Review

Advanced Molecular Surveillance of Hepatitis C Virus

Livia Maria Gonçalves Rossi [1,*], Alejandro Escobar-Gutierrez [2] and Paula Rahal [1]

[1] Department of Biology, Institute of Bioscience, Language and Exact Science, Sao Paulo State University, Sao Jose do Rio Preto, SP 15054-000, Brazil; E-Mail: rahalp@yahoo.com.br

[2] Instituto de Diagnostico y Referencia Epidemiologicos, Mexico City 01480, Mexico; E-Mail: aescobargutierrez@yahoo.com

* Author to whom correspondence should be addressed; E-Mail: liv.rossi@yahoo.com; Tel.: +55-17-3221-2379; Fax: +55-17-3221-2390.

Academic Editor: Eric O. Freed

Abstract: Hepatitis C virus (HCV) infection is an important public health problem worldwide. HCV exploits complex molecular mechanisms, which result in a high degree of intrahost genetic heterogeneity. This high degree of variability represents a challenge for the accurate establishment of genetic relatedness between cases and complicates the identification of sources of infection. Tracking HCV infections is crucial for the elucidation of routes of transmission in a variety of settings. Therefore, implementation of HCV advanced molecular surveillance (AMS) is essential for disease control. Accounting for virulence is also important for HCV AMS and both viral and host factors contribute to the disease outcome. Therefore, HCV AMS requires the incorporation of host factors as an integral component of the algorithms used to monitor disease occurrence. Importantly, implementation of comprehensive global databases and data mining are also needed for the proper study of the mechanisms responsible for HCV transmission. Here, we review molecular aspects associated with HCV transmission, as well as the most recent technological advances used for virus and host characterization. Additionally, the cornerstone discoveries that have defined the pathway for viral characterization are presented and the importance of implementing advanced HCV molecular surveillance is highlighted.

1. Introduction

Hepatitis C virus (HCV) infection is an important global public health problem. Approximately 180 million people are currently infected with HCV [1], and an alarming number of new infections occur annually [2,3]. The prevalence of hepatitis C varies significantly worldwide, imposing an important burden in highly endemic countries [2]. HCV infection is commonly associated with chronic liver disease, which frequently results in the advanced stages of cirrhosis and hepatocellular carcinoma (HCC) following years of silent infection [4,5].

HCV is a single-stranded, positive-sense, enveloped flavivirus. The viral RNA genome is ~9.6 kb in length and contains a single open reading frame encoding a large polyprotein. The polyprotein is processed by viral and host proteases, resulting in three structural proteins and seven nonstructural proteins (Figure 1A) [6,7].

HCV molecular evolution plays an important role in virus transmission, dictating in many ways the outcome of disease and therapy. Based on the nucleotide variability in the viral genome, seven HCV genotypes and multiple subtypes have been recognized (Figure 1B) [8]. These genotypes exhibit a characteristic distribution [9]: genotypes 1–3 are distributed worldwide, while genotypes 4 and 5 are mainly found in Africa and genotype 6 is endemic in Asia [9–15]. The high mutation rate is characteristic of HCV replication, resulting in a high degree of intrahost genetic diversity [16–18]. The molecular plasticity of HCV allows rapid rearrangement of the intrahost viral population under different selection pressures [19,20]. This remarkable genetic variability is one of the main factors that have prevented the development of a successful vaccine.
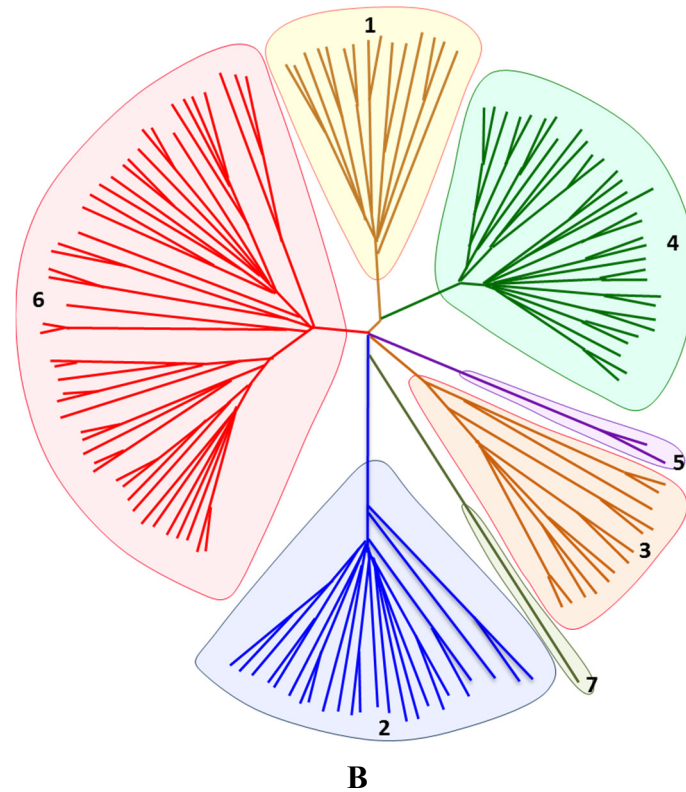

Figure 1. *Cont.*

B
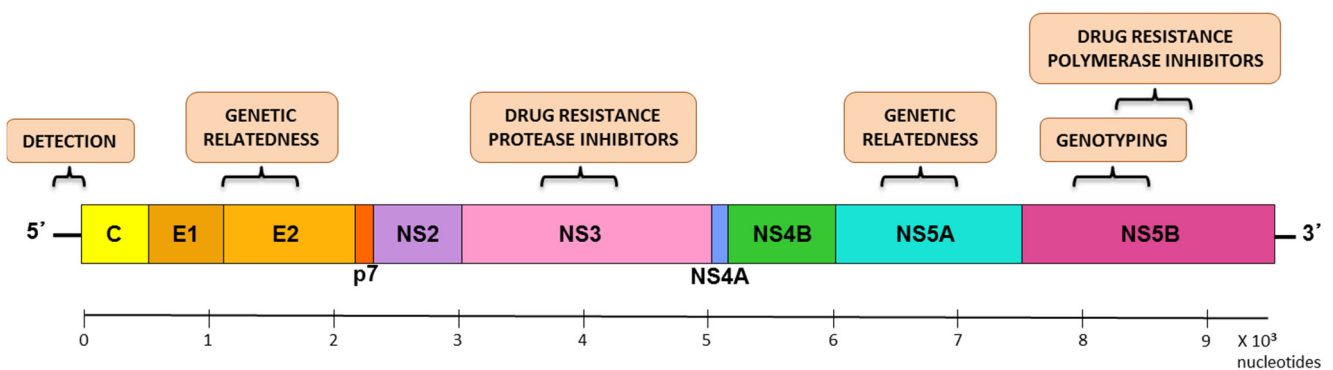
Figure 1. HCV genome organization and genotypes. (**A**) HCV genome organization is shown. The most common regions used for molecular analyses and genetic relatedness studies are highlighted; (**B**) HCV representative strains belonging to all seven genotypes and all different subtypes are depicted. Sequences from the NS5B region were chosen to illustrate the maximum diversity within a subtype. The neighbor joining tree was constructed using maximum composite likelihood nucleotide distances using MEGA5 and 1000 replications for bootstrapping.

Molecular surveillance of hepatitis C is of the utmost importance for identification of HCV transmission networks. Molecular surveillance is a critical component of HCV outbreak investigations because tracking of infections is necessary for the identification of sources of infection and accurate identification of cases associated with transmission networks [21]. Additionally, molecular surveillance can provide important information about the capacity of HCV lineages to cause disease [21].

In addition to imposing important challenges for vaccine development, HCV genetic variability also represents a major issue for the establishment of genetic relatedness between cases associated with a common source of infection. The rapidly evolving nature of the virus results in significant changes in genetic makeup of the virus, which can prevent our ability to link cases by genetic similarity [16]. However, a high degree of genetic variability is also required to track infections and facilitates genetic relatedness studies which otherwise would be hampered [22]. Here, we present a historical review of the keystone discoveries that have propelled our understanding of the mechanisms controlling HCV transmission. Additionally, we discuss the most recent information on molecular aspects related to viral spread and novel approaches designed to assess HCV intrahost genetic variability for the establishment of genetic relatedness between cases. Finally, the implementation of advanced HCV molecular surveillance as an integral component of hepatitis C global epidemiology is highlighted.

2. Historical Timeline of HCV Characterization

Since its discovery 25 years ago, the study of HCV has seen a number of cornerstone developments that have improved our understanding of HCV-related disease pathogenesis (Figure 2). The field of viral hepatitis initially began in the late 1950s with the coining of the term "infectious" or "serum" hepatitis [23]. This event was followed by the discovery of the hepatitis A (HAV) [24] and hepatitis B viruses (HBV) [25] in the late 1960s and early 1970s. However, our understanding of HCV infection dates to the late 1980s when researchers from the Centers for Disease Control and Chiron Corporation, led by Daniel Bradley and Michael Houghton, respectively, identified the virus in samples from experimentally infected chimpanzees [26]. Prior to its identification, the advent of serologic testing for HAV and HBV in the mid-1970s led to the realization that most infectious hepatitis cases were non-A, non-B (NANBH). These findings prompted the use of the chimpanzee as an animal model for the passage of the unidentified agent responsible for NANBH [27]. The subsequent development of serologic tests for the detection of HCV infection in the early 1990s allowed screening of the blood supply and successfully prevented transmission via transfusion [28]. HCV molecular divergence was recognized shortly after the discovery of the virus when Japanese strains were shown to be genetically distant from American strains [29]. Subsequently, the complete sequencing of the HCV genome allowed the elucidation of its organization [30–32]. Molecular characterization of the viral enzymes during the early 1990s served as the foundation for the development of successful anti-HCV therapies [18,33–37] and resulted in the constant reshaping of anti-HCV therapy based on interferon (IFN). The initial approval of alpha IFN (IFNα) for the treatment of HCV in 1991 was followed by the licensing of consensus IFN in 1997 and ribavirin (RBV) in 1998. In 2001, the introduction of pegylated IFN, a compound with significantly prolonged half-life in blood, increased the likelihood of a sustained virological response (SVR). Full recognition of the degree of genetic heterogeneity among HCV isolates was achieved during the mid-1990s [38,39], resulting in the proposed classification of *Hepaciviruses* as an independent genus within the *Flaviviridae* family representing a distant relative of the *Flavivirus* and *Pestivirus* genera [40]. Importantly, the fact that HCV diversity played an important role in interferon (IFN)-based therapy highlighted the relevance of assessing the genetic variability of the viral population *in vivo* [41,42].

An important advance for drug development was the generation of human hepatoma cell lines capable of replicating HCV [43–46]. Over the years, other important discoveries have been reported by several groups, especially relating to the characterization of viral proteins and their role in HCV replication [34,47–49]. Additionally, the characterization of several cellular receptors [50–53], and entry factors [54,55] has been reported.

The initial release of the first next generation sequencing (NGS) commercial platform in 2005 completely revolutionized the field of genetics. However, initial studies reporting the use of NGS approaches to assess HCV intrahost genetic variation emerged in late 2011 [56,57]. Since this time, a wealth of information has been generated using NGS to address a multitude of issues related to HCV transmission [58–71]. The challenge now resides in the study of whole genome (WG) genetic variation, which still represents a daunting task for the study of HCV molecular evolution since reconstruction of original haplotypes from NGS data is difficult [72–74].
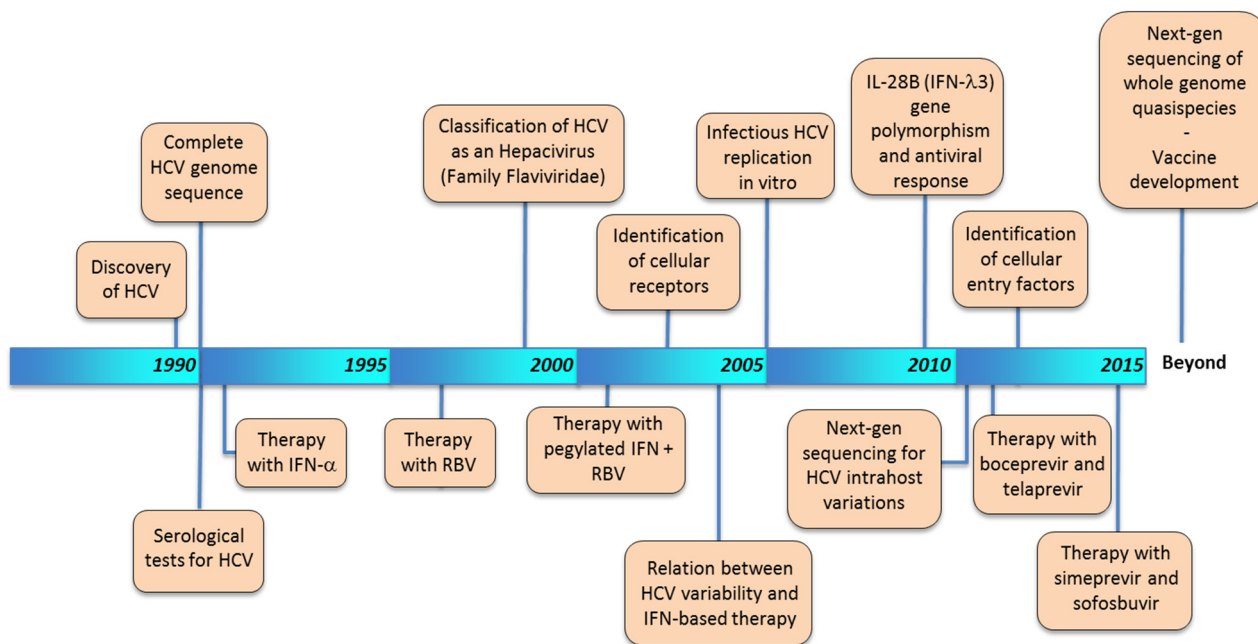
Figure 2. Chronological cornerstone developments in the field of HCV. The most relevant discoveries related to the study of HCV are shown in chronological order.

In 2009, Ge *et al*. reported a strong association between polymorphisms in the promoter region of the interleukin-28B gene (also known as IFN-λ3) and the antiviral response based on IFN [75]. Subsequently, work conducted by other groups further confirmed the relationship between the IL-28B genotype and sustained viral response (SVR) [76,77]. The relevance of these reports cannot be understated, because they resulted in extensive work on human genetics and became an important component in the study of the pathogenesis of HCV-related disease.

The arrival of the second generation of direct-acting antivirals (DAA) in 2011 resulted in improved SVR and a revolution in the field of anti-HCV therapy. Despite the development of drug resistance, the initial NS3-4A protease inhibitors (PI) telaprevir and boceprevir undoubtedly possessed advantages over dual IFN/RBV treatment [78]. The success of both drugs was subsequently eclipsed by the licensing of simeprevir and sofosbuvir [79–82]. The field of HCV therapy is rapidly evolving, and as a result, a large number of new antiviral drugs are currently being evaluated in advanced clinical trials [83].

The prevention of HCV infection has been hampered by the slow development of promising vaccine candidates. Phase I of two different vaccine trials have been completed. Testing of a prime-boost regimen with the candidate developed by GlaxoSmithKline after preliminary studies in the chimpanzee model has also been performed [84]. Importantly, conserved epitopes in HCV genotype 1 and 3 were successfully identified, suggesting the potential for cross-genotypic protection. Priming with this vaccine followed by boosting with a modified vaccinia virus is currently being evaluated in a phase I/II trial [85], with the goal of preventing persistence in HCV-naïve IDUs at high risk for infection.

3. Molecular Aspects of HCV Transmission

HCV transmission is a dynamic process that primarily occurs via parenteral routes and especially by unsafe injections, which have significantly facilitated virus spread [86]. The time of initial spread of

HCV into Western countries and the population dynamics of the epidemic can only be indirectly inferred [13]. Current evidence supports a recent spread; however, the lack of samples prior to the Second World War has hindered our ability to reconstruct the HCV epidemic. Because vertical and sexual transmission contribute little to HCV spread, the restriction of HCV transmission through primarily parenteral routes implicates unsafe injections (*i.e.*, inadequate medical treatment, large-scale vaccination programs, blood transfusion, and injecting drug use) as the main vehicle for HCV transmission [13]. Limited information is available about the transmission of HCV prior to the invention of injections. However, cultural traditions, such as tattooing and scarification, were likely to have played an important role in the spread of HCV.

