutant or therapeutic intervention) designed to further validate and then refine the model and to elucidate the mechanistic underpinnings of the systems' performance as a function of infection and disease severity.

Modeling algorithms are rapidly evolving in response to the emergence of these complex and comprehensive systems wide datasets and are beyond the focus of this review (but see [22] for more information); however, many of these approaches *de novo* assemble the networks, independent of annotated pathways or interactions. By allowing this *de novo* assembly within the context of infection, new relationships between genes (or the breaking of previously annotated relationships) emerge that allow for the identification of critical subnetworks. Such a method was recently successfully used to identify critical components of SARS-CoV induced pathogenesis following infection of mice [20••]. A *de novo* assembled network approach was used to identify *Serpine1* and other members of the Urokinase pathway as high priority candidates in regulating severe disease outcomes following lethal vs sub-lethal infections. Subsequent study of *Serpine1* knockouts as well as knockouts from other pathway members confirmed a protective role for these Urokinase pathway members in regulating severe SARS-CoV disease outcomes. Illustrating the power of these *de novo* computational algorithms, it seems unlikely that this pathway would have been otherwise implicated in SARS-CoV infection. These approaches can become even more powerful by integrating analyses across multiple large-scale datasets. Gibbs et al. [19] were able to further refine these approaches by independently assembling transcriptional and proteomic networks and then cross-contrasting these two network types. This method was able to clarify network membership and connections, as well as enhance the relationship between these joint networks and aspects of SARS-induced lung pathology. In addition, such approaches also resulted in highly prioritized list of regulators with conserved behavior for SARS-CoV and influenza A viruses (IAV) via a combined analyses, which provide valuable candidates for downstream experimental validations and therapeutic intervention [21•].

Iterative rounds of perturbation are another key component of the systems biology paradigm. These iterative perturbations are utilized in order to refine and re-evaluate networks when key members of these networks are modified. While perturbations are typically thought of as host perturbations, in some cases they can also be viral perturbations. In this way, SARS-CoV ORF6 [23] was identified as a key inhibitor of multiple antiviral cell intrinsic host genetic responses by blocking the import of targeted clusters of transcription factors into the nucleus during infection and thereby reprogramming host response networks following infection. Chromosome immunoprecipitation studies further validated the role of ORF6 expression in the nuclear import and DNA binding of select transcription factors, and loss of ORF6 attenuated virus pathogenesis. In a parallel example, the SARS-CoV E protein is a known virulence determinant [24]. Using systems biology, E protein was found to suppress the expression of 25 stress related proteins and specifically down-regulated the inositol-requiring enzyme 1 (IRE-1) signaling pathway of unfolded protein responses. In the absence of E protein, an increase in stress responses and the reduction of inflammation likely contributed to the attenuation of rSARS-CoV-ΔE, validating the systems wide predictions. In other cases, contrasting SARS-CoV with immune stimulatory molecules (e.g. interferon stimulation) or different pathogens can be used for cross-comparison. In this way, Danesh et al. [25] were able to show that in contrast to a strict interferon response in a ferret model of SARS-CoV infection, a wider variety of cell migratory and inflammatory genes were induced.