



End-point limiting-dilution real-time PCR assay for evaluation of hepatitis C virus quasispecies in serum: Performance under optimal and suboptimal conditions

Sumathi Ramachandran*, Guo-liang Xia, Lilia M. Ganova-Raeva, Omana V. Nainan¹, Yury Khudyakov

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

ARTICLE INFO

Article history:

Received 11 January 2008

Received in revised form 15 April 2008

Accepted 8 May 2008

Available online 20 June 2008

Keywords:

Hepatitis C virus

HVR1

Quasispecies

Limiting-dilution

PCR

Molecular epidemiology

ABSTRACT

An approach for determination of hepatitis C virus (HCV) quasispecies by end-point limiting-dilution real-time PCR (EPLD-PCR) is described. It involves isolation of individual coexisting sequence variants of the hypervariable region 1 (HVR1) of the HCV genome from serum specimens using a limiting-dilution protocol. EPLD-PCR applied to an HCV outbreak study provided insights into the epidemiological relationships between incident and chronic cases. When applied to samples from a longitudinal study of infected patients, HVR1 sequences from each sampling time-point were observed to group as distinct phylogenetic clusters. Melting peak analysis conducted on EPLD-PCR products generated from these patients could be used for evaluation of HVR1 sequence heterogeneity without recourse to clonal sequencing. Further, to better understand the mechanism of single-molecule PCR, experiments were conducted under optimal and suboptimal annealing temperatures. Under all temperature conditions tested, HVR1 variants from the major phylogenetic clusters of quasispecies could be amplified, revealing that successful HVR1 quasispecies analysis is not contingent to dilution of starting cDNA preparations to a single-molecule state. It was found that EPLD-PCR conducted at suboptimal annealing temperatures generated distributions of unique-sequence variants slightly different from the distribution obtained by PCR conducted at the optimal temperature. Hence, EPLD-PCR conditions can be manipulated to access different subpopulations of HCV HVR1 quasispecies, thus, improving the range of the quasispecies detection. Although EPLD-PCR conducted at different conditions detect slightly different quasispecies populations, as was shown in this study, the resulted samples of quasispecies are completely suitable for molecular epidemiological investigation in different clinical and epidemiological settings.

Published by Elsevier B.V.

1. Introduction

An estimated 3.2 million people in the United States and 170 million people worldwide are infected with the hepatitis C virus (HCV) (Armstrong et al., 2006). HCV displays a very high genetic variation both in populations and within infected individuals where it exists as a swarm of genetically related but distinct variants known as quasispecies (Martell et al., 1992). Different parts of the HCV genome are variable to different degrees (Vizmanos et al., 1998). Some regions of the genome such as the non-structural (NS) 5B region displaying moderate heterogeneity are used frequently for epidemiological reconstructions (Lohmann et al., 2000). Other parts of the HCV genome, such as the 'hypervariable' region 1

(HVR1) of the envelope (E) 2 gene, which shows much greater variability and more rapid sequence changes over time, are exploited for the evaluation of changes in the viral quasispecies composition (Sakamoto et al., 1994). Analysis of HCV quasispecies composition has significant clinical implications. For example, carriage of quasispecies constituents with high nucleotide sequence diversity in the HVR1 region has been associated with progression of liver disease (Curran et al., 2002) and resistance to interferon therapy (Zampino et al., 2004).

Different quantitative methods have been used to assess HVR1 quasispecies distribution, but many of them have inherent limitations. Measuring quasispecies spectra has always been a challenge, the two main problems being: (1) the high complexity of mixtures of quasispecies that need to be analyzed; and (2) the difficulty in detecting low-frequency or minority quasispecies constituents. All approaches to the evaluation of quasispecies diversity and complexity can be classified into two groups. One group is based on the identification of the nucleotide sequences in the sample set, and the other group evaluates quasispecies diversity and complexity without the requirement to characterize the sequence,

* Corresponding author at: Centers for Disease Control and Prevention, National Center for Infectious Diseases, Division of Viral Hepatitis, Mailstop A-33, 1600 Clifton Road NE, Atlanta, GA 30333, USA. Tel.: +1 404 639 1403; fax: +1 404 639 1563.

E-mail address: dcq6@cdc.gov (S. Ramachandran).

¹ Deceased.

e.g., single-strand conformation polymorphism (SSCP) analysis, heteroduplex mobility assay (HMA) and matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOFF).

The non-sequencing based quasispecies evaluation assays have two major disadvantages over sequencing methodology. First, they do not allow for evaluation of the primary structure of quasispecies. Second, the accuracy of quasispecies evaluation significantly depends on the physical chemical parameters of DNA molecules employed to separate and identify different quasispecies. Among the non-sequencing techniques, SSCP assays are easy to perform, but are extremely sensitive to changes in electrophoretic conditions, and quasispecies complexity identified using these methods is not proportional to sequence divergence between molecules (Nakano et al., 1997). In HMA, which evaluates diversity by providing an estimation of the Hamming distances (Rossi et al., 2003), several non-predictable factors influence the mobility and thermodynamic stability of the formed DNA hybrids; while hybrid formation is sequence-dependent, the presence of several DNA sequence variants in a given specimen can lead to formation of multiple peaks, and as a result provide an overestimation of DNA heterogeneity (Wilson et al., 1995). In mass spectrometry, the effective mass range of the MALDI instrument imposes constraints on the size of the test sequences (Ayers et al., 2002).

The sequence-based strategy provides the most comprehensive information on quasispecies distribution since in addition to the number and proportional representation of quasispecies in a specimen, these assays allow for evaluating the primary structure of quasispecies. However, there are many challenges associated with the application of these techniques. Besides technical pitfalls, the accuracy of these protocols may be compromised by strategies employed to sample the quasispecies variants. Direct PCR followed by cloning and sequencing (Kono et al., 2003) is used widely. Its usage permits measuring the rate of nucleotide and amino acid substitutions (Morishima et al., 2006). Despite the obvious advantages, post-PCR cloning has, however, serious drawbacks mostly arising from shortfalls of PCR applied to complex mixtures of sequence variants such as template jumping and allelic preference (Meyerhans et al., 1990; Simmonds et al., 1990) or stochastic variations during early cycles of amplifications that lead to PCR drift (Wagner et al., 1994). These shortfalls raise concern that quasispecies distribution evaluation may be distorted by DNA manipulation before cloning.

