Review

Epistatic connectivity among HCV genomic sites as a genetic marker of interferon resistance

James Lara1*, Yury Khudyakov1

1Division of Viral Hepatitis, Centers for Disease Control and Prevention, Atlanta, GA, USA

*Corresponding author e-mail: jlara@cdc.gov

Until recently, the standard-of-care therapy of patients with HCV infection involves treatment with interferon (IFN) and ribavirin (RBV). Host demographic and genetic factors as well as HCV genetic heterogeneity have been shown to be associated with outcomes of therapy. Although resistance to IFN/RBV remains an important clinical and public health problem, there are no reliable genetic markers for the prediction of the therapy outcomes. Recently, it was shown that adaptation to IFN, a major constituent of the host innate immunity, is reflected in the HCV genetic composition and epistatic connectivity among polymorphic genomic sites, thus providing novel genetic markers of IFN resistance. Consideration of coordinated evolution among HCV genomic sites allows for identification of these genetic markers from short regions of the HCV genome and for accurate prediction of therapeutic outcomes. HCV genomic co-evolution offers a general framework for the detection of predisposition to IFN resistance, and possibly to resistance to direct-acting antivirals.

Introduction

HCV belongs to the family Flaviviridae and contains a positive-sense, single-stranded RNA genome of approximately 9,600 nucleotides long. The HCV genome encodes a polyprotein that is processed into three structural proteins (core, E1 and E2) and seven non-structural proteins (P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B) [1]. HCV is the major aetiological agent of blood-borne non-A, non-B hepatitis [2]. Approximately 60–85% of HCV-infected patients fail to clear the virus and develop chronic hepatitis C (CHC). CHC is a risk factor for the development of cirrhosis and liver cancer [3]. At present, there is no vaccine against HCV.

The current standard-of-care therapy for CHC involves a 24- or 48-week course of treatment with pegylated interferon (IFN)-α combined with ribavirin (RBV) [4,5]. Because only 40–80% of CHC patients develop sustained virological response (SVR) to this treatment [4–6] and because patient intolerance to such therapy is common, knowing the therapeutic outcome before initiation of treatment is a major patient management objective. Recently, the IFN/RBV therapy was supplemented with HCV protease inhibitors, telaprevir or boceprevir [7]. Effectiveness of treatment with these direct-acting antiviral agents is reduced, however, in null-responders to IFN/RBV therapy, especially among patients infected with HCV subgenotype 1a [8]. As resistance to IFN is relevant to CHC therapy, detection of the resistance is clinically important. Additionally, because IFN is a main component of innate immune system [9], IFN resistance essentially defines HCV virulence; therefore, monitoring for IFN resistance in HCV populations becomes a public health objective.

Several factors are associated with HCV susceptibility to IFN/RBV treatment. HCV genotype is one of the most important factors. There are six major HCV genotypes (1–6), with genotype 1 being the most prevalent genotype worldwide [10]. Patients infected with genotype 2 achieve SVR in 70–80% of cases after treatment with IFN/RBV [5,11]. By contrast, only 40–52% of genotype-1-infected patients achieve SVR [6,11]. The dependence of IFN/RBV response rates on genotype suggests that the HCV genetic composition is an essential factor affecting therapy outcome. Host factors have also been implicated. Genetic polymorphisms in the host, for example, in the IL28B locus, play a role in defining the rate of spontaneous clearance [12] and IFN/RBV SVR [13,14]. In addition, the patient’s gender, ethnicity and age have been observed to be associated with variation of responses to IFN/RBV [15].

Because innate immunity is one of the major host selection pressures, HCV has evolved multifaceted mechanisms to avert IFN actions [16], with HCV lineages
varying in their capacity to withstand extraneous IFN. The complexity of HCV interaction with host immunity significantly impedes understanding of what underlies susceptibility to IFN, which in turn hampers discovery of reliable genetic markers of IFN resistance and development of diagnostic assays for the detection of resistance.