HCV intrahost populations frequently exist as an ensemble of genetically distinct but closely related variants [18,87]. Analysis of HCV intrahost genetic variation is the basis of genetic relatedness and epidemiological studies as well as the identification of drug resistant mutations [88,89]. While in some instances genetic relatedness might be demonstrated with consensus sequencing of some viral subgenomic regions, HCV outbreak investigation commonly requires a much deeper analysis of the infecting viral population [16,90–93]. The rapidness with which HCV intrahost populations diverge significantly affects genetic relatedness studies because molecular epidemiological links can be lost between related cases in a relatively short period of time (Figure 3) [16,94–96]. Therefore, the use of more and longer subgenomic regions, such as the NS3, NS5A, and NS5B might help alleviate this issue.
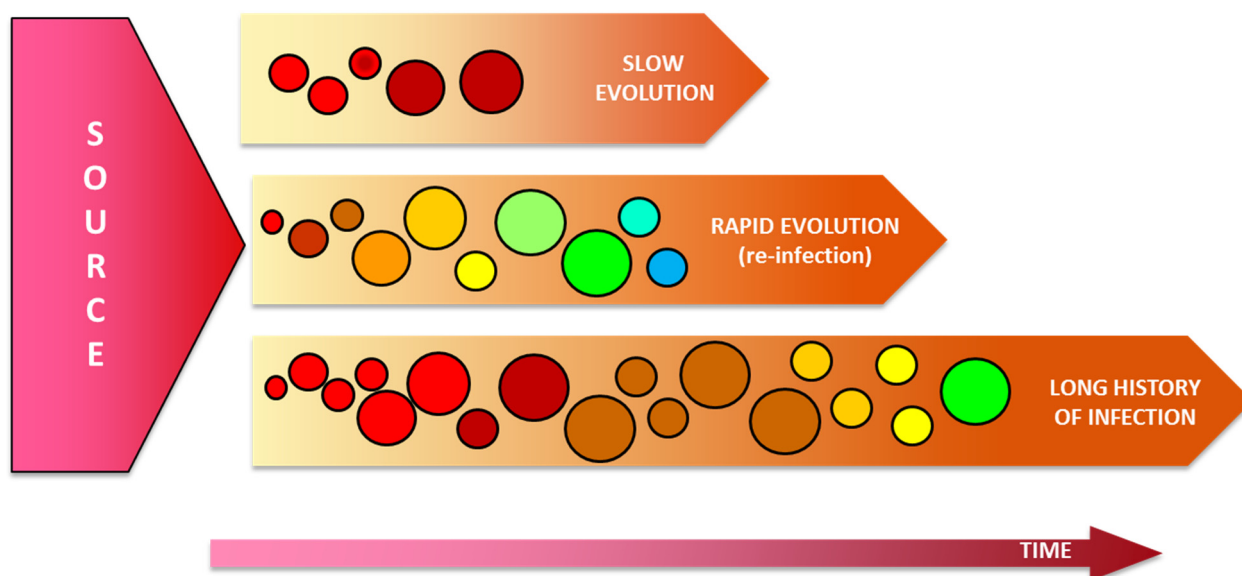
Figure 3. Rapid and slow HCV intrahost molecular evolution. The speed with which HCV evolves in different settings is shown. Viral populations are represented by circles. Viral populations are color coded to represent genetic diversity. The size of circles represents fluctuations in the size of viral populations over time. Generation of viral diversity during the course of infection is illustrated by the change in color of the viral population throughout the time.

The molecular mechanisms that shaped the evolution of HCV are quite diverse. Upon transmission, HCV undergoes a series of events that affect the architecture of the viral population (Figure 4). The "steps" taken by HCV throughout the infection process are highly complex and depend on fluctuations in the complexity and size of the intrahost population. HCV transmission events are characterized by genetic bottlenecks that inflict a strong selective pressure during acute infection [56,97–99]. Indeed, the HCV genetic bottlenecks are so intense following transmission that very few variants are able to establish infection in the new host [100,101]. Subsequently, genetic drift commonly takes place as the main force driving the molecular evolution of HCV during the early stages of the acute phase as a result of the strong founder effect [102–104]. During the later stages of HCV infection the extent of genetic drift is significantly reduced [102–105], allowing a large viral population size to come to prominence in chronic cases. To a lesser degree, genetic recombination also affects overall HCV genetic heterogeneity. HCV genetic recombination is rare [106], mainly due to the exclusion of superinfections [107–109]. However, despite limited HCV recombination, naturally occurring inter-genotype, intra-genotype, and intra-strain HCV recombinants have been reported [110–113]. Thus, HCV recombination might have important implications for clinical and epidemiological studies [114]. Staging of HCV infections has also been proposed [115]. This hypothesis suggests that during chronic infection, HCV attempts to reach a settlement stage characterized by a viral population undergoing strong negative selection. This "journey" involves complex processes including temporal variations attained by incremental changes between communities, random mutations and fluctuations in the frequency of coexisting viral subpopulations; in conjunction these factors constantly reshape the architecture of the viral population [89,95,115,116]. This staging is likely to affect HCV transmission because different viral variants that most likely possess different degrees of transmissibility are available at different time points during the course of the infection [21]. Despite the complexity of the means used by HCV to evolve in an adverse microenvironment, genetic relatedness can be successfully established in many instances if transmission has occurred relatively recently. In addition to time of infection, establishment of relatedness depends on the degree of divergence and other external factors such as a history of antiviral therapy, which can significantly alter the natural course of HCV evolution.

HCV transmission networks are difficult to be recognized for numerous reasons [117,118]. The long incubation periods and the characteristic asymptomatic nature of acute HCV infections make identification of cases a rather challenging task [117]. HCV transmissions are difficult to link to their corresponding source of infection because intrahost viral populations are often genetically related but seldom identical [95,119]. Importantly, branching in phylogenetic analyses, derived from genetic relatedness studies, do not always correspond to transmission events, particularly in those cases where not all individuals belonging to the transmission network are sampled (Figure 5). Thus, local epidemic sequences can cluster together in the absence of direct transmission [119,120].
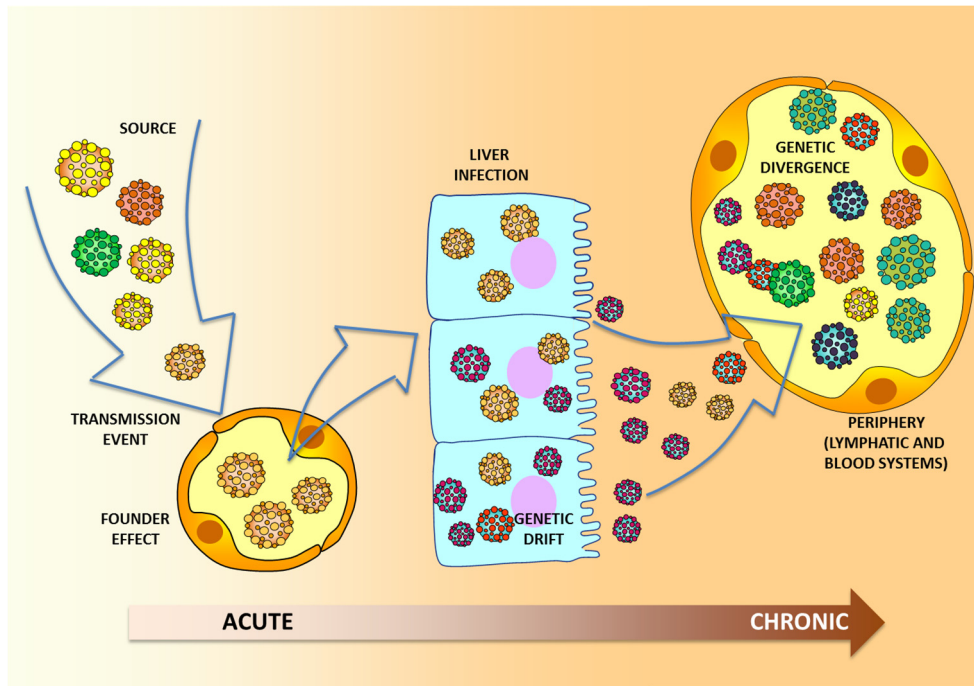
Figure 4. HCV intrahost molecular evolution. The main forces driving the molecular evolution of HCV are depicted. The generation of viral diversity during the course of natural infection is illustrated. Upon infection, the virus migrates via blood vessels to the liver where replication takes place. Viral diversity is increased throughout the infection process and reflected in the periphery over time.
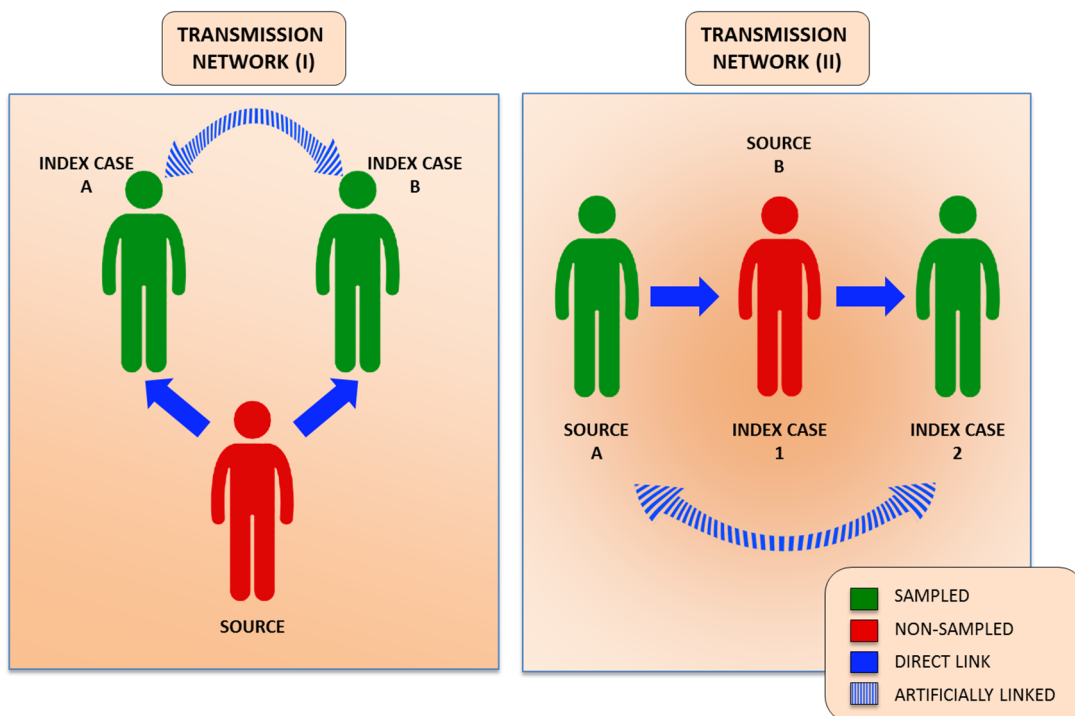

Figure 5. HCV transmission networks. Two transmission networks are depicted. The limitations of HCV molecular studies are depicted when not all individuals belonging to the transmission network are included. Artificially linked cases can be observed when not all members of the transmission network are sampled.

HCV evolution is also affected by preexisting liver conditions, infection with other viruses and also pregnancy as a direct consequence of the alterations of the immune response in the mother [17]. During pregnancy, increased viral loads and a reduction in CD8+ T cell cytotoxicity are commonly observed, favoring the emergence of more fit strains [121] as a result of the impairment of the antiviral response in the mother. Therefore, children infected perinatally might potentially bear infections with the more fit viruses, thereby facilitating persistence [122].

4. Human Genomics in HCV Advanced Molecular Surveillance

The era of human genetics is expected to play an important role in the clinical management of infectious diseases [123]. HCV infection is not the exception, and several studies have shown that host genetic makeup affects the outcome of infection and therapy [21].

HCV is spontaneously cleared by the host in up to 15% of acutely infected individuals [124–126]. It has been reported that an early functional inactivation (peripheral tolerance) of the HCV-specific CD4+ T cell response may play a major role in establishing viral persistence [127–130], and such dysfunctional T cells can be functionally reactivated [131,132]. Several studies have suggested that differences in host immune responses determine viral clearance. Indeed, the intensity of the immune response can be dictated, at least in part, by host genetic factors [133]. However, the role of host factors in HCV clearance is not well understood. Different studies have shown correlations between HCV-related disease and the genetic background of the host [75,134]. For example, certain HLA alleles have been associated with specific outcomes of HCV infection [135,136]. As a result of the completion of the Human Genome Project and the development of novel methodologies capable of rapidly assessing the host genetic background, a number of disease-associated genes have recently been reported. Consequently, mutations in these genes can be screened and used to assist in the clinical decision making process [137]. This information can also be incorporated into heightened surveillance for a number of infectious diseases, facilitating the incorporation of pharmacogenetics into global surveillance [138].

Recently, several single nucleotide polymorphisms (SNP) in immune-related genes have been associated with HCV disease outcomes. Large-scale genetic analyses have suggested the participation of different SNPs in the spontaneous clearance of HCV, including TNFSF18 (tumor necrosis factor superfamily, member 18), TANK (TRAF family member associated NF-κB activator), HAVCR1 (Hepatitis A virus cellular receptor 1) and IL-18BP (IL-18 binding protein) [139]. Additionally, SNPs occurring in the gene encoding tumor necrosis factor alpha (TNF-α) have also been proposed to participate in HCV clearance, particularly in subjects of African origin [140,141]. TNF-α is an important cytokine produced primarily by macrophages that participates in the induction of apoptosis, gene regulation and cellular proliferation [142]. Interestingly, high levels of TNF- α have been detected in the serum and liver of chronically infected patients [143]. SNPs in transforming growth factor (TGF)-β1 and IFN-γ have also been suggested to participate in HCV clearance [144,145]. TGF-β1 is a suppressor of natural killer (NK) cells that inhibits the production of IFN-γ and IL-12, thereby blocking the proliferation and cytotoxicity of NK cells [146]. Moreover, dysregulation of TGF- β1 has been shown to be involved in the progression of liver cirrhosis and HCC [147]. IFN-γ is a multifunctional cytokine produced by effector T and NK cells that participates in the development of T helper 1 (Th1) cells and

is critical for host defense against a variety of intracellular pathogens, including HCV. IFN-γ inhibits HCV replication *in vitro*, and intrahepatic levels are associated with viral clearance in animal models [148,149]. SNPs in immune-related genes encoding for interleukin-10 (IL-10) have been reported to play role in HCV-related infection [141]; however, the role of IL-10 in HCV infection is debatable. IL-10 inhibits IFN-γ production, resulting in an imbalance in the T helper response that in turn leads to viral persistence [150], while low levels of IL-10 seem to be associated with resolution of HCV infection [151]. Conversely, IL-10 has been proposed to be antifibrogenetic in chronic liver injury [152].

SNPs in the osteopontin (OPN) gene (−1748 and −443) have been associated with chronic HCV infection [153]. Response rates were higher in patients with the G/G or G/A alleles at nt −1748 in comparison to patients bearing the A/A alleles. Likewise, the response rate was higher in patients with the T/T alleles at nt −443 than in those with C/C or C/T alleles [153]. OPN is a highly phosphorylated sialoprotein and an important component of the extracellular matrix [154], which is secreted by lymphocytes, leucocytes, and macrophages. OPN interacts with cellular adhesion molecules and plays a role in different aspects of the cellular immune response, but the exact role of OPN in HVC infection is still unclear.

Recent studies have demonstrated an association between a SNP in the exon 7 splice acceptor site of the oligoadenylate synthetase 1 (OAS1) gene and SVR in HCV patients. Patients exhibiting an AA genotype showed poor SVR rates and progressed to more severe disease [155]. OAS1 is an important protein with antiviral activity [156] that is activated by double-stranded RNA. Upon activation, OAS1 inhibits viral RNA and protein synthesis. Thus, nucleotide changes affecting its function are expected to further impair virus infection control.

Importantly, a SNP in the mannan-binding lectin (MBL) gene (also known as the mannose binding protein) has been related to HCV infection. HCV patients tend to exhibit YA/YO in comparison to controls and reduced levels of MLB in plasma [157]. Additionally, the frequency of these genotypes is reduced in patients with advanced fibrosis compared to patients with moderate fibrosis. MLB is a pattern recognition receptor (PRR) that has an important function in the innate immune response [158]. Nevertheless, its role in HCV control is not well known.