Single-molecule PCR (sm-PCR) (Jeffreys et al., 1998) is a very sensitive method to detect DNA variants and, more importantly, it appears that when performed in the single-molecule mode, PCR amplification is immune to PCR errors, template jumping and allelic preference (Meyerhans et al., 1990; Simmonds et al., 1990). It was reported that human allele variants can be detected with 85–90% sensitivity by this technique (Stead and Jeffreys, 2000). Sm-PCR has been applied to mutational analysis of microsatellites (Jeffreys et al., 1990), recombination studies (Yauk et al., 2003), construction of expression libraries (Lukyanov et al., 1996), and analysis of cancers for known mutations (Vogelstein and Kinzler, 1999). End-point limiting-dilution PCR (EPLD-PCR), as an approach to sm-PCR, has been used to identify individual pro-viral DNA molecules in samples that carry human immunodeficiency virus (HIV), hepatitis B virus and simian immunodeficiency virus (Stein et al., 1994; Nainan et al., 2002; Schuitemaker and Kootstra, 2005; Greenier et al., 2005). Although the accuracy of the EPLD-PCR in evaluating quasispecies populations, as with techniques based on cloning, is affected by sampling strategies, EPLD-PCR theoretically imposes less distortion. The main reason is that the amplification step is implemented on separated single DNA molecules, whereas the cloning method requires DNA amplification before individual clones are separated.

A real-time-based EPLD-PCR was developed for the detection and sequencing of individual constituents of viral quasispecies that circumvents the need to perform post-PCR cloning to isolate single DNA molecules. It was applied to the evaluation of the HCV HVR1 heterogeneity for the sensitive and large-scale evaluation of the composition of HCV quasispecies in serum. To understand better the mechanism of sm-PCR, we conducted experiments using both optimal and suboptimal PCR conditions. The results show that the success of EPLD-PCR as applied to HVR1 quasispecies analysis is not contingent on dilution of HCV cDNA preparations down to a single DNA molecule.

2. Materials and methods

2.1. Samples and nucleic acid isolation

Assay optimization was conducted using 60 serum samples collected from patients with acute and chronic hepatitis C residing in the United States. The study included 18 specimens collected during a nosocomial HCV outbreak, and 42 serial serum specimens collected from three patients who were followed-up from acute to chronic phase over 8–18 years. Acute cases manifested discrete onset of symptoms consistent with acute viral hepatitis, jaundice and elevated serum alanine aminotransferase levels. Chronic hepatitis C infected persons were anti-HCV positive, and many have mild to severe chronic liver disease (Alter, 1997). Sera had been tested for anti-HCV IgG using the Abbott HCV EIA Version 2.0 (Abbott Laboratories) or the ORTHO HCV Version 3.0 ELISA (Ortho-Clinical Diagnostics Inc., Raritan, NJ). Reactive results were confirmed with supplemental testing using HCV Matrix (Abbott Laboratories) or RIBA 3.0 (Chiron Corp.). All patients were negative for hepatitis B surface antigen and antibodies to HIV. Total nucleic acids from anti-HCV-seropositive specimens were extracted from serum by the use of the Roche MagNA Pure LC instrument and the MagNA Pure LC Total Nucleic Acid Isolation Kit (Roche Diagnostics, Mannheim, Germany), and eluted with 50 μ l of buffer according to the manufacturer's instructions.

2.2. Reverse transcription

RNA was precipitated and reverse-transcribed using both random and specific primers as previously described (Alter et al., 1999). Reverse transcription was carried out for 60 min at 42 °C in a total volume of 20 μ l of 5 \times PCR buffer (Roche), 200 pmol of each deoxynucleotide triphosphate, 1 μ g/ μ l of random primers, 25 units of AMV reverse transcriptase (Roche) and 40 units of RNase inhibitor (Roche), followed by heating at 95 °C for 5 min.

2.3. Quantification and genotyping

HCV RNA was quantified by amplification of the 5'UTR region by real-time nested PCR using an in-house standard. To determine the HCV genotype, a segment of HCV NS5b region, encompassing positions 8275–616, was amplified using real-time nested PCR as described previously (Ramachandran et al., 2006).

2.4. Real-time nested HVR1 PCR

PCR conditions were optimized by manipulating only concentration of primers and PCR cycling conditions. All optimization experiments were conducted using two high HCV titer serum specimens, one obtained from a chronically HCV-infected patient and the other from an experimentally infected chimpanzee (Nainan et al., 2006). Specificity of the HVR1 products was confirmed by

sequence analysis. Primer concentration was titrated (1–25 μM), and an optimal concentration of 15 μM was used for amplifications. Nested “hot-start” PCR reactions were performed in a final volume of 20 μl , using the LightCycler Fast Start DNA Master^{Plus} SYBR Green 1 Kit (Roche Diagnostics, Mannheim, Germany). We added 1 μl of the RT reaction product to the master mix, and nested PCR was performed and optimized using various real-time PCR platforms—Roche Lightcycler 480, ABI 7500, Stratagene MX3005P, and Roche Lightcycler 2.0. High through-put automation of RNA extraction, assay set up and PCR amplification were made possible by the use of MagNA Pure LC system (Roche Diagnostics), automated liquid handler Biomek 3000 (Beckman Coulter) and LC480 Roche PCR system (Roche Diagnostics). The primer sequences to amplify HVR1 segment are as follows: external primers, F1-TGG CTT GGG ATA TGA TGA TGA ACT and R1-GCA GTC CTG TTG ATG TGC CA; and internal primers, F2-GGA TAT GAT GAT GAA CTG GT and R2-ATG TGC CAG CTG CCG TTG GTG T. Each PCR reaction contained 0.75 μM of the primers, and the LightCycler Fast Start DNA Master^{Plus} SYBR Green 1 Mix. After a pre-incubation step at 95 °C for 10 min to activate the Fast Start polymerase, PCR amplification was conducted for 40 cycles each cycle consisting of denaturation (95 °C for 10 s), annealing (63 °C for 10 s) with a single fluorescence measurement taken at the end of the annealing step, and extension (72 °C for 32 s). After amplification, melting-curve analysis was performed by raising the temperature to 95 °C for 15 s, heating the sample at 80 °C for 60 s followed by 95 °C for 15 s, and cooling to 40 °C. Fluorescence data were collected continuously during heating to monitor the dissociation of the product. The derivative melting curves were obtained with the LightCycler data analysis software.

2.5. EPLD-PCR assay for HVR1 quasispecies evaluation

In previous experiments the sensitivity of post-PCR cloning/sequencing and limiting-dilution PCR (LD-PCR) protocols were compared for their sensitivity in detecting differing concentrations of the hepatitis B virus (HBV) variants. LD-PCR was found to be most sensitive and could detect HBV variants at a concentration of 0.1% of the total viral population (Nainan et al., 2002). In the current work, we adapted the LD-PCR protocol to HCV, circumventing the cloning strategy.

