Hepatitis C Virus Genetic Variability and Quasispecies

Nahida Sultana¹, Mohammad Razib Ahsan², Afsana Anwar Miti³, Umme Shahera⁴, Ridwana Asma⁵ Tahmina Akther⁶, Ruksana Raihan⁷, Mamun Al Mahtab⁸, Sheikh Mohammad Fazle Akbar⁹

Abstract

Hepatitis C virus (HCV) exists as a cloud of closely related sequence variants called a quasispecies, rather than as a population of identical clones. To date there is no preventive vaccine and though antiviral therapy has been improved in the past few years the HCV cannot be eradicated in all patients as a result of its quasispecies nature due to lack of proof reading activities and high error rate of RNA-dependent RNA polymerase and the pressure exerted by host immune system. This review focuses on the genetic diversity and quasispecies nature of HCV viral genomes, and briefly reviews the principles of quasispecies dynamics and the differences with classical population genetics and discusses the biological implications of this phenomenon, focusing on the hepatitis C virus.

Keywords: Quasispecies; hepatitis C Virus; RNA; evolution; genetic diversity; compartmentalization; drug resistance.

Introduction

Hepatitis C virus (HCV) has infected over 185 million people worldwide and creates a huge disease burden due to chronic, progressive liver disease. HCV is a single stranded, positive sense, RNA virus, member of the Flaviviridae family.¹ HCV replication via RNA-dependent RNA polymerase is very error-prone and generates mutations at an estimated rate of $10^{-5}$ mutations per nucleotide per replication. This high mutation rate is the main cause of the virus's genetic diversity. The high error rate of RNA-dependent RNA polymerase and the pressure exerted by the host immune system, has driven the evolution of HCV into 7 different genotypes and more than 67

1. Research Assistant, Dept. of Virology, Bangabandhu Sheikh Mujib Medical University (BSMMU), Dhaka, Bangladesh.
2. Junior Consultant, Cardiology, District Sadar Hospital, Brahmanbaria, Bangladesh.
3. Medical officer, Dept. of Virology, Bangabandhu Sheikh Mujib Medical University (BSMMU), Dhaka, Bangladesh.
4. Medical Officer, Dept. of Virology, Bangabandhu Sheikh Mujib Medical University (BSMMU), Dhaka, Bangladesh.
5. Research Assistant, Dept. of Virology, Bangabandhu Sheikh Mujib Medical University (BSMMU), Dhaka, Bangladesh.
6. Medical officer, Dept. of Virology, Bangabandhu Sheikh Mujib Medical University (BSMMU), Dhaka, Bangladesh.
7. Associate Professor, Dept. of Microbiology US Bangla Medical College, Narayanganj, Bangladesh.
8. Associate Professor, Dept. of Virology, Bangabandhu Sheikh Mujib Medical University (BSMMU), Dhaka, Bangladesh.
9. Principal Investigator, Toshiba General Hospital, Tokyo, Japan.

Correspondence: Dr. Nahida Sultana. e-mail: nahida.sultana@gmail.com
HCV evolves by means of different mechanisms of genetic variation. On the one hand, its high mutation rates generate the production of a large number of different but closely related viral variants during infection, usually referred to as a quasispecies. This quasispecies is composed of a group of heterogeneous RNA sequences centered around a dominant nucleotide sequence that changes, under the selective pressure of the host immune system throughout the course of infection.

The great quasispecies variability of HCV has also therapeutic implications since the continuous generation and selection of resistant or fitter variants within the quasispecies spectrum might allow viruses to escape control by antiviral drugs. On the other hand, HCV exploits recombination to ensure its survival. This enormous viral diversity together with some host factors has made it difficult to control viral dispersal. Current treatment options involve ribavirin in combination with a direct-acting antiviral drug, depending on the country. Despite all the efforts put into antiviral therapy studies, eradication of the virus or the development of a preventive vaccine has been unsuccessful so far. This review focuses on current available data reported to date on the genetic mechanisms driving the molecular evolution of HCV populations and its relation with the antiviral therapies designed to control HCV infection.

**Genetic Diversity**

Genetic variability is one of the most remarkable features of HCV, contributing to evasion of host immune responses and complicating development of diagnostics, therapeutics, and effective vaccines. HCV genomic sequences can be clustered phylogenetically into related groups (genotypes and subtypes), are distinct between individuals, and are highly variable within each infected individual at any given point in time (i.e., quasispecies diversity) and over time (i.e., quasispecies divergence). Shortly after its discovery in 1989, it became clear that HCV had substantial nucleotide sequence diversity, with only 66 to 80% overall sequence similarity among strains belonging to different genotypes or subtypes. The overall sequence similarities over complete genomic sequences are at least 91% within variants of the same genotype, approximately 79% (range, 77 to 80%) between subtypes, and about 68% (range, 66 to 69%) between different genotypes. Different molecular mechanisms including mutation, genetic drift, recombination and natural selection shape the molecular evolution of HCV. Analogous to the other viruses, HCV circulates as heterogeneous but related genomes, called quasispecies. The processes of neutral and adaptive evolution of HCV operate during the course of chronic infection within an individual, leading to both continued fixation of nucleotide changes over time and the development of variable degrees of sequence diversity within the replicating population at a given time point. This viral population is composed of a dominant sequence called master sequence and a number of different sequences. Master sequence generally represents the consensus sequence of the population. This sequence might not represent an actual genome in the quasispecies, but it is useful to identify the adaptive (Darwinian) changes that affect a representative proportion (>50%) of total quasispecies population.