Another host factor associated with HCV infection is the occurrence of SNPs in the promoter region of the IL-28B gene (also known as interferon-λ 3). The type III IFN family includes IFN-λ 1, 2, 3 and 4 (IL-29, -28A, -28B, IFN-λ 4, respectively) [159,160]. Based on their molecular structures type III IFNs belong to the interleukin-10 (IL-10) superfamily, but functionally they are closely related to type I IFNs (IFNα, IFNβ), which play a major role in antiviral immunity [161]. IFN-λs are produced by dendritic cells, neuronal cells, alveolar epithelial cells, and hepatocytes [162], in response to viral infections through its activation via Toll-like receptors (TLRs) [163]. IFN-λs inhibit viral replication and modulates immune-related functions, such as the maturation and differentiation of immune cells [164–166]. The SNPs in IL-28B have been recognized as strong predictors for both spontaneous and antiviral-induced clearance of HCV. Several studies have shown that patients infected with HCV genotype 1 bearing the C/C, A/A and T/T alleles in rs12979860, rs12980275, and rs8099917, respectively, are more likely to achieve SVR [75–77]. Recently, the discovery of a new gene within the type III IFN family (IFN-λ4) situated upstream of IFNL3, and its association with a poor response to IFN-α therapy and impaired HCV clearance suggests a more complex role for the interferon response in HCV

infection [160]. However, the arrival of second and third generation treatment regimens based on combinations of different DAA with and without IFN has down-played the relevance of these SNPs in the prediction of SVR.

5. Assessment of the HCV Intrahost Genetic Variability

DNA sequencing is essential for the study of transmission, phylogenetics, and outbreaks [167]. Upon transmission, a small subset of the HCV population from the source is transferred to a new host (recipient) and then subjected to immune pressures based on the new host's genetic background. Therefore, the viral populations in the two individuals are related but genetically distinct [88,89,95,96]. Additionally, the characterization of viral variants is also useful for the study of virus tropism in host compartments [89]. The selective forces driving the molecular evolution of HCV are complex in nature and difficult to assess. Most molecular approaches, such as population sequencing, required for molecular characterization of HCV intrahost populations, are arduous, lengthy and costly [168]. However, new technologies based on NGS platforms allow a finer resolution and are more affordable [59,67,168,169], facilitating the molecular characterization of HCV intrahost viral populations [64,168,170].

The analysis of the composition of HCV intrahost populations requires nucleotide sequencing of as many variants as possible in the patient. The assessment of the complexity of the HCV intrahost population is challenging due to the difficulty in detecting low-frequency or minor variants. Establishment of genetic relatedness is affected by the variability of the genomic region analyzed, the time since transmission and the length of the subgenomic region. Moreover, sequence errors introduced by the reverse transcriptase and DNA polymerase during amplification of the template can occur. HCV population complexity can also be skewed by template re-sampling during PCR in low titer clinical samples, resulting in apparently false reduced complexity. Additionally, major variants may mask minor variants due to their respective frequencies during re-sampling. Different approaches have been developed to characterize the HCV intrahost population with different inherent limitations. Alternatives approaches developed for HCV intrahost characterization are based on the identification of nucleotide sequences or assessment of the diversity and complexity without the need to sequence the viral genome [171,172]. Complexity, cost and reliability are some of the factors that play a role in the choice of methods. Historically, some methods were not powerful enough or presented too many shortcomings to overcome. Consequently, these methods have been disused and replaced by more powerful methods capable of analyzing the composition of the HCV intrahost population to an unprecedented level.

Non-sequencing-based methods are disadvantageous because the structure of the viral population is not obtained and the performance relies on the DNA properties of the variants being analyzed. In contrast, sequence-based strategies provide a high-resolution snapshot of the circulating viral variants. Several challenges are faced during the implementation of these techniques, including the stochastic nature of PCR-based methodologies that can significantly distort the composition of the HCV intrahost population.

5.1. Genomic Regions

Tracking of HCV infection depends on sequence information originated from different subgenomic regions. The 5'-UTR region has been widely used for detection owing to its degree of conservation across genotypes, while the NS5B region is the most common target for HCV genotyping [173,174]. However, these two regions are well conserved and, therefore, do not contain sufficient sequence information to establish genetic relatedness between clinical isolates. Instead, genetic relatedness studies primarily rely on information obtained from HVR1 [16,89,93,95]. Rapid divergence in this region represents a challenge for molecular epidemiological studies and can result in the loss of genetic links between related isolates [16]. Sequencing of multiple and longer subgenomic regions has been proposed as an alternative to overcome the limitations imposed by the rapid molecular evolution of HCV [16]. NS5A has been used to establish relatedness among HCV cases [115]; the use of this gene can restore links between isolates owing to its lower nucleotide substitution rate. Despite the usefulness of different subgenomic regions for the characterization of clinical isolates, whole genome sequencing should be the ultimate goal for HCV molecular epidemiology.

5.2. Single-Stranded Conformation Polymorphism (SSCP)

Prior to the implementation of more sophisticated molecular methods for the characterization of individual variants, SSCP was a relatively popular alternative [175–177]. This approach is based on the denaturation of single-stranded amplicons that are resolved in non-denaturing conditions to differentiate molecules based on their secondary structure [178]. The popularity of this method was originally based on the assumption that different variants may adopt different secondary structures; however, several studies showed the implicit limitations of SSCP for the identification of closely genetically related variants [179]. Sensitivity is another major limitation of SSCP, because only variants present in >3% of the viral population can successfully be identified [180], and therefore, minor variants are unlikely to be detected. Additionally, the length of the regions studied imposes important limitations, because long fragments may not resolve as clearly as shorter fragments [181]. Other factors affecting the performance of SSCP include DNA concentration and GC content. Despite being an inexpensive and relatively straightforward approach, SSCP lacks the required resolution to accurately assess the composition of the HCV intrahost population. Thus, it is not surprising that its use has diminished with the arrival of more powerful technologies better suited for such demanding task.

5.3. Heteroduplex Gel Shift Assays

In many ways, heteroduplex gel shift assays resemble SSCP [182,183], and therefore, the benefits and disadvantages between the two methods are rather similar. Overall, the entire intrahost population can be sampled; however, no specific information about specific substitutions in the population can be obtained. This methodology is based on the amplification and hybridization of the viral RNA to a labeled probe. Next, the heteroduplex products are resolved by electrophoresis. Thus, different variants are expected to exhibit different migration patterns due to mismatches with the probe, resulting in differential band profiles depending on the composition of the intrahost population. A slight variation in the conventional heteroduplex assay relies on the cloning of different viral species into plasmids instead.

Genetic differences can be calculated using a heuristic approach known as the heteroduplex mobility ratio. However, the accuracy provided by this method is limited and usually requires variants to differ by ~1.5% [182].

5.4. Population Sequencing

Bacterial Cloning

Sequencing individual variants is a cumbersome and costly task that provides important information about the composition of the intrahost population. The basis of this approach includes amplification of the region of interest, followed by cloning of the amplicons into a plasmid vector and transformation of bacteria. Then, the transformed bacteria are plated. The bacteria are assumed to carry only one plasmid, and the presence of only one viral variant sequence is confirmed by colony PCR and sequencing. Thus, DNA from an individual bacterial colony represents the sequence of a single variant. Unfortunately, sampling is a major issue for this particular approach because even in large studies that analyze hundreds of clones, only a small fraction of the viral population and genome is sampled [173]. Moreover, the exact number of clones necessary to achieve a given level of confidence depends on the frequency of each incorrect nucleotide at any given nucleotide position in the sequence [184]. It has been suggested that between 20 and 100 clones per sample are required for frequency and entropy analyses. Sequencing 99 clones has been proposed to identify 95% of all variants present at a frequency of at least 3% in the population [185]. However, others have suggested that sequencing 20 clones is sufficient to cover 95% of the major variants (frequency >10% in the population). Regardless of the estimate, under these circumstances, most minor variants go undetected, significantly hampering our capacity to untangle the mechanisms involved in HCV evolution. Recent advances in single genome sequencing have improved our understanding of HCV transmission [101]. In summary, regardless of the number of clones analyzed, the depth with which population sequencing can assess the complexity of the HCV intrahost population is extremely limited.

5.5. End-Point Limiting-Dilution PCR (EPLD-PCR)

EPLD-PCR is a viable alternative to single-molecule PCR amplification and has largely been used to identify individual viral variants of different viruses including human immunodeficiency virus (HIV), hepatitis B virus (HBV) and HCV [186–188]. EPLD-PCR relies on the isolation of individual coexisting sequence variants in clinical samples using a limiting-dilution approach followed by sequencing. Overall, EPLD-PCR is less prone to distortion compared to bacterial cloning-based methods because the amplification is based on single DNA molecules. Moreover, real-time-based EPLD-PCR has been shown to be superior to conventional population sequencing methods for the detection of HCV viral variants. However, implementation of EPLD-PCR is cumbersome and time consuming, resulting in a rather expensive methodology.

5.6. Mass Spectrometry (MS)

MS was developed as an alternative to the numerous DNA sequencing technologies [171]. Originally, the accuracy, sensitivity and high-throughput methodology of MS set the basis for the

implementation of molecular methods aiming to characterize single nucleotide changes in DNA molecules [189]. MS analysis relies on the measurement of the molecular masses of different analytes, which should be ionized and separated before detection. Unfortunately, ionization may lead to partial degradation of the analytes, thereby affecting the resolution of this method. To minimize the effect of ionization on the analytes, methods such as matrix-assisted laser desorption/ionization and electrospray ionization have been developed [190,191]. MS has been used for resequencing, microbial typing, and single nucleotide polymorphism discovery [172,192,193].

Analysis of the HCV intrahost population by MS has been reported [172,192,194]. MS represents a high-performance methodology that analyzes base-specific cleaved RNA originating from PCR fragments [171,172,195]. This is a cost-effective and reproducible approach capable of high throughput. The resolution of this method easily accomplishes the identification of nucleotide polymorphisms comprising >10% of the amplicon population. The data originating from this type of analysis represents a composite pattern of the variants present in the sample [36]. Importantly, the identification of novel single-nucleotide variants depends on the proper representation of such variants in the reference database required for base calling [196]; therefore, a constant update of the reference sequences is required. Despite their implicit advantages, MS-based approaches have not been largely used for assessment of HCV intrahost populations based on the fact that *de novo* sequencing is challenging and also because of the arrival of more powerful and inexpensive methods [173].

5.7. Next Generation Sequencing (NGS)

The advent of (NGS) has revolutionized the diagnostics of infectious diseases [197]. The development of massively parallel ultra-deep pyrosequencing allows for a high-resolution snapshot of the intrahost viral population. Additionally, newer approaches have overcome the issue associated with short reads, resulting in longer sequences that facilitate the reconstruction of the viral population [93,173,198–200].

Ultra-deep sequencing is based on a limiting dilution approach and allows rapid sequencing of a large number of variants by eliminating the need to separate molecules and clone into bacterial vectors. Ultra-deep sequencing has been used to identify minor variants in different settings [59,201].

Different platforms are available, including 454 Life Science (Roche), Illumina, Ion Torrent and Pacific Biosciences. These platforms were initially developed as an alternative to the prohibitively expensive Sanger method. Comprehensive characterization of viral populations is easily accomplished with NGS. The advantages of NGS technologies in virology are numerous. Metagenomics is a growing field in virology that has allowed the characterization of viral populations from different types of samples [202–204], especially the detection and characterization of viruses associated with disease outcomes [205]. However, overcoming different challenges is required for the implementation of NGS approaches in the study of viral diseases.

NGS amplicon sequencing allows sufficient coverage in order to detect variants occurring with a frequency of <0.1%. The low abundance of nucleic acids is a major problem for the molecular characterization of viruses from clinical samples. The low ratio of viral RNA/host RNA commonly observed in clinical samples is one the main obstacles to using NGS methods. Pre-amplification and probe capture are some of the alternatives available to overcome the relatively low frequency of viral RNA commonly present in clinical specimens. Additionally, viral enrichment and concentration have

been used as alternative methods to improve sequencing quality and depth. Nevertheless, whole genome (WG) viral sequencing remains challenging. The analytical sensitivity of WG using NGS approaches is not as easy to evaluate and largely depends on the depth of sequencing [206,207]. However, increasing the depth of sequencing for an optimized sample preparation can decrease the level of detection. Moreover, the analytical sensitivity depends on the length of the genome. Longer lengths translate into a higher number of potentially available reads, as seen for some studies of viruses. This should also be the case for bacterial and fungal genomes, which could be viewed as an advantage for the detection of such microbes because their concentrations in blood can be very low even in samples from infected patients. Several studies have evaluated the diagnostic sensitivity of this technique. Analysis of the human virome in febrile and afebrile children allowed the identification of a wide range of viruses that correlated with the outcome of disease, as well as virus subtyping [208]. Another advantage of the technique is its capacity to identify co-infections, which is important for the development of adaptive therapeutics [197].

Thus, NGS will likely become a routine test for diagnostics of infectious diseases [209]. For this initiative to succeed, improvements in sample preparation, availability of sequencers in central laboratories and validated pipelines for read sorting and taxonomic assignation must be attained. The study of viral genetics in the era of next generation sequencing promises to help to unveil the most intimate details of the virus-host interface [168].

6. Bioinformatics, Phylogenetics and Data Mining

6.1. Bioinformatics

A variety of computational challenges are commonly encountered in the field of virology owing to the high diversity of viruses, their compact genome organization and their rapid rate of evolution [210]. For HCV, the increasing number of viral sequences available in public and private databases has promoted the development of novel ways to analyze considerably large data sets [211,212]. Relatively recent technological innovations have ignited an explosion in virus genome sequencing that is likely to help us understand the intimate details of HCV biology and its impact on public health. Nonetheless, any benefits derived from the generation of massive sequence data are hindered by the implementation of resources capable of identifying the sequences, as well as assembling, annotating, curating, maintaining and storing extremely large databases [117,211,212]. Importantly, virology has recently drawn the interest of the bioinformatics community, which in turn has led to the development of a variety of tools. This rapid expansion of the HCV sequence universe has forced a recalibration of the data model to better provide extant sequence representation and enhanced reference sequence products to serve the needs of the various viral communities [211–213]. However, despite the growth in viral bioinformatics, a number of questions remain unanswered, including but not restricted to identification, genome annotation, phylogeny, evolution, and genetic diversity.

Viral evolution has many implications for clinical virology. The emergence of HCV resistance mutations is one of the most important challenges for successful antiviral therapy [214]. The molecular mechanisms selecting resistance mutations are complex; therefore, new bioinformatics approaches to characterize HCV evolution both at an intra- and interhost level are required. The integration of

bioinformatics methods might lead to predictions of viral evolution in HCV chronic patients based on sequences derived from their intrahost viral populations. Hence, the ultimate goal in HCV genetics is the prediction of the course of HCV evolution that in turn could lead to the customized management of the patient and hepatitis C treatment.

6.2. Phylogenetics

Viral phylogenetics have also benefited from the exponential growth of bioinformatics. Phylogenetics is an important area in virology, particularly in HCV molecular epidemiology [21]. However, several characteristics of HCV are challenging for phylogenetics. Significant differences in evolution rates (high over the short term and much lower over the long term), gene transfer, evolutionary virus-host relationship, and the lack of physical "fossil records" of viruses (ancient viruses) remain difficult obstacles to overcome. Additionally, phylogenetic trees derived from HCV sequence analyses cannot faithfully represent complex evolutionary relationships relevant to HCV such as horizontal gene transfer, recombination or evolutionary virus-host relationships. Novel phylogenetic approaches have been developed in recent years to better represent such relationships [215,216]. However, the field of phylogenetics warrants further research to address several aspects related to the reconstruction of HCV phylogeny.

Establishing the time of infection is one of the most important and challenging tasks in any epidemiological investigation. However, infection metrics among cases with multiple risk factors and possible exposures are difficult to pinpoint. Therefore, determination of time of infection using molecular data represents an attractive line of research. Estimating the time of infection based purely on genetic data has been reported [92,217,218]. However, molecular clock estimates vary significantly, and their reliability depends on a number of factors including sampling, temporal and anatomical distribution of sampling, genome region sequenced, super- or re-infection, and the evolutionary models and algorithms used [119]. As a consequence, the implementation of such approaches for genetic relatedness studies should be undertaken cautiously [119].