EPLD-PCR was performed using serially diluted cDNA preparations. Automated liquid handler Biomek 3000 (Beckman, CA) was used to obtain serial 0.25 log-dilutions of the cDNA preparation to arrive at the approximate end-point dilution, similar to our previous approach to HBV quasispecies analysis (Nainan et al., 2002). The limiting-dilution that resulted in positivity in two out of four replicates was considered to be the one that was limiting. At this dilution, the DNA target templates are presumed to be distributed in a Poisson manner, so that most reactions (50% or less) do not carry template molecules, and thus generate no PCR product. Under such conditions, the positive reactions are most likely to have been initiated from a single template molecule. About 96 PCR reactions were routinely conducted to obtain approximately 48 clones (45–55% positivity) identified using melting-curve analysis. All the HVR1 PCR products were sequenced. The number of clones amplified varied depending on the viral titer or the scope of the investigation. To evaluate the reliability and specificity of the assay, amplifications were carried out at different annealing temperatures (63 °C, 60 °C, 66 °C, 53 °C and 70 °C), and the amplicons were sequenced and compared.

2.6. Sequence analysis

Sequencing reactions were performed using the BigDye v3.1 chemistry sequencing kit (Applied Biosystems, Foster City, CA), and

products were sequenced using an automated sequencer (3130xl Genetic Analyzer, Applied Biosystems). Preliminary sequence analysis was conducted using SeqMan and MegAlign programs from the Lasergene DNA & Protein analysis software (Version 7.0, DNASTAR Inc., Madison, WI) The Accelrys GCG Package (Genetic Computer Group, Version 11.1-UNIX, Accelrys Inc., San Diego, CA) was used for further analysis. Nucleotide sequences were aligned using the GCG multiple alignment program Pileup. HCV genotypes were classified based on the NS5b sequence (Smith et al., 1997), and by comparing each sequence with published reference sequences from GenBank. Phylogenetic trees were constructed by using the neighbor-joining algorithm based on distance matrices generated using the Kimura two-parameter model of nucleotide substitution (PHYMLIP package, v.3.6, written by J. Felsenstein, University of Washington, Seattle) (Tamura et al., 2004). Frequency distributions of pair-wise distances between nucleotide sequences were estimated using the evolution program in the Accelrys GCG Package. Shannon entropy analysis was applied to measure the extent of quasispecies complexity at suboptimal PCR conditions. This entropy measure was calculated as $S = -\sum(p_i \ln p_i)$, which takes into account the frequency of each sequence (p_i) in a given set of quasispecies. The entropy value was then normalized as $S_n = S/\ln N$, to take into consideration the total number of quasispecies sequences (N) analyzed during each PCR SAS for Windows (Version 9.12, SAS Institute Inc., Cary, NC) was used for the comparison of Shannon entropy and genetic distances.

3. Results

3.1. Application to outbreak samples

Applicability of the optimized HCV HVR1 EPLD-PCR assay for molecular epidemiological investigation was validated using 18 specimens collected during a nosocomial HCV outbreak. These specimens were collected from six incident cases, and 12 chronically infected patients belonging to genotypes 1a, 1b and 2b. The number of HVR1 sequences amplified using the EPLD-PCR from each specimen varied from 7 to 71 depending on the titer of the sample. It was found that three cases, A1 ($n=71$), A2 ($n=67$) and A3 ($n=55$), shared sequences with one chronically infected patient C1 ($n=41$), while three other cases, A4 ($n=7$), A5 ($n=51$) and A6 ($n=61$), shared sequences with another chronically infected subject C2 ($n=58$) (Fig. 1). Thus, the EPLD-PCR data suggested that the two chronically infected subjects were the source patients for this HCV outbreak. All sequence electrophoregrams for each HVR1 clone obtained by EPLD-PCR contained no double peaks indicative of no HVR1 mixtures (data not shown). This series of findings confirms that a single sequence can be efficiently obtained from a single limiting-dilution aliquot.

The number of positive clones that can be obtained depends strongly on the viral titer of the sample. For example, in the current study we were able to obtain only seven clones from the incident/acute case A4 due to a low titer. Although only a few clones were sufficient to establish a transmission link in this investigation, in more complex situations when a source case experiences a long HCV evolution and transmissions may occur over an extended period of time, the quasispecies analysis may require many more clones to delineate the link between source and incident cases. Also, for studies looking at complexity, divergence or evolutionary trend in inter- or intra-host (see the following section), a greater population of major and minor variants has to be assessed. With the increasing number of clones for evaluation, the protocol becomes very labor intensive. To make this protocol manageable the automated liquid handling stations Biomek 3000 (Beckman, CA) were used at all stages requiring extensive reagent dispensing.

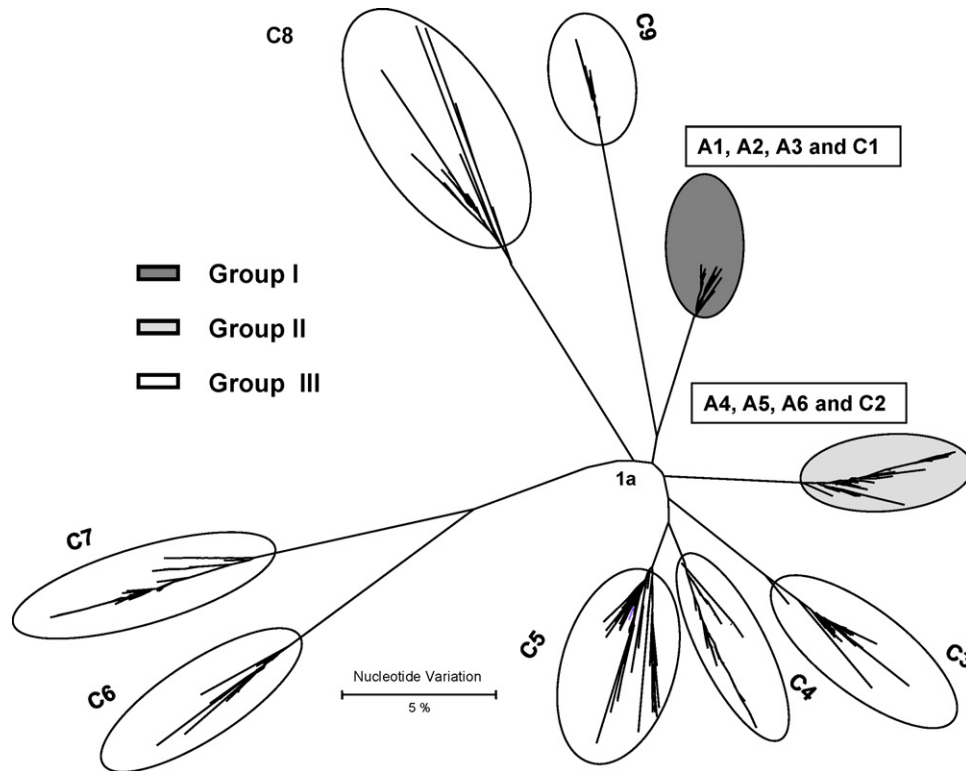


Fig. 1. Phylogenetic tree analysis of HCV HVR1 quasispecies from nine chronically infected patients (C1–9) and six incident cases (A1–6) in an outbreak investigation. Neighbor-joining dendrogram was created by using Kimura's two-parameter model. Each branch represents a different viral sequence, and small distances between branches suggest genetic relatedness. Size of each oval region represents diversity of HVR1 sequences from that specimen or group of specimens. Sequences belonging to A1–3 share sequences with C1 (Group I). Three other cases (A4–6) cluster with C2 (Group II). All other chronic cases (Group III) show independent clustering. Only unique sequence patterns are shown in the tree. All subjects represented here belong to genotype 1a. Bar at bottom denotes nucleotide variation (%).