**Mutation and Quasispecies**

HCV evolution is a highly dynamic process. Like most RNA viruses, HCV exploits all possible mechanisms of genetic variation to ensure its survival. HCV exists in each infected host as a swarm of genetically related but distinct variants,
collectively called a quasispecies. Diversity is generated by mutations introduced by the NS5B RNA-dependent RNA polymerase, which lacks a proofreading function and has an estimated error rate of $10^{-3}$ to $10^{-5}$ per nucleotide per replication cycle. Enhancing this diversity is the high rate of viral replication, with $10^{10}$ to $10^{12}$ virions produced per day. This dynamic, error-prone replication is likely to generate a vast array of mutants every day. Due to this feature and to the high replication rate of HCV, a large number of different but closely related viral variants are continuously produced during infection. Due to neutral drift and sequential selection events, HCV quasispecies sequences have motifs that gradually change over time and during passage among individuals, making sequence analysis suitable for forensic and epidemiologic linkage studies. Many viral factors, such as the error rate of the polymerase, short replication cycle, and compact genome contribute to the generation of the cloud of variants. Additionally, host factors and immune responses exert a selection pressure that contributes to evolution and diversification of the quasispecies. The variation between individual viral genomes in vivo is thought to contribute to persistence, resistance to treatment, tissue tropism, and the failure of experimental vaccines.

**Quasispecies Evolution during Acute HCV Infection**

A diverse viral population, under influence from a variety of selection pressures in a complex host environment, is an ideal situation for viral adaptation in a Darwinian manner. The quasispecies is shaped by positive selection pressure from the host (immune response) and negative selection pressure due to functional constraints imposed by requirements of the viral life cycle; therefore, each host’s HCV quasispecies directly reflects dynamic aspects of both the host and pathogen. During acute HCV infection, the diverse quasispecies may be targeted by cellular and humoral immune responses, which have the potential to reduce the fitness of variants carrying epitopes they recognize and therefore apply positive selection pressure. The rate of mutation in nonenvelope genes decline during the transition to chronicity, consistent with progressive T-cell dysfunction. Some escape mutations require compensatory changes to restore fitness, possibly accounting for additional changes observed during acute HCV infection that The envelope (E1E2) region of the HCV genome is highly variable and has a much higher rate of evolution within hosts than other regions of the genome. Humoral immune responses directed against envelope genes E1 and E2 have the potential to neutralize HCV, and the HCVpp and HCVcc systems provide the means to correlate E1E2 evolution with neutralizing antibody responses. HCV escape from neutralizing antibodies drives the evolution of envelope sequences during acute infection. This is contrasted with relative stasis of HCV envelope sequences in persons with severely impaired humoral immunity. Both stasis and driven evolution are illustrated by HVR-1 evolution in a subject with acute HCV infection progressing to chronicity in whom neutralizing antibody responses were not detected in the first 2 years of high-level viremia, during which there was no evolution of HVR-1, whereas there was rapid evolution following the detection of neutralizing antibodies. Delayed onset of neutralizing antibody responses is typical in those developing persistent HCV infection and appears to explain the acceleration of envelope evolution during the transition from acute to chronic infection.

**Clinical Significance of HCV Quasispecies**

HCV quasispecies are of research interest, but their measurement has not yet been directly applicable to HCV management or treatment. An
example of the significance of quasispecies is that, in acute HCV infection, isolates develop little genetic diversity in a particular region when they produce self-limited hepatitis, whereas they develop greater genetic diversity in the setting of persistent infection.24 Thus, the dynamics of quasispecies evolution during acute infection may reflect the future course of infection. Quasispecies diversity also has been found to be stable or to increase in patients being treated with DAA drug and ribavirin who do not respond to therapy, whereas it decreases in responders.25 HCV variants also can be used to prove linkage of infections that are associated epidemiologically. For example, molecular analysis has been used to link mother/infant pairs, to define HCV in apparently concordant sexual couples as virologically concordant or discordant, to prove nosocomial transmission of HCV between health care provider and patient, and to link needle stick recipients with the sources of infection.26

**Biological Implications of a Quasispecies Distribution for HCV Virus**

Viruses that circulate as a quasispecies present a unique set of challenges to the host. Variants may differ in their biological properties such as virulence, ability to escape the immune system, resistance to antiviral therapies, and tissue tropism. A specific variant within the quasispecies can have a phenotype that differs from the majority of the population. If such variants arise de novo they can change the course of disease.27 The quasispecies nature of viruses such as HCV and HIV may contribute to the challenge of vaccine development because the use of live attenuated viruses in vaccines is risky due to the potential of these viruses to mutate rapidly and become virulent.