6.3. Databases and Data Mining

A database is a collection of data that is organized so that its contents can easily be accessed, managed, and updated. Many databases for infectious diseases compile data obtained from clinical and/or public health laboratories, aiding the monitoring of infectious disease trends and surveillance [219,220]. However, even with our growing ability to acquire sequence data, its potential to impact disease surveillance is only fully accessed when data are translated into public health actions.

Initiatives such as the Viral Bioinformatics Resource Center (VBRC) aim to close the gaps between massive viral sequencing and bioinformatics are a welcome addition to the set of on-line bioinformatics tools available to the scientific community. These tools provide comprehensive web-based genomics resources that are useful for basic and applied virology research. The VBRC consists of relational databases and web applications for data storage, annotation, analysis, and information exchange. Curation of the data results in the ability to search for gene functions relating to biological genotypes and phenotypes with an emphasis on pathogenesis and provides a variety of analytical and visualization tools [213].

The HCV Database Project was initially funded by the Division of Microbiology and Infectious Diseases of the National Institute of Allergies and Infectious Diseases (NIAID). The HCV database is a resource for the scientific community working on HCV genetics, evolution, variability, and vaccine and drug design [211,212]. The HCV sequences deposited in GenBank composed the backbone of the database, including information such as country, sampling year, isolate names, genotype and subtype, and host species in addition to relevant annotation information associated with the corresponding publications. Annotation fields in the database include genotype, subtype, start and stop coordinates relative to the reference strain HCV-H, sampling country, sampling city, sampling date and sampling tissue. Patient information is also documented in the database, including health status, age, gender, ALT level, treatment and result, co-infection with HIV and hepatitis B, infection date, infection country, infection city, infection route, infection outcome, HLA type, and epidemiological relationship with other patients. The HCV database contains HCV-associated, hand-annotated genetic data and provides access to the central database via web-accessible search interfaces in addition to a number of analysis tools. The Los Alamos HCV database emerged as a result of the success of the HIV Database Group and uses a similar approach [212]. The infrastructure developed for the dynamic alignment of a large numbers of sequences in the HIV database was quickly implemented for the HCV database [221]. Tools designed for the manipulation of viral sequences (gene extraction, coordinates to reference strains, sequence alignment, and nucleotide or protein motifs scanning) were easily adapted for the HCV database. The information is accessed via an interface allowing for advanced searches with sorting and graphical overview capabilities. Importantly, sequence data can be retrieved as a DNA alignment. Pre-made, manually optimized alignments are also provided in a variety of flavors. Synonymous/non-synonymous substitution analysis for all genes and proteins is available. The Geography tool can plot genotype frequencies based on their geographical origin. Other functions include glycosylation site identification, *Modeltest*, *Treemaker*, *BLAST*, *PCOORD*, *Gene Cutter*, *Consensus*, *PeptGen*, *Motiscan*, *Primalign*, *Epilign*, *Seq-convert*, *OmniRead*, *SeqPublish* and *Sequence locator* [211,212]. Unfortunately, this valuable resource is no longer funded by the NIH; therefore, the herculean effort to properly maintain it has been significantly delayed by the lack of financial support.

The Virus Pathogen Database and Analysis Resource (ViPR) supports virology researchers studying select agents and other significant public health pathogens belonging to 14 virus families including HCV [222]. Cross-referencing data and integrated computational tools into the online ViPR resource allow complex analyses. ViPR captures data from external and internal sources and makes them accessible through custom searches. ViPR is primarily focused on viruses of human interest; however, families isolated from other host species are also available for comparison. ViPR features a suite of data analysis and visualization tools to perform custom correlative analyses. ViPR uses the NCBI RefSeq strains to extend the manually optimized annotations to the rest of the taxon. Thus, the ViPR resource provides the scientific community with friendly tools capable of performing complex analytical workflows [222].

Importantly, the requirement for the establishment of a database-type of molecular surveillance network that is feasible for international surveillance networks is the standardization of the input data. Several initiatives involving international partnerships and pilot studies have been initiated for an array of pathogens including HCV [219,223,224]. Data sharing is critical for an interdisciplinary approach in order to tackle a problem [223]. The introduction of NGS in routine diagnostics is likely to further boost

HCV integrative surveillance. By applying analytical tools to genomic data for HCV, public health scientists can track specific mutations that confer the ability to resist drugs or link them to transmission networks. However, the information provided by raw genomic sequences of pathogens must be integrated with knowledge about the host biology as well as societal and environmental factors in order to understand the etiology of epidemics and to anticipate their trajectories [223]. Therefore, the development of a diagnostic pipeline critically relies on database exhaustiveness that can match the rapid growth of databases prompted by NGS. A typical blast analysis of millions of sequences after *de novo* assembly into larger *contigs* against the whole NCBI database using relaxed criteria is time- and resource-consuming for diagnostics. In contrast, stringent mapping of non-assembled reads in a comprehensive database in conjunction with long sequence reads will likely improve the overall process. Thus, an HCV integrative molecular surveillance tool should ideally be web-based and include both viral and host factors as well as epidemiological data (Figure 6) [220]. In addition to international public and private sequence databases, the ideal HCV integrative database should be enriched with data obtained from ongoing molecular surveillance projects and outbreak investigations, supplemented if at all possible with sequence information from cases identified in hepatitis clinics and blood banks [117]. As a result, comprehensive analyses including human genetics, phylogenetics and data mining could be easily attained locally by researchers in the field with minimal effort at an affordable cost for local state public health laboratories (Figure 6).

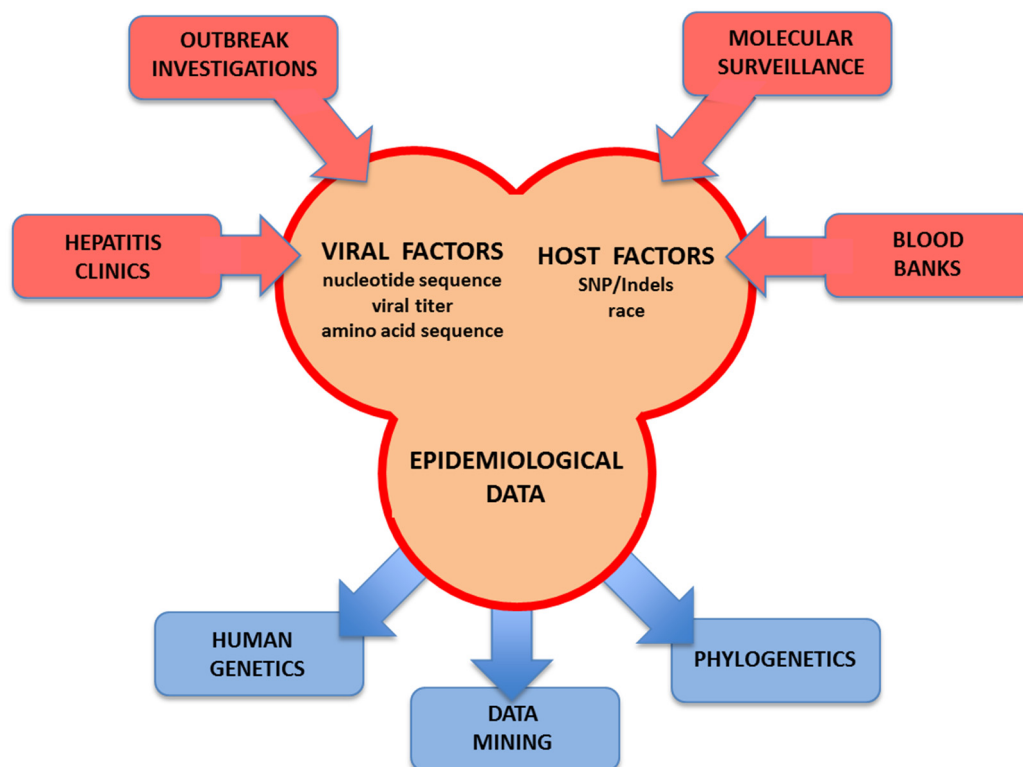**Figure 6.** Advanced HCV molecular surveillance. The main components required for advanced molecular surveillance of HCV are listed. Global databases containing comprehensive information concerning transmission events are used as a source of information to draw conclusions about viral spread. Advanced molecular tools and data mining analyses are required to accurately identify the nuances driving the transmission of HCV.

Data mining, or "Knowledge Discovery in Databases", is an interdisciplinary subfield of computer science referring to the computational process of discovering patterns in large datasets via artificial intelligence, machine learning, statistics, and database systems [225]. The ultimate goal of data mining is the analysis of tremendous volumes of data to discover hidden patterns and relationships in highly complex datasets, rendering valuable information. This is attained by a series of data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, and visualization processes [225]. The application of machine learning methods and advanced statistical modeling to laboratory data and the subsequent discovery of patterns is of relevance for the identification of factors associated with disease outcomes [226]. For HCV infection, decision-tree analysis (a core component of data mining analysis) was used to build predictive models for therapeutic outcomes to antiviral therapy in chronic HCV patients and the severity of disease [227–231]. Decision-tree analysis relies on a tree-shaped structure representing decision "calls" with classifying power, where each internal node denotes an attribute, each branch represents an outcome of the test, and each terminal node a class [227]. Decision-tree analysis facilitates the classification of patients into subgroups that can identify the possibility of an outcome of HCV therapy and, thereby, improve diagnostics [232–235].

Other approaches, such as linear progression and Bayesian network analyses, have also been shown to produce robust predictions for the progression to fibrosis [229]. Interestingly, HCV epistasis is strongly associated with host factors [230], suggesting that intrahost viral evolution is convergent and that important traits can be modeled, resulting in predictions with clinical relevance.

7. Concluding Remarks

HCV molecular evolution in many ways affects virus spread and disease transmission. The sophistication achieved by HCV in shaping its molecular evolutionary patterns is intricate and has significantly hindered our ability to fully understand the mechanisms exploited by the virus to ensure transmission. The remarkable HCV mutation rate represents a challenging task for molecular epidemiology. In this new era of advanced sequencing technologies, the implementation of enhanced molecular surveillance is of the utmost importance in order to accurately monitor the circulation of viral strains. Comprehensive molecular studies are required to uncover the key participating elements responsible for virulence. Supplementing molecular data with epidemiological information and host factors significantly improves the accuracy of HCV molecular surveillance. Therefore, the fact that comprehensive epidemiological investigations must be carried out cannot be understated.

Improved assessment of HCV intrahost genetic variability should clarify the pathway towards advanced integrative molecular surveillance and ultimately help to unveil the mechanisms driving viral transmission. For advanced molecular surveillance to be truly effective, implementation of preventive and control measures along with therapeutic interventions should be performed. The past 25 years since the discovery of HCV have been quite a journey, and the future of HCV control looks more promising than ever before.

Acknowledgments

Author Contributions

L.M.G-R. A.E-G., and P.R. designed the study. L.M.G-R., A.E-G., and P.R. wrote the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Mohd Hanafiah, K.; Groeger, J.; Flaxman, A.D.; Wiersma, S.T. Global epidemiology of hepatitis C virus infection: New estimates of age-specific antibody to HCV seroprevalence. *Hepatology* **2013**, *57*, 1333–1342.
2. Alter, M.J. Epidemiology of hepatitis C virus infection. *World J. Gastroenterol.* **2007**, *13*, 2436–2441.
3. Lavanchy, D. The global burden of hepatitis C. *Liver Int.* **2009**, *29* (Suppl. 1), 74–81.
4. Lauer, G.M.; Walker, B.D. Hepatitis C virus infection. *N. Engl. J. Med.* **2001**, *345*, 41–52.
5. McHutchison, J.G.; Bacon, B.R. Chronic hepatitis C: An age wave of disease burden. *Am. J. Manag. Care* **2005**, *11* (Suppl. 10), S286–S295.
6. Stanley, M.L.; Walker, C.; Alter, M.J.; Yi, M. Hepatitis C Virus. In *Fields Virology*; Knipe, D.M., Howley, P.M., Eds.; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2007; Volume 1, p. 1280.
7. Chevaliez, S.; Pawlotsky, J.M. HCV Genome and Life Cycle. In *Hepatitis C Viruses: Genomes and Molecular Biology*; Tan, S.L., Ed.; Horizon Bioscience: Norfolk, UK, 2006.
8. Smith, D.B.; Bukh, J.; Kuiken, C.; Muerhoff, A.S.; Rice, C.M.; Stapleton, J.T.; Simmonds, P. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: Updated criteria and genotype assignment web resource. *Hepatology* **2014**, *59*, 318–327.
9. Lavanchy, D. Evolving epidemiology of hepatitis C virus. *Clin. Microbiol. Infect.* **2011**, *17*, 107–115.
10. Agha, S.; Tanaka, Y.; Saudy, N.; Kurbanov, F.; Abo-Zeid, M.; El-Malky, M.; Khalaf, M.; Ohta, N.; Yoshizawa, H.; Mizokami, M. Reliability of hepatitis C virus core antigen assay for detection of viremia in HCV genotypes 1, 2, 3, and 4 infected blood donors: A collaborative study between Japan, Egypt, and Uzbekistan. *J. Med. Virol.* **2004**, *73*, 216–222.
11. McOmish, F.; Yap, P.L.; Dow, B.C.; Follett, E.A.; Seed, C.; Keller, A.J.; Cobain, T.J.; Krusius, T.; Kolho, E.; Naukkarinen, R.; *et al*. Geographical distribution of hepatitis C virus genotypes in blood donors: An international collaborative survey. *J. Clin. Microbiol.* **1994**, *32*, 884–892.

12. Kao, J.H.; Chen, P.J.; Lai, M.Y.; Yang, P.M.; Sheu, J.C.; Wang, T.H.; Chen, D.S. Genotypes of hepatitis C virus in Taiwan and the progression of liver disease. *J. Clin. Gastroenterol.* **1995**, *21*, 233–237.

13. Simmonds, P. The origin of hepatitis C virus. *Curr. Top. Microbiol. Immunol.* **2013**, *369*, 1–15.

14. Iles, J.C.; Raghwani, J.; Harrison, G.L.; Pepin, J.; Djoko, C.F.; Tamoufe, U.; LeBreton, M.; Schneider, B.S.; Fair, J.N.; Tshala, F.M.; *et al*. Phylogeography and epidemic history of hepatitis C virus genotype 4 in Africa. *Virology* **2014**, *464–465*, 233–243.

15. Pybus, O.G.; Barnes, E.; Taggart, R.; Lemey, P.; Markov, P.V.; Rasachak, B.; Syhavong, B.; Phetsouvanah, R.; Sheridan, I.; Humphreys, I.S.; *et al*. Genetic history of hepatitis C virus in East Asia. *J. Virol.* **2009**, *83*, 1071–1082.

16. Cruz-Rivera, M.; Carpio-Pedroza, J.C.; Escobar-Gutierrez, A.; Lozano, D.; Vergara-Castaneda, A.; Rivera-Osorio, P.; Martinez-Guarneros, A.; Chacon, C.A.; Fonseca-Coronado, S.; Vaughan, G. Rapid hepatitis C virus divergence among chronically infected individuals. *J. Clin. Microbiol.* **2013**, *51*, 629–632.

17. Honegger, J.R.; Kim, S.; Price, A.A.; Kohout, J.A.; McKnight, K.L.; Prasad, M.R.; Lemon, S.M.; Grakoui, A.; Walker, C.M. Loss of immune escape mutations during persistent HCV infection in pregnancy enhances replication of vertically transmitted viruses. *Nat. Med.* **2013**, *19*, 1529–1533.

18. Martell, M.; Esteban, J.I.; Quer, J.; Genesca, J.; Weiner, A.; Esteban, R.; Guardia, J.; Gomez, J. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: Quasispecies nature of HCV genome distribution. *J. Virol.* **1992**, *66*, 3225–3229.

19. Von Hahn, T.; Yoon, J.C.; Alter, H.; Rice, C.M.; Rehermann, B.; Balfe, P.; McKeating, J.A. Hepatitis C virus continuously escapes from neutralizing antibody and T-cell responses during chronic infection *in vivo*. *Gastroenterology* **2007**, *132*, 667–678.