3.2. Analysis of serial samples from chronically infected patients

The EPLD-PCR strategy was applied to study sequence diversity and complexity in specimens collected from three patients (Patients 1–3). HCV genotyping revealed that Patients 1 and 3 were infected with HCV genotype 1a, and Patient 2 with genotype 1b. Phylogenetic analysis of the HVR1 sequences recovered from amplified PCR fragments showed that these sequences could be grouped into distinct clusters specific for each patient (Fig. 2). The phylogenetic tree in Fig. 2 shows unique variants only. In total, Patient 1 had 70 unique variants out of 122 clones, Patient 2—136 unique variants from 173 clones and Patient 3—115 unique variants from 145 clones. No less than 10 HVR1 variants could be detected at each time-point. It is noted that divergence between HVR1 sequences observed in this study in a given patient reached up to 17%, while a previous study showed that divergence between HVR1 sequences from unrelated cases may be as low as 5% (Xia et al., 2006).

3.3. Melting peak analysis

Comparison of melting curves obtained at various time-points for Patients 1–3 revealed a distinct pattern of independent melting peaks that was specific to each time-point (Fig. 3). For Patient 1, each peak corresponded to a specific cluster of HVR1 quasispecies in the phylogenetic tree, with three clusters being found at separate time-points (Fig. 2). Thus, the melting peak pattern relates to changes in quasispecies diversity over time as phylogenetically determined. Further, the melting-curve patterns obtained immediately at the end of the amplification process allowed an assessment of the

heterogeneity of the HVR1 variants without recourse to sequence analysis.

3.4. Performance at suboptimal PCR conditions

In the EPLD-PCR assay, templates are amplified in separate reactions in a single-molecule mode. It implies that any deviation from optimal conditions may lead to the loss of PCR sensitivity and, as a result, to amplification of a mixed DNA population derived from more than one DNA molecule (Kraytsberg and Khrapko, 2005). To evaluate the effect of suboptimal conditions on performance of the HCV HVR1 EPLD-PCR, we carried out PCR experiments at a range of annealing temperatures (53–70 °C) using serum from Patient 1.

At the optimal annealing temperature of 63 °C, PCR sensitivity was an order of magnitude higher than at 53 °C and 70 °C (Fig. 4). At 60 °C and 66 °C, there was a similar level of detection sensitivity to that of 63 °C, but the amplification take-off-cycles were delayed (data not shown). A total of 123 HVR1 sequences were obtained from four experiments (one each at 53 °C and 70 °C, duplicate experiments at optimal temperature of 63 °C) from the study sample, which represented 65 sequence variants (Fig. 5). The two experiments conducted at the optimal annealing temperature generated 17 and 21 sequence variants, and those at 53 °C and 70 °C generated 34 and 14 sequence variants, respectively. As shown in Fig. 5, only 11 out of 65 sequence variants were shared between experiments.

Phylogenetic analysis of the HVR1 sequences obtained by EPLD-PCR revealed three sequence clusters designated A, B and C in Fig. 5. There was no statistically significant difference in the distribution of quasispecies obtained from the different experiments

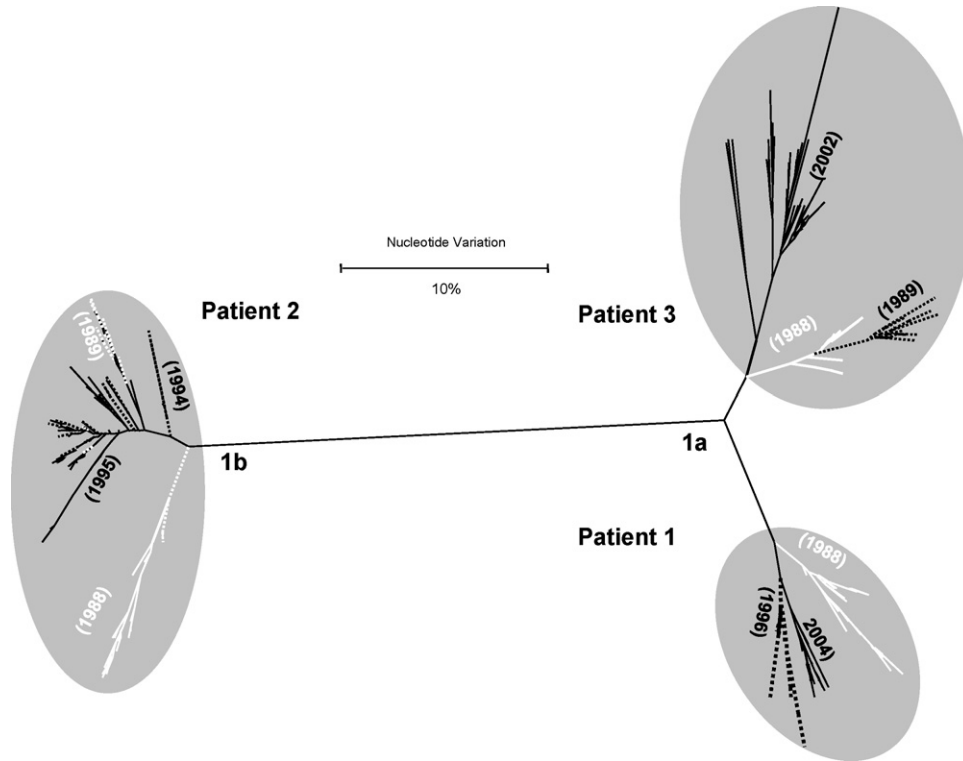


Fig. 2. Phylogenetic tree analysis of HCV HVR1 quasispecies from three chronic HCV patients followed over a period of 8–15 years. The phylogenetic tree shows unique variants only.

