Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Virology

Functional genomics approaches to understand cytomegalovirus replication, latency and pathogenesis

Igor Landais and Jay A Nelson

Cytomegalovirus (CMV) is a species-specific herpesvirus that is ubiquitous in the population and has the potential to cause significant disease in immunocompromised individuals as well as in congenitally infected infants. CMV establishes latency in cells of the myeloid lineage following primary infection. High-throughput functional genomics approaches have provided insight into the mechanisms of CMV replication, but although CMV latency cell models have been useful in elucidating the mechanisms of viral latency and reactivation, omics approaches have proven challenging in these cell systems. This review will summarize the current state of knowledge concerning the use of functional genomics technologies to understand mechanisms of CMV replication, latency and pathogenesis.

Address
VGTI, OHSU West Campus, 505 NW 185th Avenue, Beaverton, OR, 97006, USA

Corresponding author: Nelson, Jay A (nelsonj@ohsu.edu)

Current Opinion in Virology 2013, **3**:408–415

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

Edited by **Michael Katze**

For a complete overview see the Issue and the Editorial

Available online 29th May 2013

1879-6257/$ – see front matter,

http://dx.doi.org/10.1016/j.coviro.2013.06.002

Introduction

Human cytomegalovirus (HCMV) is a β-herpesvirus that is ubiquitous in the population with the highest seroprevalence in individuals in lower socioeconomic conditions [1]. Similar to other herpesviruses, CMV persists in the immune-competent host for life after primary acute infection by establishing a latent state in which only a restricted set of viral genes are expressed and infectious virus is not produced [2•,3•]. Reactivation of the virus from latency is considered to be the major source of virus in seropositive individuals. Although most HCMV infections are asymptomatic, the virus is considered a significant pathogen in immunosuppressed patients, either undergoing immunosuppressive therapy following solid organ or bone marrow transplantation, or with AIDS patients [4]. In addition HCMV is one of the leading viral causes of birth defects [5] and has been implicated in the acceleration of long-term vascular diseases such as atherosclerosis [6].

A major roadblock to our understanding of CMV latency is largely due to the fact that latent infection is the result of complex, intricate and dynamic mechanisms that occur at multiple levels (viral, cellular and organism). To overcome the challenges that the multi-layered complexity of CMV latency poses, novel, non-reductionist and systems-level approaches are needed (Figure 1). In this review, we will assess the contribution of selected functional genomics studies to our current understanding of CMV productive replication and latency, and how systems biology will contribute to solve the remaining challenges.