20. Ralston, R.; Jacobson, I.; Scull, M. The conundrum of relapse in STAT-C therapy: Does HCV play the Red Queen or Rip Van Winkle? *Semin. Liver Dis.* **2011**, *31*, 410–419.

21. Khudyakov, Y. Molecular surveillance of hepatitis C. *Antivir. Ther.* **2012**, *17*, 1465–1470.

22. Depledge, D.P.; Gray, E.R.; Kundu, S.; Cooray, S.; Poulsen, A.; Aaby, P.; Breuer, J. Evolution of cocirculating varicella-zoster virus genotypes during a chickenpox outbreak in Guinea-Bissau. *J. Virol.* **2014**, *88*, 13936–13946.

23. Krugman, S.; Ward, R.; Giles, J.P. The natural history of infectious hepatitis. *Am. J. Med.* **1962**, *32*, 717–728.

24. Feinstone, S.M.; Kapikian, A.Z.; Purceli, R.H. Hepatitis A: Detection by immune electron microscopy of a viruslike antigen associated with acute illness. *Science* **1973**, *182*, 1026–1028.

25. Bayer, M.E.; Blumberg, B.S.; Werner, B. Particles associated with Australia antigen in the sera of patients with leukaemia, Down's Syndrome and hepatitis. *Nature* **1968**, *218*, 1057–1059.

26. Choo, Q.L.; Kuo, G.; Weiner, A.J.; Overby, L.R.; Bradley, D.W.; Houghton, M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* **1989**, *244*, 359–362.

27. Hollinger, F.B.; Gitnick, G.L.; Aach, R.D.; Szmuness, W.; Mosley, J.W.; Stevens, C.E.; Peters, R.L.; Weiner, J.M.; Werch, J.B.; Lander, J.J. Non-A, non-B hepatitis transmission in chimpanzees: A project of the transfusion-transmitted viruses study group. *Intervirology* **1978**, *10*, 60–68.

28. Miyamura, T.; Saito, I.; Katayama, T.; Kikuchi, S.; Tateda, A.; Houghton, M.; Choo, Q.L.; Kuo, G. Detection of antibody against antigen expressed by molecularly cloned hepatitis C virus cDNA: Application to diagnosis and blood screening for posttransfusion hepatitis. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 983–987.

29. Kubo, Y.; Takeuchi, K.; Boonmar, S.; Katayama, T.; Choo, Q.L.; Kuo, G.; Weiner, A.J.; Bradley, D.W.; Houghton, M.; Saito, I.; *et al.* A cDNA fragment of hepatitis C virus isolated from an implicated donor of post-transfusion non-A, non-B hepatitis in Japan. *Nucleic Acids Res.* **1989**, *17*, 10367–10372.

30. Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Coit, D.; Medina-Selby, R.; Barr, P.J.; *et al.* Genetic organization and diversity of the hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2451–2455.

31. Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimotohno, K. Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9524–9528.

32. Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Shimotohno, K. Sequence diversity of hepatitis C viral genomes. *Mol. Biol. Med.* **1990**, *7*, 495–501.

33. Grakoui, A.; McCourt, D.W.; Wychowski, C.; Feinstone, S.M.; Rice, C.M. Characterization of the hepatitis C virus-encoded serine proteinase: Determination of proteinase-dependent polyprotein cleavage sites. *J. Virol.* **1993**, *67*, 2832–2843.

34. Hijikata, M.; Mizushima, H.; Akagi, T.; Mori, S.; Kakiuchi, N.; Kato, N.; Tanaka, T.; Kimura, K.; Shimotohno, K. Two distinct proteinase activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus. *J. Virol.* **1993**, *67*, 4665–4675.

35. Eckart, M.R.; Selby, M.; Masiarz, F.; Lee, C.; Berger, K.; Crawford, K.; Kuo, C.; Kuo, G.; Houghton, M.; Choo, Q.L. The hepatitis C virus encodes a serine protease involved in processing of the putative nonstructural proteins from the viral polyprotein precursor. *Biochem. Biophys. Res. Commun.* **1993**, *192*, 399–406.

36. Suzich, J.A.; Tamura, J.K.; Palmer-Hill, F.; Warrener, P.; Grakoui, A.; Rice, C.M.; Feinstone, S.M.; Collett, M.S. Hepatitis C virus NS3 protein polynucleotide-stimulated nucleoside triphosphatase and comparison with the related pestivirus and flavivirus enzymes. *J. Virol.* **1993**, *67*, 6152–6158.

37. Lohmann, V.; Korner, F.; Herian, U.; Bartenschlager, R. Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *J. Virol.* **1997**, *71*, 8416–8428.

38. Simmonds, P. Variability of hepatitis C virus. *Hepatology* **1995**, *21*, 570–583.

39. Bukh, J.; Purcell, R.H.; Miller, R.H. Sequence analysis of the core gene of 14 hepatitis C virus genotypes. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 8239–8243.

40. Robertson, B.; Myers, G.; Howard, C.; Brettin, T.; Bukh, J.; Gaschen, B.; Gojobori, T.; Maertens, G.; Mizokami, M.; Nainan, O.; *et al.* Classification, nomenclature, and database development for hepatitis C virus (HCV) and related viruses: Proposals for standardization. International Committee on Virus Taxonomy. *Arch. Virol.* **1998**, *143*, 2493–2503.

41. Patel, K.; McHutchison, J.G. Initial treatment for chronic hepatitis C: Current therapies and their optimal dosing and duration. *Clevel. Clin. J. Med.* **2004**, *71* (Suppl. 3), S8–S12.

42. Perales, C.; Beach, N.M.; Sheldon, J.; Domingo, E. Molecular basis of interferon resistance in hepatitis C virus. *Curr. Opin. Virol.* **2014**, *8*, 38–44.

43. Lohmann, V.; Korner, F.; Koch, J.; Herian, U.; Theilmann, L.; Bartenschlager, R. Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line. *Science* **1999**, *285*, 110–113.

44. Pietschmann, T.; Lohmann, V.; Rutter, G.; Kurpanek, K.; Bartenschlager, R. Characterization of cell lines carrying self-replicating hepatitis C virus RNAs. *J. Virol.* **2001**, *75*, 1252–1264.

45. Blight, K.J.; Kolykhalov, A.A.; Rice, C.M. Efficient initiation of HCV RNA replication in cell culture. *Science* **2000**, *290*, 1972–1974.

46. Wakita, T.; Pietschmann, T.; Kato, T.; Date, T.; Miyamoto, M.; Zhao, Z.; Murthy, K.; Habermann, A.; Krausslich, H.G.; Mizokami, M.; *et al.* Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat. Med.* **2005**, *11*, 791–796.

47. Jones, C.T.; Murray, C.L.; Eastman, D.K.; Tassello, J.; Rice, C.M. Hepatitis C virus p7 and NS2 proteins are essential for production of infectious virus. *J. Virol.* **2007**, *81*, 8374–8383.

48. Luik, P.; Chew, C.; Aittoniemi, J.; Chang, J.; Wentworth, P., Jr.; Dwek, R.A.; Biggin, P.C.; Venien-Bryan, C.; Zitzmann, N. The 3-dimensional structure of a hepatitis C virus p7 ion channel by electron microscopy. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12712–12716.

49. Tellinghuisen, T.L.; Foss, K.L.; Treadaway, J. Regulation of hepatitis C virion production via phosphorylation of the NS5A protein. *PLOS Pathog.* **2008**, *4*, e1000032.

50. Barth, H.; Schafer, C.; Adah, M.I.; Zhang, F.; Linhardt, R.J.; Toyoda, H.; Kinoshita-Toyoda, A.; Toida, T.; van Kuppevelt, T.H.; Depla, E.; *et al.* Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. *J. Biol. Chem.* **2003**, *278*, 41003–41012.

51. Scarselli, E.; Ansuini, H.; Cerino, R.; Roccasecca, R.M.; Acali, S.; Filocamo, G.; Traboni, C.; Nicosia, A.; Cortese, R.; Vitelli, A. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J.* **2002**, *21*, 5017–5025.

52. Evans, M.J.; von Hahn, T.; Tscherne, D.M.; Syder, A.J.; Panis, M.; Wolk, B.; Hatziioannou, T.; McKeating, J.A.; Bieniasz, P.D.; Rice, C.M. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature* **2007**, *446*, 801–805.

53. Ploss, A.; Evans, M.J.; Gaysinskaya, V.A.; Panis, M.; You, H.; de Jong, Y.P.; Rice, C.M. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature* **2009**, *457*, 882–886.

54. Lupberger, J.; Zeisel, M.B.; Xiao, F.; Thumann, C.; Fofana, I.; Zona, L.; Davis, C.; Mee, C.J.; Turek, M.; Gorke, S.; *et al.* EGFR and EphA2 are host factors for hepatitis C virus entry and possible targets for antiviral therapy. *Nat. Med.* **2011**, *17*, 589–595.

55. Sainz, B., Jr.; Barretto, N.; Martin, D.N.; Hiraga, N.; Imamura, M.; Hussain, S.; Marsh, K.A.; Yu, X.; Chayama, K.; Alrefai, W.A.; *et al.* Identification of the Niemann-Pick C1-like 1 cholesterol absorption receptor as a new hepatitis C virus entry factor. *Nat. Med.* **2012**, *18*, 281–285.

56. Bull, R.A.; Luciani, F.; McElroy, K.; Gaudieri, S.; Pham, S.T.; Chopra, A.; Cameron, B.; Maher, L.; Dore, G.J.; White, P.A.; Lloyd, A.R., Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. *PLOS Pathog.* **2011**, *7*, e1002243.

57. Nasu, A.; Marusawa, H.; Ueda, Y.; Nishijima, N.; Takahashi, K.; Osaki, Y.; Yamashita, Y.; Inokuma, T.; Tamada, T.; Fujiwara, T.; *et al.* Genetic heterogeneity of hepatitis C virus in association with antiviral therapy determined by ultra-deep sequencing. *PLOS ONE* **2011**, *6*, e24907.

58. Ninomiya, M.; Ueno, Y.; Funayama, R.; Nagashima, T.; Nishida, Y.; Kondo, Y.; Inoue, J.; Kakazu, E.; Kimura, O.; Nakayama, K.; *et al.* Use of illumina deep sequencing technology to differentiate hepatitis C virus variants. *J. Clin. Microbiol.* **2012**, *50*, 857–866.

59. Forbi, J.C.; Purdy, M.A.; Campo, D.S.; Vaughan, G.; Dimitrova, Z.E.; Ganova-Raeva, L.M.; Xia, G.L.; Khudyakov, Y.E. Epidemic history of hepatitis C virus infection in two remote communities in Nigeria, West Africa. *J. Gen. Virol.* **2012**, *93*, 1410–1421.

60. Svarovskaia, E.S.; Martin, R.; McHutchison, J.G.; Miller, M.D.; Mo, H. Abundant drug-resistant NS3 mutants detected by deep sequencing in hepatitis C virus-infected patients undergoing NS3 protease inhibitor monotherapy. *J. Clin. Microbiol.* **2012**, *50*, 3267–3274.

61. Dietz, J.; Schelhorn, S.E.; Fitting, D.; Mihm, U.; Susser, S.; Welker, M.W.; Fuller, C.; Daumer, M.; Teuber, G.; Wedemeyer, H.; *et al.* Deep sequencing reveals mutagenic effects of ribavirin during monotherapy of hepatitis C virus genotype 1-infected patients. *J. Virol.* **2013**, *87*, 6172–6181.

62. Akuta, N.; Suzuki, F.; Seko, Y.; Kawamura, Y.; Sezaki, H.; Suzuki, Y.; Hosaka, T.; Kobayashi, M.; Hara, T.; Saitoh, S.; *et al.* Emergence of telaprevir-resistant variants detected by ultra-deep sequencing after triple therapy in patients infected with HCV genotype 1. *J. Med. Virol.* **2013**, *85*, 1028–1036.

63. Trimoulet, P.; Pinson, P.; Papuchon, J.; Foucher, J.; Vergniol, J.; Chermak, F.; Wittkop, L.; Castaing, N.; Merrouche, W.; Reigadas, S.; *et al.* Dynamic and rapid changes in viral quasispecies by UDPS in chronic hepatitis C patients receiving telaprevir-based therapy. *Antivir. Ther.* **2013**, *18*, 723–727.

64. Caraballo Cortes, K.; Zagordi, O.; Laskus, T.; Ploski, R.; Bukowska-Osko, I.; Pawelczyk, A.; Berak, H.; Radkowski, M. Ultradeep pyrosequencing of hepatitis C virus hypervariable region 1 in quasispecies analysis. *BioMed Res. Int.* **2013**, *2013*, e626083.

65. Akuta, N.; Suzuki, F.; Fukushima, T.; Kawamura, Y.; Sezaki, H.; Suzuki, Y.; Hosaka, T.; Kobayashi, M.; Hara, T.; Saitoh, S.; *et al.* Prediction of treatment efficacy and telaprevir-resistant variants after triple therapy in patients infected with hepatitis C virus genotype 1. *J. Clin. Microbiol.* **2013**, *51*, 2862–2868.

66. Kirst, M.E.; Li, E.C.; Wang, C.X.; Dong, H.J.; Liu, C.; Fried, M.W.; Nelson, D.R.; Wang, G.P. Deep sequencing analysis of HCV NS3 resistance-associated variants and mutation linkage in liver transplant recipients. *PLOS ONE* **2013**, *8*, e69698.

67. Miura, M.; Maekawa, S.; Takano, S.; Komatsu, N.; Tatsumi, A.; Asakawa, Y.; Shindo, K.; Amemiya, F.; Nakayama, Y.; Inoue, T.; *et al.* Deep-sequencing analysis of the association between the quasispecies nature of the hepatitis C virus core region and disease progression. *J. Virol.* **2013**, *87*, 12541–12551.

68. Bartolini, B.; Giombini, E.; Zaccaro, P.; Selleri, M.; Rozera, G.; Abbate, I.; Comandini, U.V.; Ippolito, G.; Solmone, M.; Capobianchi, M.R. Extent of HCV NS3 protease variability and resistance-associated mutations assessed by next generation sequencing in HCV monoinfected and HIV/HCV coinfected patients. *Virus Res.* **2013**, *177*, 205–208.

69. Wu, S.; Kanda, T.; Nakamoto, S.; Jiang, X.; Miyamura, T.; Nakatani, S.M.; Ono, S.K.; Takahashi-Nakaguchi, A.; Gonoi, T.; Yokosuka, O. Prevalence of hepatitis C virus subgenotypes 1a and 1b in Japanese patients: Ultra-deep sequencing analysis of HCV NS5B genotype-specific region. *PLOS ONE* **2013**, *8*, e73615.

70. Campo, D.S.; Skums, P.; Dimitrova, Z.; Vaughan, G.; Forbi, J.C.; Teo, C.G.; Khudyakov, Y.; Lau, D.T. Drug resistance of a viral population and its individual intrahost variants during the first 48 h of therapy. *Clin. Pharmacol. Ther.* **2014**, *95*, 627–635.

71. Abdelrahman, T.; Hughes, J.; Main, J.; McLauchlan, J.; Thursz, M.; Thomson, E. Next generation sequencing sheds light on the natural history of hepatitis C infection in patients that fail treatment. *Hepatology* **2014**, *61*, 88–97.

72. Zagordi, O.; Bhattacharya, A.; Eriksson, N.; Beerenwinkel, N. ShoRAH: Estimating the genetic diversity of a mixed sample from next-generation sequencing data. *BMC Bioinform.* **2011**, *12*, e119.

73. Mancuso, N.; Tork, B.; Skums, P.; Ganova-Raeva, L.; Mandoiu, I.; Zelikovsky, A. Reconstructing viral quasispecies from NGS amplicon reads. *Silico Biol.* **2011**, *11*, 237–249.

74. Skums, P.; Mancuso, N.; Artyomenko, A.; Tork, B.; Mandoiu, I.; Khudyakov, Y.; Zelikovsky, A. Reconstruction of viral population structure from next-generation sequencing data using multicommodity flows. *BMC Bioinform.* **2013**, *14* (Suppl. 9), eS2.

75. Ge, D.; Fellay, J.; Thompson, A.J.; Simon, J.S.; Shianna, K.V.; Urban, T.J.; Heinzen, E.L.; Qiu, P.; Bertelsen, A.H.; Muir, A.J.; *et al*. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* **2009**, *461*, 399–401.