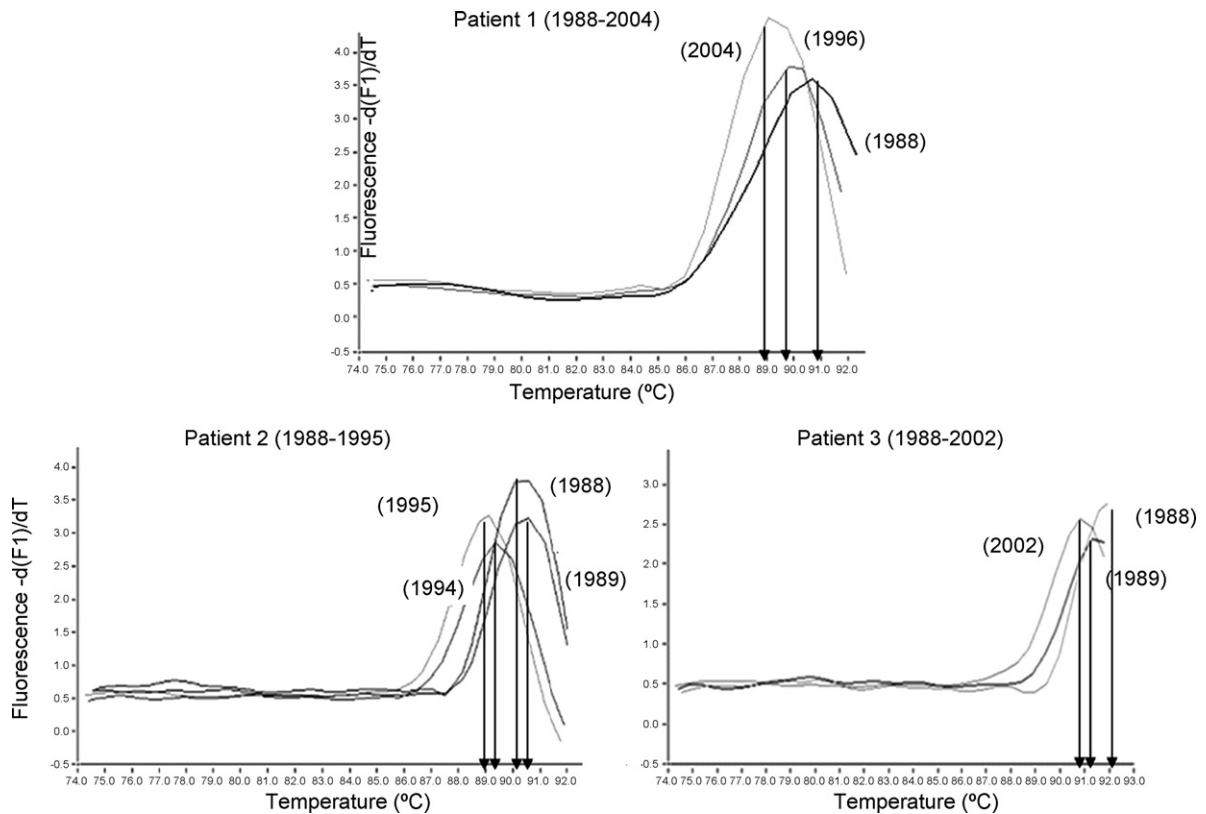


Fig. 3. Melting curves for HVR1 PCR products from Patients 1–3 at different time-points of infection. EPLD-PCR was used to amplify PCR product from isolates of HCV, which were subsequently subjected to melting-curve analysis. The negative differential of the relative fluorescence divided by temperature ($-d(RFU)/dT$) was plotted against temperature. Time-dependent melting-curve traces are shown. Vertical arrows denote the precise melting temperature.

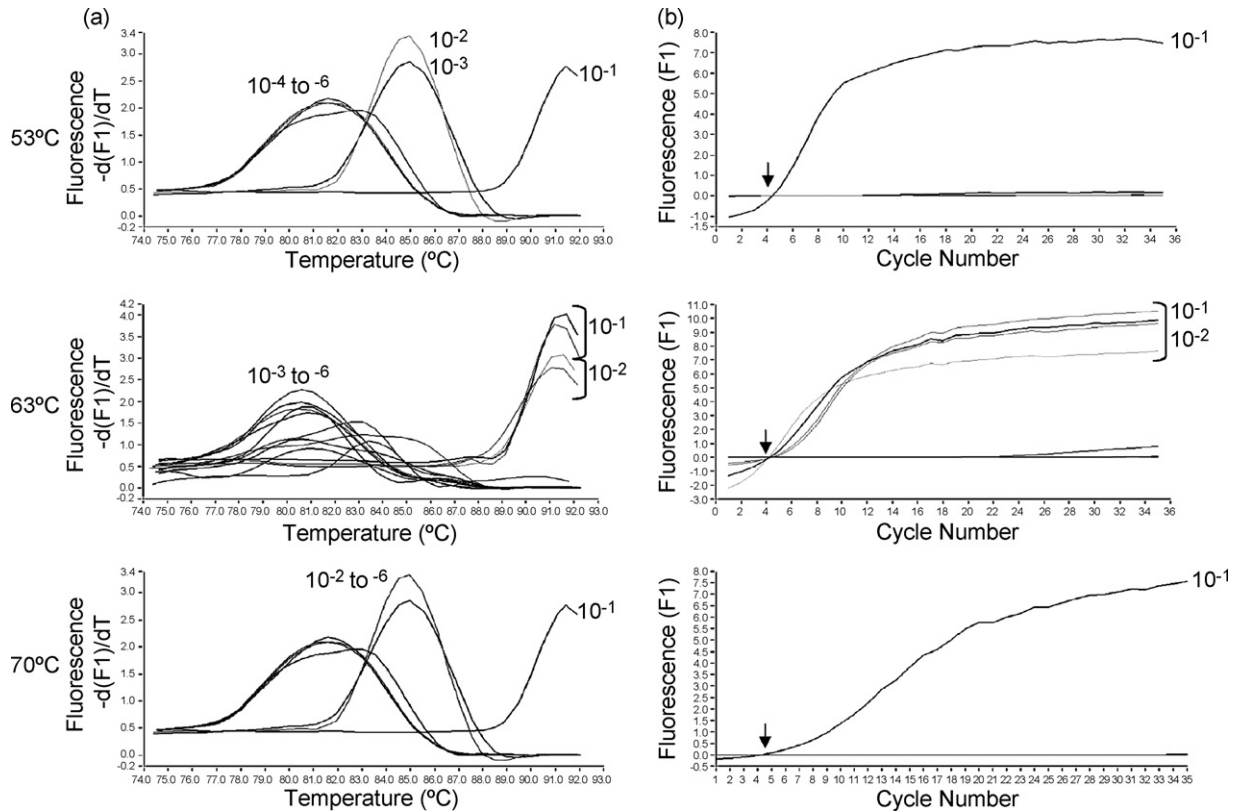


Fig. 4. Melting curves (a) and quantification curves (b) for HVR1 quasispecies analyzed from Patient 1 at optimal (63 °C) and suboptimal temperatures (53 °C and 70 °C). Fluorescence data were converted into melting peaks by the LightCycler software (Ver. 3.39) to plot the negative derivative of fluorescence over temperature vs. temperature ($-dF/dT$ vs. T). Fluorescence of SYBR Green I was measured once per cycle to monitor template amplification and represented as quantification curves. Arrows indicate amplification take-off-cycles. Duplicate experiments were performed at 63 °C to confirm the observation.

into these clusters. Only about 15–41% of sequence variants were shared between any two experiments. While PCR at 63 °C or 70 °C generated about 50–60% of unique sequences in each experiment, more than 70% of unique variants were generated at 53 °C. S_n values of sequences from PCR at 53 °C and 63 °C varied between 0.020 and 0.022, and a S_n value of 0.016 was obtained for PCR at 70 °C, but differences in these values were not statistically significant.