**HCV genetic heterogeneity and IFN/RBV resistance**

The mechanisms of IFN or RBV action against HCV are not fully understood. Studies have shown that treatment with IFN activates the host’s innate antiviral immune responses by inducing IFN-stimulated genes [17,18]. For RBV, several mechanisms of therapeutic action have been proposed, including viral mutagenesis induction [19], inhibition of inosine monophosphate dehydrogenase [20], mRNA methylation inhibition [21] and facilitation of Th-1-mediated immunoreponses [22]. Feld and colleagues [23] recently reported that combined treatment with this nucleoside analogue improves early responses to IFN, thus supporting its role in enhancing IFN signalling [24] and emphasizing the leading role of IFN in combination therapy.

IFN-inducible, double-stranded, RNA-activated protein kinase R (PKR) is known to be involved in the IFN-induced antiviral response [25]. Two HCV protein domains, the PKR-eIF2α phosphorylation homology domain (PePHD) of the E2 protein and the PKR-binding domain located in the C-terminal region of NS5A, have been implicated in binding to PKR [26]. Genetic heterogeneity within these domains was suggested to be related to IFN resistance. The E2 PePHD of IFN-resistant strains were observed to share greater similarity to the autophosphorylation sites of PKR and the phosphorylation site of the translation initiation factor (eIF2α) than variants susceptible to IFN [26]. This similarity was greater for HCV genotype 1 than genotypes 2 or 3. However, these findings have not been supported in other studies [27,28], indicating that PePHD is not a reliable marker of IFN resistance.

Sequence analysis of HCV genotype 1b showed that a 40-amino-acid stretch in NS5A, known as the IFN-sensitivity-determining region (ISDR), was associated with response to IFN therapy. The number of mutations in ISDR was found to correlate with response [29–31], suggesting that ISDR may be used as a genetic marker of IFN resistance. However, these findings were refuted in other studies [32,33]. Castelain and colleagues [34] observed no binding between PKR and genotype 3a NS5A from IFN-resistant HCV strains. Moreover, intrahost HCV variants sharing the ISDR sequence were found to display differential susceptibility to IFN/RBV treatment [30]. Similar observations for HCV quasispecies sharing PePHD sequence [29] suggest that genetic diversity at regions other than ISDR and PePHD may be responsible for IFN resistance.

Particular variations at other genomic regions have been reported to have association with differences in IFN/RBV susceptibility among HCV strains [16]. Specific mutations in the E2 [35], core [29,36,37], NS2 [29,35], NS5A [29,35,38], NS5B [16,35] and p7 [35,39] were found to correlate with response to IFN/ RBV therapy. Although detectable, these genetic associations were inconsistent among HCV strains. Changes in mutation rates corresponding to IFN response were observed in some regions of the HCV genome, including E2 [40,41], NS5A [31,36], P7 [39] and NS2 [42]. However, no strong linkage between the HCV genetic diversity and IFN/RBV sensitivity has been detected [29,32,33,36,38]. All these observations indicate a complex association between HCV genetic variability and response to IFN/RBV treatment. Resistance to IFN is most probably defined by variation at many HCV genomic sites rather than by specific mutations in a particular genomic region. Such a genetic basis of IFN resistance hinders identification of HCV genomic markers for prediction of outcomes of IFN/RBV therapy.

**Coordinated evolution of the HCV genome**

Host selection pressures shape HCV evolution. Since innate immunity exerts one of the most intensive selection pressures on viruses, HCV evolution is significantly defined by IFN actions. Therefore, adaptation to IFN should be reflected in the HCV genetic composition and epistatic connectivity among polymorphic genomic sites. Epistasis plays a very important role in viral evolution [43,44] and development of drug resistance [45]. Pervasive epistatic connectivity among viral genomic sites is frequently detected in the form of compensatory substitutions [45,46]. This widespread connectivity is organized into a network of coordinated substitutions, reflecting extensive global co-evolution among sites across the entire HCV genome [47]. The network has a very specific topology, with only a few sites being connected to many other sites and many sites having only a few connections.

Global coordination among HCV sites indicates that selection pressures acting at one site are distributed to other sites according to their degree and strength of connectivity in the network. Thus, selection affecting sites with many connections should have a much greater effect on the network state than selection at sites with very few connections. Since several HCV proteins are involved in mitigating IFN-mediated antiviral activity [41], it should be expected that co-evolution among many sites across the entire HCV genome is linked to selection pressures exerted by IFN. Indeed, recent studies showed that the topology of networks of coordinated substitutions is different for HCV strains that
are resistant or sensitive to IFN/RBV treatment [48] and co-evolution among sites in each HCV protein is associated with IFN response [49].