76. Suppiah, V.; Moldovan, M.; Ahlenstiel, G.; Berg, T.; Weltman, M.; Abate, M.L.; Bassendine, M.; Spengler, U.; Dore, G.J.; Powell, E.; *et al*. IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat. Genet.* **2009**, *41*, 1100–1104.

77. Tanaka, Y.; Nishida, N.; Sugiyama, M.; Kurosaki, M.; Matsuura, K.; Sakamoto, N.; Nakagawa, M.; Korenaga, M.; Hino, K.; Hige, S.; *et al*. Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat. Genet.* **2009**, *41*, 1105–1109.

78. Tungol, A.; Rademacher, K.; Schafer, J.A. Formulary management of the protease inhibitors boceprevir and telaprevir for chronic hepatitis C virus. *J. Manag. Care Pharm.* **2011**, *17*, 685–694.

79. Chao, D.; Botwin, G.J.; Morgan, T.R. Update on recently approved treatments for hepatitis C. *Curr. Treat. Options Gastroenterol.* **2014**, *12*, 211–228.

80. Lawitz, E.; Mangia, A.; Wyles, D.; Rodriguez-Torres, M.; Hassanein, T.; Gordon, S.C.; Schultz, M.; Davis, M.N.; Kayali, Z.; Reddy, K.R.; *et al*. Sofosbuvir for previously untreated chronic hepatitis C infection. *N. Engl. J. Med.* **2013**, *368*, 1878–1887.

81. Gane, E.J.; Stedman, C.A.; Hyland, R.H.; Ding, X.; Svarovskaia, E.; Symonds, W.T.; Hindes, R.G.; Berrey, M.M. Nucleotide polymerase inhibitor sofosbuvir plus ribavirin for hepatitis C. *N. Engl. J. Med.* **2013**, *368*, 34–44.

82. Hayashi, N.; Izumi, N.; Kumada, H.; Okanoue, T.; Tsubouchi, H.; Yatsuhashi, H.; Kato, M.; Ki, R.; Komada, Y.; Seto, C.; *et al*. Simeprevir with peginterferon/ribavirin for treatment-naive hepatitis C genotype 1 patients in Japan: CONCERTO-1, a phase III trial. *J. Hepatol.* **2014**, *61*, 219–227.

83. Pawlotsky, J.M. New hepatitis C therapies: The toolbox, strategies, and challenges. *Gastroenterology* **2014**, *146*, 1176–1192.

84. Barnes, E.; Folgori, A.; Capone, S.; Swadling, L.; Aston, S.; Kurioka, A.; Meyer, J.; Huddart, R.; Smith, K.; Townsend, R.; *et al*. Novel adenovirus-based vaccines induce broad and sustained T cell responses to HCV in man. *Sci. Trans. Med.* **2012**, *4*, e115ra1.

85. Honegger, J.R.; Zhou, Y.; Walker, C.M. Will there be a vaccine to prevent HCV infection? *Semin. Liver Dis.* **2014**, *34*, 79–88.

86. Armstrong, G.L. Commentary: Modelling the epidemiology of hepatitis C and its complications. *Int. J. Epidemiol.* **2003**, *32*, 725–726.

87. Wilke, C.O. Quasispecies theory in the context of population genetics. *BMC Evol. Biol.* **2005**, *5*, e44.

88. Gismondi, M.I.; Becker, P.D.; Diaz Carrasco, J.M.; Guzman, C.A.; Campos, R.H.; Preciado, M.V. Evolution of hepatitis C virus hypervariable region 1 in immunocompetent children born to HCV-infected mothers. *J. Viral Hepat.* **2009**, *16*, 332–339.

89. Gismondi, M.I.; Diaz Carrasco, J.M.; Valva, P.; Becker, P.D.; Guzman, C.A.; Campos, R.H.; Preciado, M.V. Dynamic changes in viral population structure and compartmentalization during chronic hepatitis C virus infection in children. *Virology* **2013**, *447*, 187–196.

90. Fischer, G.E.; Schaefer, M.K.; Labus, B.J.; Sands, L.; Rowley, P.; Azzam, I.A.; Armour, P.; Khudyakov, Y.E.; Lin, Y.; Xia, G.; *et al*. Hepatitis C virus infections from unsafe injection practices at an endoscopy clinic in Las Vegas, Nevada, 2007–2008. *Clin. Infect. Dis.* **2010**, *51*, 267–273.

91. Warner, A.E.; Schaefer, M.K.; Patel, P.R.; Drobeniuc, J.; Xia, G.; Lin, Y.; Khudyakov, Y.; Vonderwahl, C.W.; Miller, L.; Thompson, N.D. Outbreak of hepatitis C virus infection associated with narcotics diversion by an hepatitis C virus-infected surgical technician. *Am. J. Infect. Control* **2015**, *43*, 53–58.

92. Gonzalez-Candelas, F.; Bracho, M.A.; Wrobel, B.; Moya, A. Molecular evolution in court: Analysis of a large hepatitis C virus outbreak from an evolving source. *BMC Biol.* **2013**, *11*, e76.

93. Escobar-Gutierrez, A.; Vazquez-Pichardo, M.; Cruz-Rivera, M.; Rivera-Osorio, P.; Carpio-Pedroza, J.C.; Ruiz-Pacheco, J.A.; Ruiz-Tovar, K.; Vaughan, G. Identification of hepatitis C virus transmission using a next-generation sequencing approach. *J. Clin. Microbiol.* **2012**, *50*, 1461–1463.

94. Ray, S.C.; Fanning, L.; Wang, X.H.; Netski, D.M.; Kenny-Walsh, E.; Thomas, D.L. Divergent and convergent evolution after a common-source outbreak of hepatitis C virus. *J. Exp. Med.* **2005**, *201*, 1753–1759.

95. Escobar-Gutierrez, A.; Soudeyns, H.; Larouche, A.; Carpio-Pedroza, J.C.; Martinez-Guarneros, A.; Vazquez-Chacon, C.A.; Fonseca-Coronado, S.; Yamasaki, L.H.; Ruiz-Tovar, K.; Cruz-Rivera, M.Y. Vertical transmission of hepatitis C virus: A tale of multiple outcomes. *Infect. Genet. Evol.* **2013**, *20*, 465–470.

96. Larouche, A.; Gaetan, G.; El-Bilali, N.; Quesnel-Vallieres, M.; Martin, S.R.; Alvarez, F.; Shoukry, N.H.; Soudeyns, H. Seronegative hepatitis C virus infection in a child infected via mother-to-child transmission. *J. Clin. Microbiol.* **2012**, *50*, 2515–2519.

97. Wang, G.P.; Sherrill-Mix, S.A.; Chang, K.M.; Quince, C.; Bushman, F.D. Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. *J. Virol.* **2010**, *84*, 6218–6228.

98. Cubero, M.; Gregori, J.; Esteban, J.I.; Garcia-Cehic, D.; Bes, M.; Perales, C.; Domingo, E.; Rodriguez-Frias, F.; Sauleda, S.; Casillas, R.; *et al*. Identification of host and viral factors involved in a dissimilar resolution of a hepatitis C virus infection. *Liver Int.* **2014**, *34*, 896–906.

99. Manrubia, S.C.; Escarmis, C.; Domingo, E.; Lazaro, E. High mutation rates, bottlenecks, and robustness of RNA viral quasispecies. *Gene* **2005**, *347*, 273–282.

100. D'Arienzo, V.; Moreau, A.; D'Alteroche, L.; Gissot, V.; Blanchard, E.; Gaudy-Graffin, C.; Roch, E.; Dubois, F.; Giraudeau, B.; Plantier, J.C.; *et al*. Sequence and functional analysis of the envelope glycoproteins of hepatitis C virus variants selectively transmitted to a new host. *J. Virol.* **2013**, *87*, 13609–13618.

101. Li, H.; Stoddard, M.B.; Wang, S.; Blair, L.M.; Giorgi, E.E.; Parrish, E.H.; Learn, G.H.; Hraber, P.; Goepfert, P.A.; Saag, M.S.; *et al*. Elucidation of hepatitis C virus transmission and early diversification by single genome sequencing. *PLOS Pathog.* **2012**, *8*, e1002880.

102. Kato, N.; Ootsuyama, Y.; Sekiya, H.; Ohkoshi, S.; Nakazawa, T.; Hijikata, M.; Shimotohno, K. Genetic drift in hypervariable region 1 of the viral genome in persistent hepatitis C virus infection. *J. Virol.* **1994**, *68*, 4776–4784.

103. Van Doorn, L.J.; Quint, W.; Tsiquaye, K.; Voermans, J.; Paelinck, D.; Kos, T.; Maertens, G.; Schellekens, H.; Murray, K. Longitudinal analysis of hepatitis C virus infection and genetic drift of the hypervariable region. *J. Infect. Dis.* **1994**, *169*, 1226–1235.

104. Okamoto, H.; Kojima, M.; Okada, S.; Yoshizawa, H.; Iizuka, H.; Tanaka, T.; Muchmore, E.E.; Peterson, D.A.; Ito, Y.; Mishiro, S. Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: Variability and stability. *Virology* **1992**, *190*, 894–899.

105. Allain, J.P.; Dong, Y.; Vandamme, A.M.; Moulton, V.; Salemi, M. Evolutionary rate and genetic drift of hepatitis C virus are not correlated with the host immune response: Studies of infected donor-recipient clusters. *J. Virol.* **2000**, *74*, 2541–2549.

106. Shi, W.; Freitas, I.T.; Zhu, C.; Zheng, W.; Hall, W.W.; Higgins, D.G. Recombination in hepatitis C virus: Identification of four novel naturally occurring inter-subtype recombinants. *PLOS ONE* **2012**, *7*, e41997.

107. Tscherne, D.M.; Evans, M.J.; von Hahn, T.; Jones, C.T.; Stamataki, Z.; McKeating, J.A.; Lindenbach, B.D.; Rice, C.M. Superinfection exclusion in cells infected with hepatitis C virus. *J. Virol.* **2007**, *81*, 3693–3703.

108. Webster, B.; Ott, M.; Greene, W.C. Evasion of superinfection exclusion and elimination of primary viral RNA by an adapted strain of hepatitis C virus. *J. Virol.* **2013**, *87*, 13354–13369.

109. Webster, B.; Wissing, S.; Herker, E.; Ott, M.; Greene, W.C. Rapid intracellular competition between hepatitis C viral genomes as a result of mitosis. *J. Virol.* **2013**, *87*, 581–596.

110. Kalinina, O.; Norder, H.; Mukomolov, S.; Magnius, L.O. A natural intergenotypic recombinant of hepatitis C virus identified in St. Petersburg. *J. Virol.* **2002**, *76*, 4034–4043.

111. Sentandreu, V.; Jimenez-Hernandez, N.; Torres-Puente, M.; Bracho, M.A.; Valero, A.; Gosalbes, M.J.; Ortega, E.; Moya, A.; Gonzalez-Candelas, F. Evidence of recombination in intrapatient populations of hepatitis C virus. *PLOS ONE* **2008**, *3*, e3239.

112. Colina, R.; Casane, D.; Vasquez, S.; Garcia-Aguirre, L.; Chunga, A.; Romero, H.; Khan, B.; Cristina, J. Evidence of intratypic recombination in natural populations of hepatitis C virus. *J. Gen. Virol.* **2004**, *85*, 31–37.

113. Kalinina, O.; Norder, H.; Magnius, L.O. Full-length open reading frame of a recombinant hepatitis C virus strain from St Petersburg: Proposed mechanism for its formation. *J. Gen. Virol.* **2004**, *85*, 1853–1857.

114. Gonzalez-Candelas, F.; Lopez-Labrador, F.X.; Bracho, M.A. Recombination in hepatitis C virus. *Viruses* **2011**, *3*, 2006–2024.

115. Ramachandran, S.; Campo, D.S.; Dimitrova, Z.E.; Xia, G.L.; Purdy, M.A.; Khudyakov, Y.E. Temporal variations in the hepatitis C virus intrahost population during chronic infection. *J. Virol.* **2011**, *85*, 6369–6380.

116. Gray, R.R.; Salemi, M.; Klenerman, P.; Pybus, O.G. A new evolutionary model for hepatitis C virus chronic infection. *PLOS Pathog.* **2012**, *8*, e1002656.

117. Goncalves Rossi, L.M.; Rahal, P. Challenges in molecular epidemiology of hepatitis C virus. *J. Clin. Virol.* **2014**, *60*, 174–176.

118. Araujo, A.C. Antibody- and genome-based identification of recent HCV infection. *Antivir. Ther.* **2012**, *17*, 1459–1464.

119. Vandamme, A.M.; Pybus, O.G. Viral phylogeny in court: The unusual case of the Valencian anesthetist. *BMC Biol.* **2013**, *11*, e83.

120. Abecasis, A.B.; Geretti, A.M.; Albert, J.; Power, L.; Weait, M.; Vandamme, A.M. Science in court: The myth of HIV fingerprinting. *Lancet Infect. Dis.* **2011**, *11*, 78–79.

121. Roque-Afonso, A.M.; Ducoulombier, D.; Di Liberto, G.; Kara, R.; Gigou, M.; Dussaix, E.; Samuel, D.; Feray, C. Compartmentalization of hepatitis C virus genotypes between plasma and peripheral blood mononuclear cells. *J. Virol.* **2005**, *79*, 6349–6357.

122. Indolfi, G.; Azzari, C.; Resti, M. Hepatitis: Immunoregulation in pregnancy and perinatal transmission of HCV. *Nat. Rev. Gastroenterol. Hepatol.* **2014**, *11*, 6–7.

123. Imran, M.; Manzoor, S.; Ashraf, J.; Khalid, M.; Tariq, M.; Khaliq, H.M.; Azam, S. Role of viral and host factors in interferon based therapy of hepatitis C virus infection. *Virol. J.* **2013**, *10*, 299.

124. Villano, S.A.; Vlahov, D.; Nelson, K.E.; Cohn, S.; Thomas, D.L. Persistence of viremia and the importance of long-term follow-up after acute hepatitis C infection. *Hepatology* **1999**, *29*, 908–914.

125. Micallef, J.M.; Kaldor, J.M.; Dore, G.J. Spontaneous viral clearance following acute hepatitis C infection: A systematic review of longitudinal studies. *J. Viral Hepat.* **2006**, *13*, 34–41.

126. Saito, T.; Ueno, Y. Transmission of hepatitis C virus: Self-limiting hepatitis or chronic hepatitis? *World J. Gastroenterol.* **2013**, *19*, 6957–6961.

127. Diepolder, H.M.; Zachoval, R.; Hoffmann, R.M.; Wierenga, E.A.; Santantonio, T.; Jung, M.C.; Eichenlaub, D.; Pape, G.R. Possible mechanism involving T-lymphocyte response to non-structural protein 3 in viral clearance in acute hepatitis C virus infection. *Lancet* **1995**, *346*, 1006–1007.

128. Ulsenheimer, A.; Gerlach, J.T.; Gruener, N.H.; Jung, M.C.; Schirren, C.A.; Schraut, W.; Zachoval, R.; Pape, G.R.; Diepolder, H.M. Detection of functionally altered hepatitis C virus-specific CD4 T cells in acute and chronic hepatitis C. *Hepatology* **2003**, *37*, 1189–1198.

129. Urbani, S.; Amadei, B.; Fisicaro, P.; Tola, D.; Orlandini, A.; Sacchelli, L.; Mori, C.; Missale, G.; Ferrari, C. Outcome of acute hepatitis C is related to virus-specific CD4 function and maturation of antiviral memory CD8 responses. *Hepatology* **2006**, *44*, 126–139.

130. Puig, M.; Mihalik, K.; Tilton, J.C.; Williams, O.; Merchlinsky, M.; Connors, M.; Feinstone, S.M.; Major, M.E. CD4+ immune escape and subsequent T-cell failure following chimpanzee immunization against hepatitis C virus. *Hepatology* **2006**, *44*, 736–745.

131. Brooks, D.G.; McGavern, D.B.; Oldstone, M.B. Reprogramming of antiviral T cells prevents inactivation and restores T cell activity during persistent viral infection. *J. Clin. Investig.* **2006**, *116*, 1675–1685.

132. Bes, M.; Sauleda, S.; Casamitjana, N.; Piron, M.; Campos-Varela, I.; Quer, J.; Cubero, M.; Puig, L.; Guardia, J.; Esteban, J.I. Reversal of nonstructural protein 3-specific CD4(+) T cell dysfunction in patients with persistent hepatitis C virus infection. *J. Viral Hepat.* **2012**, *19*, 283–294.