4. Discussion

This study reports the development and validation of an efficient and easy-to-perform real-time PCR to assess the HCV quasispecies population in HCV-infected patients. The usability of this method was shown using specimens obtained from an outbreak investigation, as well as serial serum specimens obtained from three chronically HCV-infected patients. To investigate HCV sequence diversity and complexity in these study subjects, PCR amplification and sequencing of viral genes are essential steps. However, a limitation of PCR-based methods for detection of viral variants is that due to stochastic variation in the early cycles of amplification, during direct amplification from clinical specimens one target DNA template may be preferentially amplified when multiple templates are present in the study sample (Ngui and Teo, 1997). This process is known as PCR drift (Wagner et al., 1994). In viral studies, such preferential amplification presents a problem to the making of inferences about viral sequence diversity and complexity (Meyerhans et al., 1990; Simmonds et al., 1990). Sm-PCR, however, is a simple alternative that circumvents this disadvantage inherent to the post-PCR cloning approach (Kraytsberg and Khrapko,

2005; Yauk et al., 2003). The sm-PCR strategy has been adopted for the study of mixed populations of variants in HIV (Blaak, 2005), recombination studies (Kraytsberg et al., 2004), mutational analysis of microsatellites (Jeffreys et al., 1990) and digital genotyping and haplotyping with polymerase colonies (Mitra et al., 2003).

In the current study, the concept of how single molecules are made available for sm-PCR was extended. An important consideration is that if PCR reactions were to generate sequences that are truly representative of the single-molecule state, they would be highly dependent on the conditions under which PCR was conducted. Deviations from optimal conditions should decrease PCR sensitivity and thus may require more than a single molecule for the amplification of a PCR product. Therefore, PCR conducted under suboptimal conditions may result potentially in amplification of more than one template at the end-point dilution. The outcome of the experiments presented in the current study challenges this notion.

It was determined that EPLD-PCR conducted under suboptimal annealing temperatures (53 °C and 70 °C) could indeed successfully generate HCV HVR1 fragments bearing unique sequences, albeit with a 10-fold lower sensitivity. Another observation was that a similar population of quasispecies could be sampled at optimal as well as suboptimal conditions. Some HVR1 sequences generated at the optimal and suboptimal temperatures were shared (Fig. 5). Such an outcome is unexpected because PCR at suboptimal temperatures should not, following the rationale behind sm-PCR as proposed by Kraytsberg and Khrapko (2005), amplify unique sequences from a mixed population of target templates. Our observation suggests that the function of end-point limiting-dilution in the EPLD-PCR is

to limit the number of molecules to a small critical number, from which one molecule is then stochastically amplified. The size of this pool depends on the sensitivity of PCR. Thus, it is not necessary, although desirable, to dilute a specimen to attain a single DNA molecule concentration in order to successfully sample a population of sequence variants. Since single molecules may be amplified from a pool of DNA, it leaves, however, a potential for recombination between different molecules (Kraytsberg and Khrapko, 2005). While recombination is likely still to occur at a low rate or between progeny of a single DNA molecule, use of EPLD-PCR in evaluating recombination and mutation events should continue to be assessed cautiously. However, when validated and applied for characterization and evolution of viral quasispecies involving short sequences, the likelihood of recombinant artifacts confounding the assessment is negligible.

Although about the same number of end-point limiting-dilution aliquots was sampled at each optimal and suboptimal condition, the number of quasispecies variants obtained varied, with 20 variants

being found at 70 °C and up to 41 at 53 °C, and 27 and 35 variants in the duplicate experiments at 63 °C. It seems that EPLD-PCR conducted at the lower than optimal annealing temperature samples larger number of different HCV variants, whereas PCR at optimal annealing temperature is more sensitive and can be performed on low-viral-titer specimens. Such slight differences in sampling quasispecies at different PCR conditions should be taken into consideration for the comparative analysis of quasispecies obtained using different PCR methods. Nonetheless, all subpopulations sampled at different PCR conditions are representative of the entire population of quasispecies. Taking into account these considerations, quasispecies amplification at different conditions is suitable for analyzing HVR1 heterogeneity and relatedness in clinical and outbreak investigation settings.

As the shape and position of the DNA melting curve are functions of the GC/AT ratio, GC distribution and base mismatches among other factors, it can be used to differentiate amplification products separated by less than 2 °C in melting temperatures (Ririe et