Functional genomics analysis of CMV replication

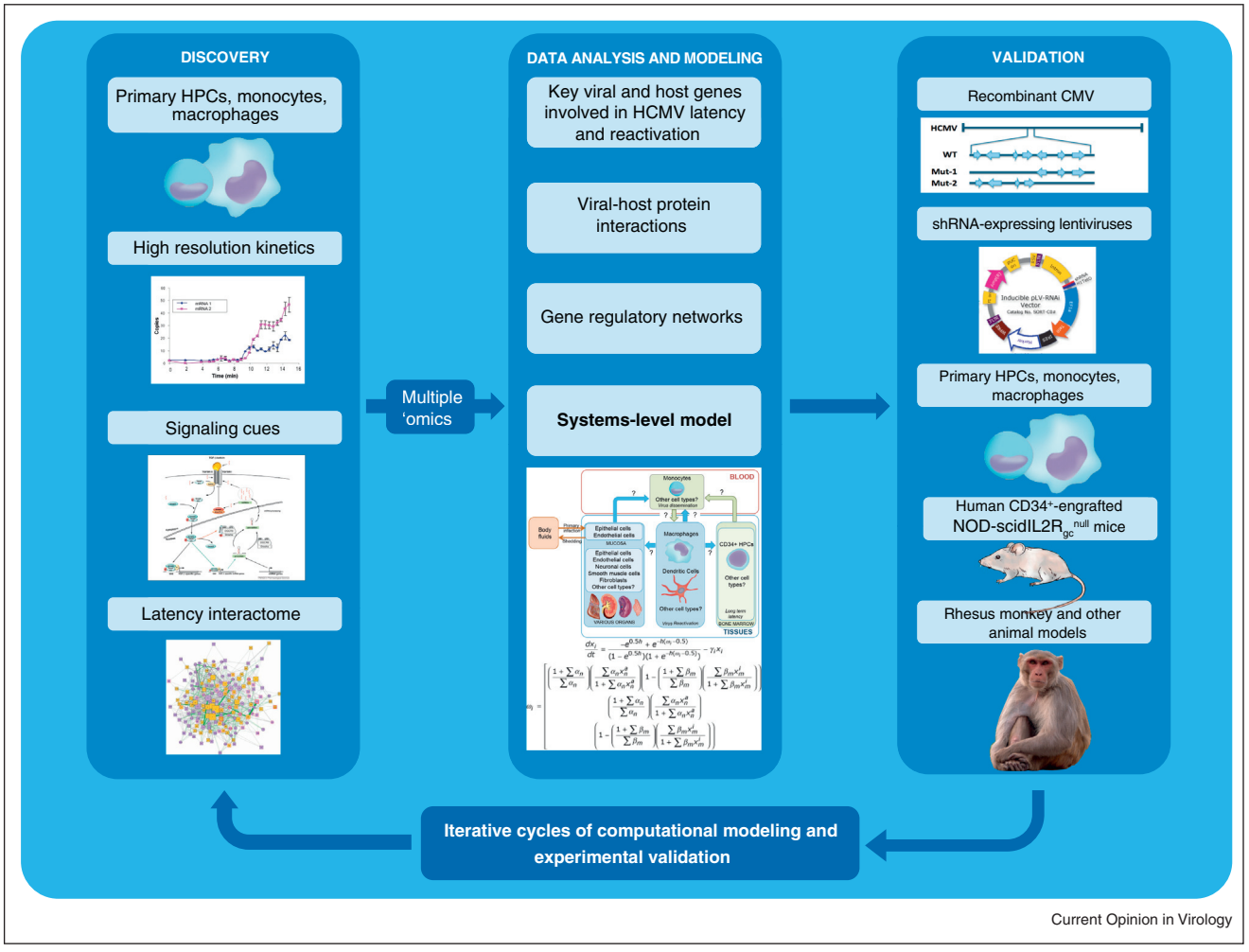
CMV transcriptome

Analysis of the coding potential of cytomegaloviruses has proven to be a remarkably complex task because of the large genome size and the frequent use of alternative transcription initiation sites. Some of the first CMV transcriptomics studies used homemade DNA chips with probes to the 150–200 open reading frames (ORFs) that were predicted from genomics studies [7,8•,9•]. More recent studies using Next Generation Sequencing (NGS) observed that RNA splicing was much more common than previously recognized [10•]. Recently, a sensitive ribosome profiling approach coupled with NGS and bioinformatics was used to analyze the coding potential of HCMV [11••]. This study revealed that translation from the HCMV transcriptome is far more complex than previously anticipated with over 700 viral translation products. Most notably, pervasive use of alternative transcript start sites and alternative splicing allows HCMV to significantly increase viral coding capacity and enables tight temporal control of protein expression [10•,11••]. Interestingly, these recent studies revealed that long non-coding RNAs (lncRNA) and antisense transcripts (AST) together represent >60% of all viral transcripts during productive infection but the function of these non-coding RNAs are unknown.

CMV modulation of the cellular transcriptome

Some of the first studies to examine the effect of HCMV on the cellular transcriptome determined that the virus not only induces the α-interferon response but also alters genes involved in regulation of the cell cycle [12–16]. Other studies have examined the effect of individual HCMV genes on the cellular transcriptome. In one of these early works recombinant HCMV gB was observed to activate cellular genes involved in the Type I interferon response [17]. These observations correlated with

Figure 1



Systems biology approach to CMV latency and reactivation. Systems-level understanding of CMV latency and reactivation requires the use of multiple discovery, modeling and validation platforms. In the initial discovery phase, high-throughput and high-resolution kinetics and interactomics analyses of viral and cellular gene expression are conducted in latently infected and reactivating primary HPCs and monocytes. In the second phase (data analysis and modeling), the large amount of data generated by the discovery platforms are analyzed and used to build a systems-level computational model of CMV latency and reactivation. In the third phase (validation), predictions made by the computational model (such as the identification of viral and cellular factors, virus–host protein interactions and regulatory networks important for CMV latency and reactivation) are tested using a variety of tools including mutant CMV viruses, shRNA-expressing lentiviruses and cell-based or animal infection systems. Iterative cycles of discovery, computational modeling and validation are crucial for continual refinement of the model.

earlier published transcriptome studies [14] and were later explained by the fact that gB binds TLR2 [18]. In another study expression of the transcriptional activator IE72 in cells was observed to activate proinflammatory cytokine genes as well as STAT1, a central mediator of interferon signaling [19].

Recently, a 4-thiouridine (4sU) metabolic labeling approach to tag newly transcribed RNAs coupled with microarray and bioinformatics pathway analysis was used to determine in real-time the transcriptional profile of cellular and viral gene expression during the early phases of productive MCMV infection [20••]. This study observed the upregulation of a cluster of cellular genes involved in immune and inflammatory processes immediately upon virus entry that was followed by a transient DNA damage response and a delayed ER response. All 3 clusters were rapidly and dynamically counter-regulated by viral gene expression. Two clusters of cellular genes were down-regulated in the same time frame, with a rapid repression of genes involved in cell proliferation and differentiation followed by a delayed down-modulation of chromatin assembly and cell cycle genes. Interestingly, promoter analysis revealed that each cluster was targeted by distinct transcription factors such as NFκB and IRF-1 (cluster 1), Elk-1 and YY1 (cluster 2), c-Myc and AP-4 (cluster 3), Mzf1 and AP-2 (cluster 4), as well as E2F and C/EBP (cluster 5). In summary these functional genomics