Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Virology

High-resolution genetic profile of viral genomes: why it matters

Hangfei Qi[1,4], Nicholas C Wu[2,3,4], Yushen Du[1], Ting-Ting Wu[1] and Ren Sun[1,2]

The approval of novel antiviral treatments for hepatitis C virus (HCV) infection provides a great example of research driven medicine in action. However, the emergence of resistant viral strains to existing treatments reminds us of the ongoing challenge that we still face in HCV therapy. What can be done to minimize the health risk posed by viral variants that develop resistance and cause failure of therapy? Here we propose that a high-resolution genetic profiling approach that can assess the function at a single nucleotide/amino acid resolution, may provide a solution. We further discuss the potential application of this methodology in resolving viral resistance through the following three aspects: (1) high-resolution mapping of inflexible regions on the viral genome to identify better drug targets; (2) exhaustive drug resistance profiles to facilitate next-generation drug design; (3) coupled with closely monitoring within-host virus quasi-species, drug resistance profiles can aid in optimized drug combination and personalized medicine in HCV treatments.

Addresses
[1] Department of Molecular and Medical Pharmacology, University of California Los Angeles, Los Angeles, CA 90095, United States
[2] The Molecular Biology Institute, University of California Los Angeles, Los Angeles, CA 90095, United States
[3] Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, United States

Corresponding author: Sun, Ren (rsun@mednet.ucla.edu)
[4] These two authors contribute equally to the manuscript.!

Current Opinion in Virology 2015, **14**:62–70

This review comes from a themed issue on **Engineering for viral resistance**

Edited by **Albrecht von Brunn**

For a complete overview see the Issue and the Editorial

Available online 10th September 2015

http://dx.doi.org/10.1016/j.coviro.2015.08.005

1879-6257/Published by Elsevier B.V.

Introduction

Hepatitis C virus (HCV) infection has emerged as a major cause of human hepatocellular carcinoma (HCC), a major type of liver malignancy, with an estimated 130–150 million people chronically infected globally (2014 World Health Organization; URL: http://www.who.int/mediacentre/factsheets/fs164/en/). Fortunately, HCV antiviral research has entered into a productive era. Since the development of HCV replicon and infectious experimental systems [1–5], tremendous efforts have been devoted to searching for interferon (IFN)-free antiviral therapeutic strategies to avoid the side effects caused by IFN treatment. These efforts have led to the identification of three major classes of direct-acting antivirals (DAAs) targeting the viral non-structural (NS) proteins (NS), including NS3/4A (protease), NS5B (polymerase) and NS5A [6,7]. However, the application of each has been hindered by the emergence of viral resistance.

In 2011, the first two protease inhibitors (boceprevir and telaprevir) were approved in combination with pegylated interferon and ribavirin based therapy as a standard treatment for genotype 1 patients [8–11]. This specific triple-therapy has not been widely adopted due to the emergence of viral resistance and the poor tolerability of the side effects caused by combination therapy of IFN-α [12]. Although a more effective version of protease inhibitor, simeprevir was approved in December of 2013 [13], this improved protease inhibitor displayed an overlapping resistance profile with boceprevir and telaprevir, in which viral susceptibility is significantly diminished by a naturally occurring polymorphism at position 80 of the viral protein [14]. Sofosbuvir is the first approved nucleotide analog that targets the viral polymerase, NS5B (US Food and Drug Administration. http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm377888.htm). This drug is effective against a broad spectrum of HCV genotypes, but the efficacy is significantly reduced by a polymerase substitution S282T [15]. NS5A inhibitors, for example, daclatasvir and ledipasvir, a new family of antivirals against HCV were identified to inhibit NS5A functions through multiple mechanisms [16]. This class of compounds has a potent inhibitory activity against different genotypes of HCV. The resistance barrier to this class of DAA seems to be low. For example, a handful of single mutations were reported both *in vitro* and in *vivo* to confer resistance to the single use of daclatasvir [16–20].