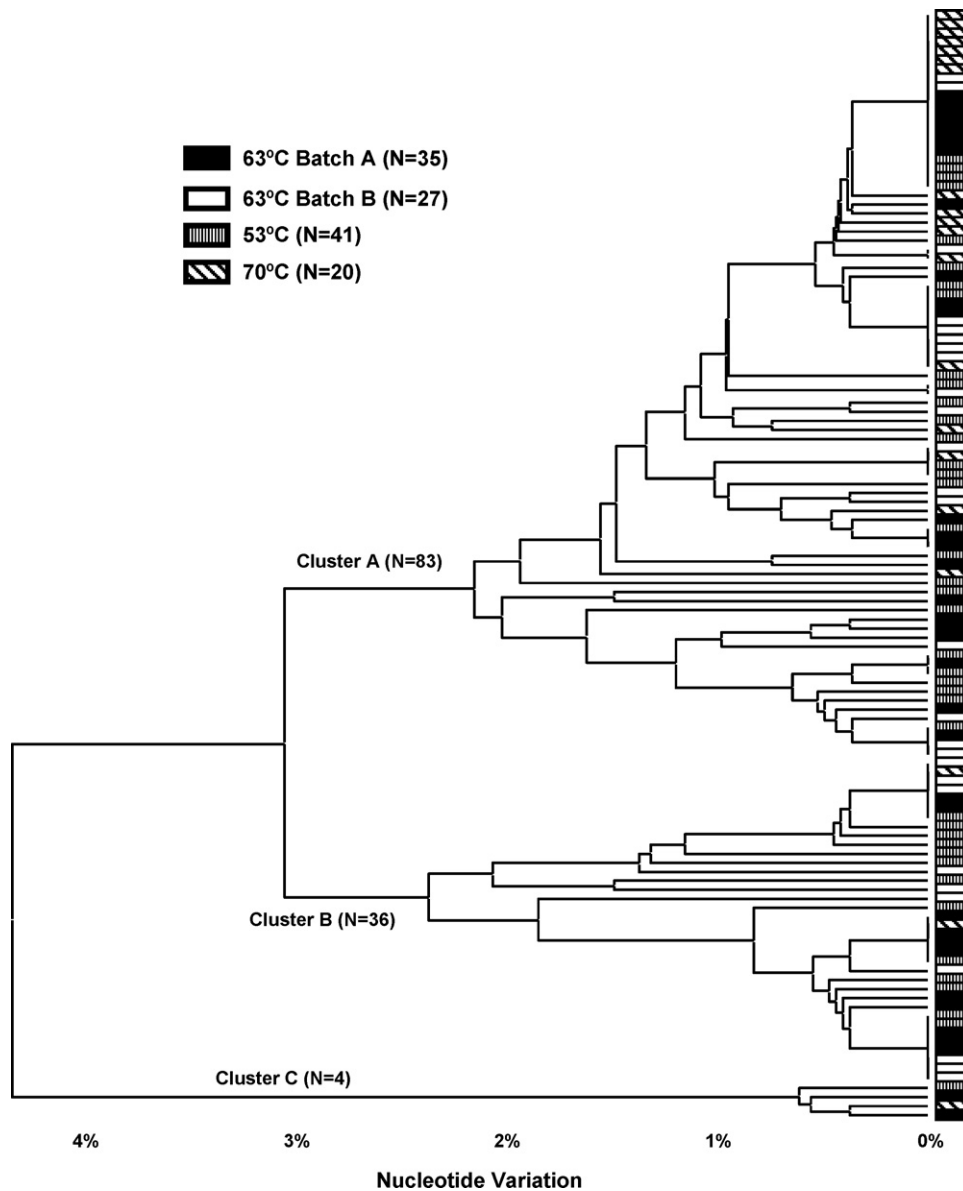


Fig. 5. Effect of different annealing temperatures on HVR1 quasispecies distribution of Patient 1 represented as an unrooted-dendrogram. A total of 123 HVR1 sequences were obtained from four experimental runs. Phylogenetic analysis of the HVR1 sequences obtained by EPLD-PCR at the three temperatures revealed existence of three clusters of quasispecies.

al., 1997). We used this phenomenon to our advantage to reduce the need for time-consuming gel electrophoresis and post-PCR manipulations to confirm positive amplicons. Moreover, the fine resolution of the melting peaks paves the way to gain an insight into the intra-individual heterogeneity of quasispecies in HCV-infected individuals prior to sequence analysis. The analysis of melting peak separation shown in Fig. 3 allows for a rapid identification of three major clusters of HVR1 quasispecies in HCV-infected patients studied in this work.

The present study also revealed that intra-person divergence may be so extensive as to exceed the divergence between unrelated HCV-infected patients. Previous studies have described a rapid sequence drift of HCV over time, a process of diversification that leads ultimately to the existence of identifiably separate strains or isolates within human populations (Okamoto et al., 1992; Smith et al., 1997). Measurements of cross-sectional population diversity and intra-sample substitution rates provide a snapshot of the complexity of a virus population, and indicates the strength of selective pressure and subsequent response to this made by the virus population. The dynamics of viral quasispecies mandate careful evaluation, and molecular epidemiological studies should take into consideration the rapid evolution of HCV HVR1 sequences, which potentially confounds genetic relatedness analysis.