133. Thursz, M.; Yallop, R.; Goldin, R.; Trepo, C.; Thomas, H.C. Influence of MHC class II genotype on outcome of infection with hepatitis C virus. *Lancet* **1999**, *354*, 2119–2124.

134. Thomas, D.L.; Astemborski, J.; Rai, R.M.; Anania, F.A.; Schaeffer, M.; Galai, N.; Nolt, K.; Nelson, K.E.; Strathdee, S.A.; Johnson, L.; *et al*. The natural history of hepatitis C virus infection: Host, viral, and environmental factors. *JAMA* **2000**, *284*, 450–456.

135. Thio, C.L.; Thomas, D.L.; Goedert, J.J.; Vlahov, D.; Nelson, K.E.; Hilgartner, M.W.; O'Brien, S.J.; Karacki, P.; Marti, D.; Astemborski, J.; *et al*. Racial differences in HLA class II associations with hepatitis C virus outcomes. *J. Infect. Dis.* **2001**, *184*, 16–21.

136. Khakoo, S.I.; Thio, C.L.; Martin, M.P.; Brooks, C.R.; Gao, X.; Astemborski, J.; Cheng, J.; Goedert, J.J.; Vlahov, D.; Hilgartner, M.; *et al*. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science* **2004**, *305*, 872–874.

137. Grody, W.W. Molecular genetic risk screening. *Ann. Rev. Med.* **2003**, *54*, 473–490.

138. Marsh, S.; van Booven, D.J.; McLeod, H.L. Global pharmacogenetics: Giving the genome to the masses. *Pharmacogenomics* **2006**, *7*, 625–631.

139. Mosbruger, T.L.; Duggal, P.; Goedert, J.J.; Kirk, G.D.; Hoots, W.K.; Tobler, L.H.; Busch, M.; Peters, M.G.; Rosen, H.R.; Thomas, D.L.; *et al*. Large-scale candidate gene analysis of spontaneous clearance of hepatitis C virus. *J. Infect. Dis.* **2010**, *201*, 1371–1380.

140. Thio, C.L.; Goedert, J.J.; Mosbruger, T.; Vlahov, D.; Strathdee, S.A.; O'Brien, S.J.; Astemborski, J.; Thomas, D.L. An analysis of tumor necrosis factor alpha gene polymorphisms and haplotypes with natural clearance of hepatitis C virus infection. *Genes Immun.* **2004**, *5*, 294–300.

141. Lio, D.; Caruso, C.; di Stefano, R.; Colonna Romano, G.; Ferraro, D.; Scola, L.; Crivello, A.; Licata, A.; Valenza, L.M.; Candore, G.; *et al*. IL-10 and TNF-alpha polymorphisms and the recovery from HCV infection. *Hum. Immunol.* **2003**, *64*, 674–680.

142. Vassalli, P. The pathophysiology of tumor necrosis factors. *Ann. Rev. Immunol.* **1992**, *10*, 411–452.

143. Koziel, M.J.; Dudley, D.; Afdhal, N.; Grakoui, A.; Rice, C.M.; Choo, Q.L.; Houghton, M.; Walker, B.D. HLA class I-restricted cytotoxic T lymphocytes specific for hepatitis C virus. Identification of multiple epitopes and characterization of patterns of cytokine release. *J. Clin. Investig.* **1995**, *96*, 2311–2321.

144. Kimura, T.; Saito, T.; Yoshimura, M.; Yixuan, S.; Baba, M.; Ji, G.; Muramatsu, M.; Kawata, S. Association of transforming growth factor-beta 1 functional polymorphisms with natural clearance of hepatitis C virus. *J. Infect. Dis.* **2006**, *193*, 1371–1374.

145. Huang, Y.; Yang, H.; Borg, B.B.; Su, X.; Rhodes, S.L.; Yang, K.; Tong, X.; Tang, G.; Howell, C.D.; Rosen, H.R.; *et al*. A functional SNP of interferon-gamma gene is important for interferon-alpha-induced and spontaneous recovery from hepatitis C virus infection. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 985–990.

146. Rook, A.H.; Kehrl, J.H.; Wakefield, L.M.; Roberts, A.B.; Sporn, M.B.; Burlington, D.B.; Lane, H.C.; Fauci, A.S. Effects of transforming growth factor beta on the functions of natural killer cells: Depressed cytolytic activity and blunting of interferon responsiveness. *J. Immunol.* **1986**, *136*, 3916–3920.

147. Okumoto, K.; Hattori, E.; Tamura, K.; Kiso, S.; Watanabe, H.; Saito, K.; Saito, T.; Togashi, H.; Kawata, S. Possible contribution of circulating transforming growth factor-beta1 to immunity and prognosis in unresectable hepatocellular carcinoma. *Liver Int.* **2004**, *24*, 21–28.

148. Frese, M.; Schwarzle, V.; Barth, K.; Krieger, N.; Lohmann, V.; Mihm, S.; Haller, O.; Bartenschlager, R. Interferon-gamma inhibits replication of subgenomic and genomic hepatitis C virus RNAs. *Hepatology* **2002**, *35*, 694–703.

149. Woollard, D.J.; Grakoui, A.; Shoukry, N.H.; Murthy, K.K.; Campbell, K.J.; Walker, C.M. Characterization of HCV-specific Patr class II restricted CD4+ T cell responses in an acutely infected chimpanzee. *Hepatology* **2003**, *38*, 1297–1306.

150. Piazzolla, G.; Tortorella, C.; Schiraldi, O.; Antonaci, S. Relationship between interferon-gamma, interleukin-10, and interleukin-12 production in chronic hepatitis C and *in vitro* effects of interferon-alpha. *J. Clin. Immunol.* **2000**, *20*, 54–61.

151. Edwards-Smith, C.J.; Jonsson, J.R.; Purdie, D.M.; Bansal, A.; Shorthouse, C.; Powell, E.E. Interleukin-10 promoter polymorphism predicts initial response of chronic hepatitis C to interferon alfa. *Hepatology* **1999**, *30*, 526–530.

152. Nelson, D.R.; Lauwers, G.Y.; Lau, J.Y.; Davis, G.L. Interleukin 10 treatment reduces fibrosis in patients with chronic hepatitis C: A pilot trial of interferon nonresponders. *Gastroenterology* **2000**, *118*, 655–660.

153. Naito, M.; Matsui, A.; Inao, M.; Nagoshi, S.; Nagano, M.; Ito, N.; Egashira, T.; Hashimoto, M.; Mishiro, S.; Mochida, S.; *et al*. SNPs in the promoter region of the osteopontin gene as a marker predicting the efficacy of interferon-based therapies in patients with chronic hepatitis C. *J. Gastroenterol.* **2005**, *40*, 381–388.

154. McKee, M.D.; Nanci, A. Osteopontin: An interfacial extracellular matrix protein in mineralized tissues. *Connect. Tissue Res.* **1996**, *35*, 197–205.

155. El Awady, M.K.; Anany, M.A.; Esmat, G.; Zayed, N.; Tabll, A.A.; Helmy, A.; el Zayady, A.R.; Abdalla, M.S.; Sharada, H.M.; el Raziky, M.; *et al*. Single nucleotide polymorphism at exon 7 splice acceptor site of OAS1 gene determines response of hepatitis C virus patients to interferon therapy. *J. Gastroenterol. Hepatol.* **2011**, *26*, 843–850.

156. Hamano, E.; Hijikata, M.; Itoyama, S.; Quy, T.; Phi, N.C.; Long, H.T.; Ha, L.D.; Ban, V.V.; Matsushita, I.; Yanai, H.; *et al*. Polymorphisms of interferon-inducible genes OAS-1 and MxA associated with SARS in the Vietnamese population. *Biochem. Biophys. Res. Commun.* **2005**, *329*, 1234–1239.

157. Alves Pedroso, M.L.; Boldt, A.B.; Pereira-Ferrari, L.; Steffensen, R.; Strauss, E.; Jensenius, J.C.; Ioshii, S.O.; Messias-Reason, I. Mannan-binding lectin MBL2 gene polymorphism in chronic hepatitis C: Association with the severity of liver fibrosis and response to interferon therapy. *Clin. Exp. Immunol.* **2008**, *152*, 258–264.

158. Apostolopoulos, V.; McKenzie, I.F. Role of the mannose receptor in the immune response. *Curr. Mol. Med.* **2001**, *1*, 469–474.

159. Sheppard, P.; Kindsvogel, W.; Xu, W.; Henderson, K.; Schlutsmeyer, S.; Whitmore, T.E.; Kuestner, R.; Garrigues, U.; Birks, C.; Roraback, J.; *et al*. IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nat. Immunol.* **2003**, *4*, 63–68.

160. Prokunina-Olsson, L.; Muchmore, B.; Tang, W.; Pfeiffer, R.M.; Park, H.; Dickensheets, H.; Hergott, D.; Porter-Gill, P.; Mumy, A.; Kohaar, I.; *et al*. A variant upstream of IFNL3 (IL28B) creating a new interferon gene IFNL4 is associated with impaired clearance of hepatitis C virus. *Nat. Genet.* **2013**, *45*, 164–171.

161. Gad, H.H.; Dellgren, C.; Hamming, O.J.; Vends, S.; Paludan, S.R.; Hartmann, R. Interferon-lambda is functionally an interferon but structurally related to the interleukin-10 family. *J. Biol. Chem.* **2009**, *284*, 20869–20875.

162. Iversen, M.B.; Ank, N.; Melchjorsen, J.; Paludan, S.R. Expression of type III interferon (IFN) in the vaginal mucosa is mediated primarily by dendritic cells and displays stronger dependence on NF-kappaB than type I IFNs. *J. Virol.* **2010**, *84*, 4579–4586.

163. Marcello, T.; Grakoui, A.; Barba-Spaeth, G.; Machlin, E.S.; Kotenko, S.V.; MacDonald, M.R.; Rice, C.M. Interferons alpha and lambda inhibit hepatitis C virus replication with distinct signal transduction and gene regulation kinetics. *Gastroenterology* **2006**, *131*, 1887–1898.

164. Megjugorac, N.J.; Gallagher, G.E.; Gallagher, G. Modulation of human plasmacytoid DC function by IFN-lambda1 (IL-29). *J. Leukoc. Biol.* **2009**, *86*, 1359–1363.

165. Mennechet, F.J.; Uze, G. Interferon-lambda-treated dendritic cells specifically induce proliferation of FOXP3-expressing suppressor T cells. *Blood* **2006**, *107*, 4417–4423.

166. Wolk, K.; Witte, K.; Witte, E.; Proesch, S.; Schulze-Tanzil, G.; Nasilowska, K.; Thilo, J.; Asadullah, K.; Sterry, W.; Volk, H.D.; *et al*. Maturing dendritic cells are an important source of IL-29 and IL-20 that may cooperatively increase the innate immunity of keratinocytes. *J. Leukoc. Biol.* **2008**, *83*, 1181–1193.

167. Preciado, M.V.; Valva, P.; Escobar-Gutierrez, A.; Rahal, P.; Ruiz-Tovar, K.; Yamasaki, L.; Vazquez-Chacon, C.; Martinez-Guarneros, A.; Carpio-Pedroza, J.C.; Fonseca-Coronado, S.; *et al*. Hepatitis C virus molecular evolution: Transmission, disease progression and antiviral therapy. *World J. Gastroenterol.* **2014**, *20*, 15992–16013.

168. Cruz-Rivera, M.; Forbi, J.C.; Yamasaki, L.H.; Vazquez-Chacon, C.A.; Martinez-Guarneros, A.; Carpio-Pedroza, J.C.; Escobar-Gutierrez, A.; Ruiz-Tovar, K.; Fonseca-Coronado, S.; Vaughan, G. Molecular epidemiology of viral diseases in the era of next generation sequencing. *J. Clin. Virol.* **2013**, *57*, 378–380.

169. Lauck, M.; Alvarado-Mora, M.V.; Becker, E.A.; Bhattacharya, D.; Striker, R.; Hughes, A.L.; Carrilho, F.J.; O'Connor, D.H.; Pinho, J.R. Analysis of hepatitis C virus intrahost diversity across the coding region by ultradeep pyrosequencing. *J. Virol.* **2012**, *86*, 3952–3960.

170. Beerenwinkel, N.; Gunthard, H.F.; Roth, V.; Metzner, K.J. Challenges and opportunities in estimating viral genetic diversity from next-generation sequencing data. *Front. Microbiol.* **2012**, *3*, e329.

171. Ganova-Raeva, L.M.; Dimitrova, Z.E.; Campo, D.S.; Khudyakov, Y. Application of mass spectrometry to molecular surveillance of hepatitis B and C viral infections. *Antivir. Ther.* **2012**, *17*, 1477–1482.

172. Ganova-Raeva, L.M.; Dimitrova, Z.E.; Campo, D.S.; Lin, Y.; Ramachandran, S.; Xia, G.L.; Honisch, C.; Cantor, C.R.; Khudyakov, Y.E. Detection of hepatitis C virus transmission by use of DNA mass spectrometry. *J. Infect. Dis.* **2013**, *207*, 999–1006.

173. Fonseca-Coronado, S.; Escobar-Gutierrez, A.; Ruiz-Tovar, K.; Cruz-Rivera, M.Y.; Rivera-Osorio, P.; Vazquez-Pichardo, M.; Carpio-Pedroza, J.C.; Ruiz-Pacheco, J.A.; Cazares, F.; Vaughan, G. Specific detection of naturally occurring hepatitis C virus mutants with resistance to telaprevir and boceprevir (protease inhibitors) among treatment-naive infected individuals. *J. Clin. Microbiol.* **2012**, *50*, 281–287.

174. Fonseca-Coronado, S.; Vaughan, G.; Cruz-Rivera, M.Y.; Carpio-Pedroza, J.C.; Ruiz-Tovar, K.; Ruiz-Pacheco, J.A.; Escobar-Gutierrez, A. Interleukin-28B genotyping by melt-mismatch amplification mutation assay PCR analysis using single nucleotide polymorphisms rs12979860 and rs8099917, a useful tool for prediction of therapy response in hepatitis C patients. *J. Clin. Microbiol.* **2011**, *49*, 2706–2710.

175. Enomoto, N.; Kurosaki, M.; Tanaka, Y.; Marumo, F.; Sato, C. Fluctuation of hepatitis C virus quasispecies in persistent infection and interferon treatment revealed by single-strand conformation polymorphism analysis. *J. Gen. Virol.* **1994**, *75*, 1361–1369.

176. Kurosaki, M.; Enomoto, N.; Marumo, F.; Sato, C. Evolution and selection of hepatitis C virus variants in patients with chronic hepatitis C. *Virology* **1994**, *205*, 161–169.

177. Moribe, T.; Hayashi, N.; Kanazawa, Y.; Mita, E.; Fusamoto, H.; Negi, M.; Kaneshige, T.; Igimi, H.; Kamada, T.; Uchida, K. Hepatitis C viral complexity detected by single-strand conformation polymorphism and response to interferon therapy. *Gastroenterology* **1995**, *108*, 789–795.

178. Hayashi, K. PCR-SSCP: A simple and sensitive method for detection of mutations in the genomic DNA. *PCR Methods Appl.* **1991**, *1*, 34–38.

179. Vera-Otarola, J.; Barria, M.I.; Leon, U.; Carvallo, P.; Soza, A.; Lopez-Lastra, M. Is single-strand conformation polymorphism analysis of the full 5' untranslated region an adequate approach to study hepatitis C virus quasispecies distribution? *J. Virol.* **2009**, *83*, 9018–9021.

180. Laskus, T.; Wang, L.F.; Rakela, J.; Vargas, H.; Pinna, A.D.; Tsamandas, A.C.; Demetris, A.J.; Fung, J. Dynamic behavior of hepatitis C virus in chronically infected patients receiving liver graft from infected donors. *Virology* **1996**, *220*, 171–176.

181. Kukita, Y.; Tahira, T.; Sommer, S.S.; Hayashi, K. SSCP analysis of long DNA fragments in low pH gel. *Hum. Mutat.* **1997**, *10*, 400–407.

182. Delwart, E.L.; Gordon, C.J. Tracking changes in HIV-1 envelope quasispecies using DNA heteroduplex analysis. *Methods* **1997**, *12*, 348–354.