To make the treatment more effective and combat viral resistance, combination therapies were developed. In 2014, Food and Drug Administration (FDA) in the United States approved two DAA combination therapies, Harvoni (combination of sofosbuvir and ledipasvir) (FDA website; URL: http://www.fda.gov/newsevents/newsroom/pressannouncements/ucm418365.htm) and Viekira Pak (combination of Paritaprevir (PI), ombitasvir

(NS5A inhibitor), and dasabuvir (NS5B inhibitor)) (FDA website; URL: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm427530.htm). Despite the breakthroughs in HCV combination therapy, which raised the prospect of curing chronic HCV infection, there are still cases of relapse observed in clinical trials, which is caused by emerging resistant mutations and drug resistant polymorphisms at baseline of treatments [21–26]. The evolution of drug resistance is a painful lesson that we have learned from battles against rapidly mutating microbes and cancer cells, and the emergence of drug resistance may become a major obstacle for eradication of HCV [27]. Here we summarize how the recent development of high-resolution genetics strategies can be applied to analyze the functional residues on viral genomes, and propose how this information will be useful in conquering viral resistance.

Profiling genome flexibility and its significance to drug development

The HCV RNA polymerase NS5B lacks proofreading capacity, and the error rate of genome replication can be between 10^{-5} and 10^{-4} per nucleotide per replication round [28,29]. In combination with high viral load (replication ranging between 10^{10} and 10^{12} virons per day) [30], NS5B creates an extremely large and divergent genetic background for HCV, known as quasispecies. In theory, all sequence variants may exist within a single infected individual [31]. Studies have shown that patients often harbor preexisting drug resistant variants as minor variants within the population, which then can arise and rapidly dominate the population under a drug selection pressure [32]. Clinical observations indicate the emergence of drug resistance within days of initiating therapy with protease inhibitors, suggesting drug-resistant polymorphisms preexist *in vivo* [33]. This genetic plasticity drives the ongoing challenge of antiviral resistance in HCV therapy.

An ideal antiviral drug should maximize the genetic barrier to the emergence of escape mutations by targeting a viral region that has a high fitness cost upon mutation. A systematic approach to analyze the functional consequence of every possible amino acid change at each position on the virus genome to determine the flexibility of the virus genome would be highly beneficial in providing insight for drug development and antiviral resistance. Coupled with functional assays, mutagenesis analysis of viral protein-coding sequences is a powerful method [34••,35] to understand protein functionality within viral replication and potential host interactions. Reverse genetic approaches, such as alanine scanning, which characterizes the phenotypic effects of individual alanine substitutions in a set of residues on a virus genome, is oftentimes limited to a low throughput level [36,37]. Although the concept of evaluating the intrinsic genome flexibility of microorganisms at a systems-level has been a long-term interest in the microbiology field, full characterization of viral genomes on a massive scale has been traditionally limited to low throughput approaches until the advent of high-resolution genetic platforms have emerged.

High-resolution genetic profile to determine viral genome flexibility

The viral assembly process enables linkage of an observed phenotype, viral replication capacity characterized by counting the copy number of viral variants, to a genotype through DNA sequencing of the genome. Therefore, next-generation sequencing technology, which serves both the DNA sequencing of diverse viral populations and the quantification of individual viral variants, is an effective way to characterize every species in a population. The underlying concept of high-resolution genetic profiling approach is to mutagenize the viral genome toward a saturation level, aiming to interrogate every position on the genome, passage the mutant library in any desired selection condition, and quantitate the frequency change of each mutant viral species under the imposed selective pressure to evaluate the phenotype (Figure 1).

Transposon-mediated insertional mutagenesis was established and introduced to studies on small viral genomes to construct mutant libraries and profile essential regions on virus genomes, including HCV, Venezuelan equine encephalitis virus (VEEV), norovirus and influenza A virus [38–43]. For example, an HCV virus genome was subjected to *Mu* transposon mediated insertional mutagenesis to generate a mutant virus library with 15-nucleotide (nt) insertion randomly placed on the virus genome. The mutant virus library was selected in cell culture for genetic selection. The change in frequency of each variant from input to selection output was then calculated resulting in phenotypic determination of each mutant [39,43].

However, an insertion of 15 nt usually disrupts the function of an entire protein domain and gives blunt interrogation of the functional scanning of the proteins. Over the past several years, a series of technologies have greatly advanced DNA sequencing, which has made it practical to functionally scan viral genomes at single amino acid resolution. This technological advance has enabled the mutagenic interrogation into the flexibility of multiple viral genomes to a high resolution, including the human immunodeficiency virus (HIV) [44•] and influenza virus [45,46•,47••,48••].

Identification of effective therapeutic target

In the past decade, the structures of many proteins from different viruses, such as influenza, HIV, and HCV, have been revealed [49–51]. The structural information greatly complements to the increasing amount of high-resolution genetic profiling data within different viruses of relevance to human health. By projecting the fitness data from