Co-evolution and genetic markers of IFN/RBV response

Genetic analysis of IFN/RBV resistance is frequently focused on host or viral markers separately. It seems reasonable to assume that since the host provides the environment for viral replication, it should play a dominant role in controlling outcomes of HCV infection. Accrued data strongly support the essential contribution of host genetic variation to outcomes of IFN/RBV therapy [15]. There is also an equally large body of evidence supporting the crucial role of HCV genetic heterogeneity in overcoming antiviral actions of extraneous IFN [29–31,36,37,39–42]. However, it is only recently that the integrative framework of host–virus interactions in the form of co-evolution among HCV genomic sites was considered for the identification of genetic markers of IFN/RBV response [48–50].

As aforementioned, coordinated evolution among HCV genomic sites reflects many host selection pressures, including innate and adaptive immune responses [47–50], thus associating host with viral genetic factors. Analysis of individual genomic sites without consideration of their interrelationships impedes detection of this association. There is no one single mutation that is responsible for IFN resistance. Rather, genome-wide coordination among HCV sites is strongly linked to IFN response and can serve as a complex genetic marker of the resistance [49].

Examination of the entire HCV genome to assess the state of genome-wide coordination is clearly impractical in clinical and public health settings. However, the extensive epistatic connectivity discovered among HCV genomic sites [47–49] suggests that co-evolution among sites in any genomic region reflects selection pressures acting on the entire HCV genome in an infected host. Indeed, it was recently reported that each HCV protein [48,49] or a small genomic region, such as hypervariable region 1 in the E2 gene and a segment of the NS5A gene [50], has strong association with IFN resistance.

Although all HCV proteins were found to contribute sites into the network of coordinated substitutions associated with IFN response, the topological properties of sites differed among proteins [49] and, therefore, their involvement in IFN resistance. Two proteins, E2 and NS5A, contributed approximately 40% of all sites and approximately 62% of all links to the HCV genome-wide network and effectively defined the state of the entire network [49]. The full-length sequences [49] or small regions [49,50] of these proteins were used to predict outcomes of IFN/RBV therapy with 83–90% accuracy. Interpretation of epistatic connectivity within these two regions into IFN/RBV resistance was achieved using mathematical models that associated various combinations of nucleotides or amino acids at the specifically selected HCV sites with therapy outcomes.

Conclusions

The HCV IFN resistance is a complex phenotypic trait defined by many host and viral factors. Contribution of these factors to the outcome of IFN/RBV therapy can be mathematically modelled based on coordinated substitutions within the HCV genome. Co-evolution among HCV sites offers novel genetic indicators of IFN/RBV resistance, which nonetheless are detectable using specially devised computational models. Complex genetic markers comprising more than one site are commonly formulated as ‘signatures’ or ‘profiles’. Although assuming interrelationships among sites, such formulation obscures the underlying epistatic connectivity and encourages the search for the combination of specific nucleotides or amino acids at certain genomic or protein positions that predict the desirable phenotypic trend. Because predisposition to IFN resistance is encoded in the complex genome-wide interactions among numerous sites, a vast number of possible genomic signatures can be expected. Identification of all signatures is a daunting task, but integration of all relevant signatures may be achieved by formulation of IFN resistance as a function of epistatic connectivity. Mathematical modelling of this connectivity presents a valuable opportunity for accurate prediction of the IFN/RBV therapy outcomes [49,50].

When possible, reduction of signatures to a single site presents the best option for molecular detection in view of the simplicity of genetic testing. However, detection of interrelationships among several genomic sites suitable for the accurate prediction of IFN resistance should not be any more complex than detection of polymorphisms at a single site. Recent findings indicate that sequencing of short regions of the HCV genome provides sufficient information for the prediction of therapy outcomes [50]. Complexity of the sequence interpretation into IFN resistance requires the application of accurate mathematical models that translate interrelationships among genomic sites into novel genetic markers.

Disclosure statement

The authors declare no competing interests.

References


Accepted 25 May 2012; published online 7 December 2012