The major implications of a quasispecies distribution are discussed below with examples from studies of HCV.

**Transmission and quasispecies divergence**

HCV viruses that circulate as a quasispecies, there are usually multiple variants present at the time of transmission of infection. Transmission of all variants within a quasispecies is not uniform.28-30 Often transmission results in a population bottleneck, with only a small fraction of the variants in the original quasispecies passing to the new host. The dominant variants in the inoculum are often poorly adapted to the new environment. As a result, a minor variant in the quasispecies often becomes dominant in the new host. When multiple variants successfully make the transition into the new host, there is high quasispecies diversity during the early phase of infection.31 Within an infected individual HCV quasispecies diversity can vary greatly. Herring et al., reported sequences differing by 1 to 7.8% at the nucleotide level in HVR1 in the Quasispecies of 12 subjects during the early phase of infection.31

**Compartmentalization**

Analysis of variants isolated from different body compartments show that the members of the quasispecies are not randomly distributed. Variants with different tissue tropisms and compartmentalization of genomes have been observed for a number of RNA viruses including HIV and HCV. Sequence variants that are restricted to a particular body compartment have been found in the serum,22 peripheral blood mononuclear cells (PBMCs),34-36 CNS,37-40 and other extrahepatic sites in specimens from patients with HCV, suggesting that some portions of the quasispecies may replicate in isolation from other portions. Individuals infected with multiple distinct quasispecies may show complete segregation of the two populations. In extreme cases, sequences of one HCV genotype are isolated from a specific compartment (e.g., liver)
and sequences of a different genotype are isolated from a different compartment (e.g., PBMCs). We previously described a patient infected with both genotype 1a and genotype 1b HCV. Genotype 1b variants were found exclusively in the liver, while genotype 1a variants were found in the liver, plasma, and brain tissue. It is also possible to demonstrate that the genetic distances observed between variants isolated from different compartments are significantly greater than the genetic distances observed between variants in the same compartment.

**Viral persistence and progression of disease**

A few RNA viruses, including HCV, can establish a chronic infection. It is thought that changes in antigen epitopes may contribute to persistence by allowing the virus to escape from the adaptive immune system. According to this hypothesis, in the absence of immune pressure, there should be little or no antigenic diversity in the viral population. Indeed, one small study of HCV patients with agammaglobulinemia, found limited quasispecies diversity. More recent studies of HCV patients undergoing liver transplantation for HCV related cirrhosis showed that post transplant viral complexity and diversity was lower in HVR1 than pre-transplant, suggesting that immunosuppression can decrease the introduction of new variants into the quasispecies.

It remains unclear if the increasing complexity and diversity of the quasispecies, and in particular immune escape mutants are the cause or result of chronic infection. Studies of HCV show that progression from acute to chronic infection is associated with an increase in complexity in HVR1 sequences of the quasispecies. In patients with transfusion acquired HCV, viral clearance was associated with stasis of the quasispecies during acute phase, while progression to chronic infection was associated with evolution of the quasispecies. In serum samples taken before seroconversion, there was a significant decrease in the diversity of the quasispecies relative to pre-seroconversion serum samples in patients who cleared the virus spontaneously. No such decrease in quasispecies diversity was observed in those who progressed to develop chronic infection. These changes in the quasispecies were observed only in the HVR1 region, underscoring the importance of HVR1 in HCV-host interactions.

**Implications to new antiviral strategy**

In viruses that circulate as a quasispecies, the polymerase error rate is such that on average one point mutation is made per replication cycle. This mutation rate works to the advantage of the virus. Among mammalian viruses, the error rate contributes to fitness by helping the virus evade the immune system.

However, due to the quasispecies nature of HCV, it is difficult to develop durable small molecule inhibitors of HCV. Variants harboring resistant mutations to protease inhibitors have been observed in treatment naïve subjects. The emergence of minor variants with drug resistant phenotypes can be anticipated and has been observed in early trials of the STAT-C (specifically targeted anti-viral therapy for HCV) drugs. Although resistance mutations have generally been associated with reduced viral fitness relative to the wild-type, these variants can become major species during treatment. As a consequence, resistant mutants are likely to become more prevalent in the population, reducing the success rate of newly developed drugs. Combination therapy of protease inhibitors with interferon and ribavirin should increase the likelihood of achieving a sustained virological response.
Conclusions

The quasispecies nature of HCV in particular, has significant implications for the behaviors of these viruses in vivo. The exact mechanisms by which the variants in the population contribute to compartmentalization, viral persistence and progression, and transmission events remain open questions.

References