183. Wilson, J.J.; Polyak, S.J.; Day, T.D.; Gretch, D.R. Characterization of simple and complex hepatitis C virus quasispecies by heteroduplex gel shift analysis: Correlation with nucleotide sequencing. *J. Gen. Virol.* **1995**, *76*, 1763–1771.

184. Bower, M.A.; Spencer, M.; Matsumura, S.; Nisbet, R.E.; Howe, C.J. How many clones need to be sequenced from a single forensic or ancient DNA sample in order to determine a reliable consensus sequence? *Nucleic Acids Res.* **2005**, *33*, 2549–2556.

185. McCaughan, G.W.; Laskus, T.; Vargas, H.E. Hepatitis C virus quasispecies: Misunderstood and mistreated? *Liver Transplant.* **2003**, *9*, 1048–1052.

186. Stein, C.A.; Levantis, P.; Oxford, J.S. Sequence analysis of proviral HIV RT amplified directly by a semi-quantitative technique from AZT treated patients. *J. Med. Virol.* **1994**, *44*, 115–121.

187. Ison, M.G.; Llata, E.; Conover, C.S.; Friedewald, J.J.; Gerber, S.I.; Grigoryan, A.; Heneine, W.; Millis, J.M.; Simon, D.M.; Teo, C.G.; *et al*. Transmission of human immunodeficiency virus and hepatitis C virus from an organ donor to four transplant recipients. *Am. J. Transplant.* **2011**, *11*, 1218–1225.

188. Ramachandran, S.; Zhai, X.; Thai, H.; Campo, D.S.; Xia, G.; Ganova-Raeva, L.M.; Drobeniuc, J.; Khudyakov, Y.E. Evaluation of intra-host variants of the entire hepatitis B virus genome. *PLOS ONE* **2011**, *6*, e25232.

189. Leushner, J.; Chiu, N.H. Automated mass spectrometry: A revolutionary technology for clinical diagnostics. *Mol. Diagn.* **2000**, *5*, 341–348.

190. Karas, M.; Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **1988**, *60*, 2299–2301.

191. Fenn, J.B.; Mann, M.; Meng, C.K.; Wong, S.F.; Whitehouse, C.M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246*, 64–71.

192. Dimitrova, Z.; Campo, D.S.; Ramachandran, S.; Vaughan, G.; Ganova-Raeva, L.; Lin, Y.; Forbi, J.C.; Xia, G.; Skums, P.; Pearlman, B.; *et al*. Evaluation of viral heterogeneity using next-generation sequencing, end-point limiting-dilution and mass spectrometry. *Silico biol.* **2011**, *11*, 183–192.

193. Bocker, S. SNP and mutation discovery using base-specific cleavage and MALDI-TOF mass spectrometry. *Bioinformatics* **2003**, *19* (Suppl. 1), i44–i53.

194. Yea, C.; Bukh, J.; Ayers, M.; Roberts, E.; Krajden, M.; Tellier, R. Monitoring of hepatitis C virus quasispecies in chronic infection by matrix-assisted laser desorption ionization-time of flight mass spectrometry mutation detection. *J. Clin. Microbiol.* **2007**, *45*, 1053–1057.

195. Hellberg, R.S.; Li, F.; Sampath, R.; Yasuda, I.J.; Carolan, H.E.; Wolfe, J.M.; Brown, M.K.; Alexander, R.C.; Williams-Hill, D.M.; Martin, W.B. Rapid detection and differentiation of human noroviruses using RT-PCR coupled to electrospray ionization mass spectrometry. *Food Microbiol.* **2014**, *44*, 71–80.

196. Dunne, E.M.; Ong, E.K.; Moser, R.J.; Siba, P.M.; Phuanukoonnon, S.; Greenhill, A.R.; Robins-Browne, R.M.; Mulholland, E.K.; Satzke, C. Multilocus sequence typing of *Streptococcus pneumoniae* by use of mass spectrometry. *J. Clin. Microbiol.* **2011**, *49*, 3756–3760.

197. Lecuit, M.; Eloit, M. The diagnosis of infectious diseases by whole genome next generation sequencing: A new era is opening. *Front. Cell. Infect. Microbiol.* **2014**, *4*, e25.

198. Tombacz, D.; Sharon, D.; Olah, P.; Csabai, Z.; Snyder, M.; Boldogkoi, Z. Strain Kaplan of Pseudorabies Virus Genome Sequenced by PacBio Single-Molecule Real-Time Sequencing Technology. *Genome Announc.* **2014**, *2*, 8–14.

199. Giallonardo, F.D.; Topfer, A.; Rey, M.; Prabhakaran, S.; Duport, Y.; Leemann, C.; Schmutz, S.; Campbell, N.K.; Joos, B.; Lecca, M.R.; *et al*. Full-length haplotype reconstruction to infer the structure of heterogenous virus populations. *Nucleic Acids Res.* **2014**, *42*, e115.

200. Brinzevich, D.; Young, G.R.; Sebra, R.; Ayllon, J.; Maio, S.M.; Deikus, G.; Chen, B.K.; Fernandez-Sesma, A.; Simon, V.; Mulder, L.C. HIV-1 interacts with human endogenous retrovirus K (HML-2) envelopes derived from human primary lymphocytes. *J. Virol.* **2014**, *88*, 6213–6223.

201. Rodriguez-Frias, F.; Tabernero, D.; Quer, J.; Esteban, J.I.; Ortega, I.; Domingo, E.; Cubero, M.; Camos, S.; Ferrer-Costa, C.; Sanchez, A.; *et al*. Ultra-deep pyrosequencing detects conserved genomic sites and quantifies linkage of drug-resistant amino acid changes in the hepatitis B virus genome. *PLOS ONE* **2012**, *7*, e37874.

202. Aw, T.G.; Howe, A.; Rose, J.B. Metagenomic approaches for direct and cell culture evaluation of the virological quality of wastewater. *J. Virol. Methods* **2014**, *210C*, 15–21.

203. Idris, A.; Al-Saleh, M.; Piatek, M.J.; Al-Shahwan, I.; Ali, S.; Brown, J.K. Viral metagenomics: Analysis of begomoviruses by illumina high-throughput sequencing. *Viruses* **2014**, *6*, 1219–1236.

204. Chen, G.Q.; Zhuang, Q.Y.; Wang, K.C.; Liu, S.; Shao, J.Z.; Jiang, W.M.; Hou, G.Y.; Li, J.P.; Yu, J.M.; Li, Y.P.; *et al*. Identification and survey of a novel avian coronavirus in ducks. *PLOS ONE* **2013**, *8*, e72918.

205. Callegaro, A.; Di Filippo, E.; Astuti, N.; Ortega, P.A.; Rizzi, M.; Farina, C.; Valenti, D.; Maggiolo, F. Early clinical response and presence of viral resistant minority variants: A proof of concept study. *J. Int. AIDS Soc.* **2014**, *17*, e19759.

206. Gregori, J.; Esteban, J.I.; Cubero, M.; Garcia-Cehic, D.; Perales, C.; Casillas, R.; Alvarez-Tejado, M.; Rodriguez-Frias, F.; Guardia, J.; Domingo, E.; *et al*. Ultra-deep pyrosequencing (UDPS) data treatment to study amplicon HCV minor variants. *PLOS ONE* **2013**, *8*, e83361.

207. Gregori, J.; Salicru, M.; Domingo, E.; Sanchez, A.; Esteban, J.I.; Rodriguez-Frias, F.; Quer, J. Inference with viral quasispecies diversity indices: Clonal and NGS approaches. *Bioinformatics* **2014**, *30*, 1104–1111.

208. Wylie, K.M.; Mihindukulasuriya, K.A.; Sodergren, E.; Weinstock, G.M.; Storch, G.A. Sequence analysis of the human virome in febrile and afebrile children. *PLOS ONE* **2012**, *7*, e27735.

209. Quer, J.; Gregori, J.; Rodriguez-Frias, F.; Buti, M.; Madejon, A.; Perez-del-Pulgar, S.; Garcia-Cehic, D.; Casillas, R.; Blasi, M.; Homs, M.; *et al*. High-resolution hepatitis C virus subtyping using NS5B deep sequencing and phylogeny, an alternative to current methods. *J. Clin. Microbiol.* **2015**, *53*, 219–226.

210. Marz, M.; Beerenwinkel, N.; Drosten, C.; Fricke, M.; Frishman, D.; Hofacker, I.L.; Hoffmann, D.; Middendorf, M.; Rattei, T.; Stadler, P.F.; *et al*. Challenges in RNA virus bioinformatics. *Bioinformatics* **2014**, *30*, 1793–1799.

211. Kuiken, C.; Hraber, P.; Thurmond, J.; Yusim, K. The hepatitis C sequence database in Los Alamos. *Nucleic Acids Res.* **2008**, *36*, D512–D516.

212. Kuiken, C.; Yusim, K.; Boykin, L.; Richardson, R. The Los Alamos hepatitis C sequence database. *Bioinformatics* **2005**, *21*, 379–384.

213. Brister, J.R.; Ako-Adjei, D.; Bao, Y.; Blinkova, O. NCBI Viral Genomes Resource. *Nucleic Acids Res.* **2014**, doi:10.1093/nar/gku1207.

214. Salvatierra, K.; Fareleski, S.; Forcada, A.; Lopez-Labrador, F.X. Hepatitis C virus resistance to new specifically-targeted antiviral therapy: A public health perspective. *World J. Virol.* **2013**, *2*, 6–15.

215. Huson, D.H.; Scornavacca, C. A survey of combinatorial methods for phylogenetic networks. *Genome Biol. Evol.* **2011**, *3*, 23–35.

216. Scornavacca, C.; Zickmann, F.; Huson, D.H. Tanglegrams for rooted phylogenetic trees and networks. *Bioinformatics* **2011**, *27*, i248–i256.

217. Wrobel, B.; Torres-Puente, M.; Jimenez, N.; Bracho, M.A.; Garcia-Robles, I.; Moya, A.; Gonzalez-Candelas, F. Analysis of the overdispersed clock in the short-term evolution of hepatitis C virus: Using the E1/E2 gene sequences to infer infection dates in a single source outbreak. *Mol. Biol. Evol.* **2006**, *23*, 1242–1253.

218. De Oliveira, T.; Pybus, O.G.; Rambaut, A.; Salemi, M.; Cassol, S.; Ciccozzi, M.; Rezza, G.; Gattinara, G.C.; D'Arrigo, R.; Amicosante, M.; *et al*. Molecular epidemiology: HIV-1 and HCV sequences from Libyan outbreak. *Nature* **2006**, *444*, 836–837.

219. Niesters, H.G.; Rossen, J.W.; van der Avoort, H.; Baas, D.; Benschop, K.; Claas, E.C.; Kroneman, A.; van Maarseveen, N.; Pas, S.; van Pelt, W.; *et al*. Laboratory-based surveillance in the molecular era: The TYPENED model, a joint data-sharing platform for clinical and public health laboratories. *Euro Surveill.* **2013**, *18*, 20387.

220. Squires, R.B.; Noronha, J.; Hunt, V.; Garcia-Sastre, A.; Macken, C.; Baumgarth, N.; Suarez, D.; Pickett, B.E.; Zhang, Y.; Larsen, C.N.; *et al*. Influenza research database: An integrated bioinformatics resource for influenza research and surveillance. *Influenza Respir. Viruses* **2012**, *6*, 404–416.

221. Gaschen, B.; Kuiken, C.; Korber, B.; Foley, B. Retrieval and on-the-fly alignment of sequence fragments from the HIV database. *Bioinformatics* **2001**, *17*, 415–418.

222. Pickett, B.E.; Greer, D.S.; Zhang, Y.; Stewart, L.; Zhou, L.; Sun, G.; Gu, Z.; Kumar, S.; Zaremba, S.; Larsen, C.N.; *et al*. Virus pathogen database and analysis resource (ViPR): A comprehensive bioinformatics database and analysis resource for the coronavirus research community. *Viruses* **2012**, *4*, 3209–3226.

223. Janies, D.A.; Voronkin, I.O.; Das, M.; Hardman, J.; Treseder, T.W.; Studer, J. Genome informatics of influenza A: From data sharing to shared analytical capabilities. *Anim. Health Res. Rev.* **2010**, *11*, 73–79.

224. Heisey-Grove, D.M.; Church, D.R.; Haney, G.A.; Demaria, A., Jr. Enhancing surveillance for hepatitis C through public health informatics. *Public Health Rep.* **2011**, *126*, 13–18.

225. Glasgow, J.; Jurisica, I.I.; Ng, R. Data mining and knowledge discovery in molecular databases. *Pac. Symp. Biocomput.* **2000**, *5*, 365–366.

226. Shang, G.; Richardson, A.; Gahan, M.E.; Easteal, S.; Ohms, S.; Lidbury, B.A. Predicting the presence of hepatitis B virus surface antigen in Chinese patients by pathology data mining. *J. Med. Virol.* **2013**, *85*, 1334–1339.

227. Zayed, N.; Awad, A.B.; El-Akel, W.; Doss, W.; Awad, T.; Radwan, A.; Mabrouk, M. The assessment of data mining for the prediction of therapeutic outcome in 3719 Egyptian patients with chronic hepatitis C. *Clin. Res. Hepatol. Gastroenterol.* **2013**, *37*, 254–261.

228. Kurosaki, M.; Hiramatsu, N.; Sakamoto, M.; Suzuki, Y.; Iwasaki, M.; Tamori, A.; Matsuura, K.; Kakinuma, S.; Sugauchi, F.; Sakamoto, N.; *et al.* Data mining model using simple and readily available factors could identify patients at high risk for hepatocellular carcinoma in chronic hepatitis C. *J. Hepatol.* **2012**, *56*, 602–608.

229. Lara, J.; Lopez-Labrador, F.; Gonzalez-Candelas, F.; Berenguer, M.; Khudyakov, Y.E. Computational models of liver fibrosis progression for hepatitis C virus chronic infection. *BMC Bioinform.* **2014**, *15* (Suppl. 8), eS5.

230. Lara, J.; Tavis, J.E.; Donlin, M.J.; Lee, W.M.; Yuan, H.J.; Pearlman, B.L.; Vaughan, G.; Forbi, J.C.; Xia, G.L.; Khudyakov, Y.E. Coordinated evolution among hepatitis C virus genomic sites is coupled to host factors and resistance to interferon. *Silico Biol.* **2011**, *11*, 213–224.

231. Lara, J.; Xia, G.; Purdy, M.; Khudyakov, Y. Coevolution of the hepatitis C virus polyprotein sites in patients on combined pegylated interferon and ribavirin therapy. *J. Virol.* **2011**, *85*, 3649–3663.

232. Kurosaki, M.; Sakamoto, N.; Iwasaki, M.; Sakamoto, M.; Suzuki, Y.; Hiramatsu, N.; Sugauchi, F.; Yatsuhashi, H.; Izumi, N. Pretreatment prediction of response to peginterferon plus ribavirin therapy in genotype 1 chronic hepatitis C using data mining analysis. *J. Gastroenterol.* **2011**, *46*, 401–409.

233. Hiramatsu, N.; Kurosaki, M.; Sakamoto, N.; Iwasaki, M.; Sakamoto, M.; Suzuki, Y.; Sugauchi, F.; Tamori, A.; Kakinnuma, S.; Matsuura, K.; *et al.* Pretreatment prediction of anemia progression by pegylated interferon alpha-2b plus ribavirin combination therapy in chronic hepatitis C infection: Decision-tree analysis. *J. Gastroenterol.* **2011**, *46*, 1111–1119.

234. Kurosaki, M.; Sakamoto, N.; Iwasaki, M.; Sakamoto, M.; Suzuki, Y.; Hiramatsu, N.; Sugauchi, F.; Tamori, A.; Nakagawa, M.; Izumi, N. Sequences in the interferon sensitivity-determining region and core region of hepatitis C virus impact pretreatment prediction of response to PEG-interferon plus ribavirin: Data mining analysis. *J. Med. Virol.* **2011**, *83*, 445–452.

235. Araujo, A.C.; Astrakhantseva, I.V.; Fields, H.A.; Kamili, S. Distinguishing acute from chronic hepatitis C virus (HCV) infection based on antibody reactivities to specific HCV structural and nonstructural proteins. *J. Clin. Microbiol.* **2011**, *49*, 54–57.