References

- Alter, M.J., 1997. Epidemiology of hepatitis C. *Hepatology* 26, 62S–65S.
- Alter, M.J., Kruszon-Moran, D., Nainan, O.V., McQuillan, G.M., Gao, F., Moyer, L.A., Kaslow, R.A., Margolis, H.S., 1999. The prevalence of hepatitis C virus infection in the United States, 1988 through 1994. *N. Engl. J. Med.* 341, 556–562.
- Armstrong, G.L., Wasley, A., Simard, E.P., McQuillan, G.M., Kuhnert, W.L., Alter, M.J., 2006. The prevalence of hepatitis C virus infection in the United States: 1999 through 2002. *Ann. Intern. Med.* 144, 705–714.
- Ayers, M., Siu, K., Roberts, E., Garvin, A.M., Tellier, R., 2002. Characterization of hepatitis C virus quasispecies by matrix-assisted laser desorption/ionization-time-of-flight (mass spectrometry) mutation detection. *J. Clin. Microbiol.* 40, 3455–3462.
- Blaak, H., 2005. Isolation of human immunodeficiency virus type 2 biological clones from peripheral blood lymphocytes. *Methods Mol. Biol.* 304, 95–112.
- Curran, R., Jameson, C.L., Craggs, J.K., Grabowska, A.M., Thomson, B.J., Robins, A., Irving, W.L., Ball, J.K., 2002. Evolutionary trends of the first hypervariable region of the hepatitis C virus E2 protein in individuals with differing liver disease severity. *J. Gen. Virol.* 83, 11–23.
- Greenier, J.L., Van Rompay, K.K., Montefiori, D., Earl, P., Moss, B., Marthas, M.L., 2005. Simian immunodeficiency virus (SIV) envelope quasispecies transmission and evolution in infant rhesus macaques after oral challenge with uncloned SIVmac251: increased diversity is associated with neutralizing antibodies and improved survival in previously immunized animals. *Virology* 41, 2–11.
- Jeffreys, A.J., Neumann, R., Wilson, V., 1990. Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell* 60, 473–485.
- Jeffreys, A.J., Wilson, V., Neumann, R., Keyte, J., 1998. Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. *Nucleic Acids Res.* 16, 10953–10971.
- Kono, Y., Hayashida, K., Tanaka, H., Ishibashi, H., Harada, M., 2003. High-density lipoprotein binding rate differs greatly between genotypes 1b and 2a/2b of hepatitis C virus. *J. Med. Virol.* 70, 42–48.
- Kraytsov, Y., Schwartz, M., Brown, T.A., Ebralidze, K., Kunz, W.S., Clayton, D.A., Vissing, J., Khrapko, K., 2004. Recombination of human mitochondrial DNA. *Science* 304, 981.
- Kraytsov, Y., Khrapko, K., 2005. Single-molecule PCR: an artifact-free PCR approach for the analysis of somatic mutations. *Expert Rev. Mol. Diagn.* 5, 809–815.
- Lohmann, V., Roos, A., Korner, F., Koch, J.O., Bartenschlager, R., 2000. Biochemical and structural analysis of the NS5B RNA-dependent RNA polymerase of the hepatitis C virus. *J. Viral Hepat.* 7, 167–174.
- Lukyanov, K.A., Matz, M.V., Bogdanova, E.A., Gurskaya, N.G., Lukyanov, S.A., 1996. Molecule by molecule PCR amplification of complex DNA mixtures for direct sequencing: an approach to in vitro cloning. *Nucleic Acids Res.* 24, 2194–2195.
- Martell, M., Esteban, J.L., Quer, J., Genesca, J., Weiner, A., Esteban, R., Guardia, J., Gomez, J., 1992. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *J. Virol.* 66, 3225–3229.
- Meyerhans, A., Vartanian, J.P., Wain-Hobson, S., 1990. DNA recombination during PCR. *Nucleic Acids Res.* 18, 1687–1691.
- Mitra, R.D., Butty, V.L., Shendure, J., Williams, B.R., Housman, D.E., Church, G.M., 2003. Digital genotyping and haplotyping with polymerase colonies. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5926–5931.
- Morishima, C., Polyak, S.J., Ray, R., Doherty, M.C., Di Bisceglie, A.M., Malet, P.F., Bonkovsky, H.L., Sullivan, D.G., Gretch, D.R., Rothman, A.L., Koziel, M.J., Lindsay, K.L., Hepatitis C Antiviral Long-Term Treatment Against Cirrhosis Trial Group, 2006. Hepatitis C virus-specific immune responses and quasi-species variability at baseline are associated with nonresponse to antiviral therapy during advanced hepatitis C. *J. Infect. Dis.* 193, 931–940.
- Nainan, O.V., Khristova, M.L., Byun, K., Xia, G., Taylor, P.E., Stevens, C.E., Margolis, H.S., 2002. Genetic variation of hepatitis B surface antigen coding region among infants with chronic hepatitis B virus infection. *J. Med. Virol.* 68, 319–327.
- Nainan, O.V., Lu, L., Gao, F.X., Meeks, E., Robertson, B.H., Margolis, H.S., 2006. Selective transmission of hepatitis C virus genotypes and quasispecies in humans and experimentally infected chimpanzees. *J. Gen. Virol.* 87, 83–91.
- Nakano, I., Maertens, G., Major, M.E., Vitvitski, L., Dubuisson, J., Fournillier, A., De Martynoff, G., Trepo, C., Inchauspe, G., 1997. Immunization with plasmid DNA encoding hepatitis C virus envelope E2 antigenic domains induces antibodies whose immune reactivity is linked to the injection mode. *J. Virol.* 71, 7101–7109.
- Ngui, S.L., Teo, C.G., 1997. Hepatitis B virus genomic heterogeneity: variation between quasispecies may confound molecular epidemiological analyses of transmission incidents. *J. Viral Hepat.* 4, 309–315.
- Okamoto, H., Kojima, M., Okada, S., Yoshizawa, H., Iizuka, H., Tanaka, T., Muchmore, E.E., Peterson, D.A., Ito, Y., Mishiro, S., 1992. Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability. *Virology* 190, 894–899.
- Ramachandran, S., Xia, G., Nainan, O.V., Khudyakov, Y., 2006. Temporal patterns of quasispecies in chronic HCV infected patients. In: Proceedings of the 12th International Symposium on Viral Hepatitis and Liver Disease, 2006, Conference Program and Book of Abstracts, Paris, France.
- Ririe, K.M., Rasmussen, R.P., Wittwer, C.T., 1997. Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Anal. Biochem.* 245, 154–160.
- Rossi, L., Leveri, M., Marinelli, D., Belli, L., Cividari, E., Silini, E.M., 2003. A heteroduplex mobility analysis assay based on capillary electrophoresis for the study of HCV quasispecies. *J. Virol. Methods* 110, 37–49.
- Sakamoto, M., Akahane, Y., Tsuda, F., Tanaka, T., Woodfield, D.G., Okamoto, H., 1994. Entire nucleotide sequence and characterization of a hepatitis C virus of genotype V/3a. *J. Gen. Virol.* 75, 1761–1768.
- Schuitmaker, H., Kootstra, N.A., 2005. Isolation, propagation, and titration of human immunodeficiency virus type 1 from peripheral blood of infected individuals. *Methods Mol. Biol.* 304, 17–24.
- Simmonds, P., Zhang, L.Q., Watson, H.G., Rebus, S., Ferguson, E.D., Balfe, P., Leadbetter, G.H., Yap, P.L., Peutherer, J.F., Ludlam, C.A., 1990. Hepatitis C quantification and sequencing in blood products, haemophiliacs, and drug users. *Lancet* 336, 1469–1472.
- Smith, D.B., Pathirana, S., Davidson, F., Lawlor, E., Power, J., Yap, P.L., Simmonds, P., 1997. The origin of hepatitis C virus genotypes. *J. Gen. Virol.* 78, 321–328.
- Stead, J.D., Jeffreys, A.J., 2000. Allele diversity and germline mutation at the insulin minisatellite. *Hum. Mol. Genet.* 9, 713–723.
- Stein, C.A., Levantis, P., Oxford, J.S., 1994. Sequence analysis of proviral HIV RT amplified directly by a semi-quantitative technique from AZT treated patients. *J. Med. Virol.* 44, 115–121.
- Tamura, K., Nei, M., Kumar, S., 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11030–11035.
- Vizmanos, J.L., Gonzalez-Navarro, C.J., Novo, F.J., Civeira, M.P., Prieto, J., Gullon, A., Garcia-Delgado, M., 1998. Degree and distribution of variability in the 5' untranslated, E1, E2/NS1 and NS5 regions of the hepatitis C virus (HCV). *J. Viral Hepat.* 5, 227–240.
- Vogelstein, B., Kinzler, K.W., 1999. Digital PCR. *Proc. Natl. Acad. Sci. U.S.A.* 96, 9236–9241.
- Wagner, A., Blackstone, N., Cartwright, P., 1994. Surveys of gene families using polymerase chain reactions: PCR sequences and PCR drift. *Syst. Biol.* 43, 250–261.
- Wilson, J.J., Polyak, S.J., Day, T.D., Gretch, D.R., 1995. Characterization of simple and complex hepatitis C virus quasispecies by heteroduplex gel shift analysis: correlation with nucleotide sequencing. *J. Gen. Virol.* 76, 1763–1771.
- Xia, G., Gao, F., Chai, F., Nainan, O.V., Alter, M.J., Margolis, H.S., 2006. Genetic relatedness between hepatitis C virus variants: lessons from a general population survey. In: Proceedings of the 13th International Meeting on Hepatitis C Virus & Related Viruses, 2006, Conference Program & Book of Abstracts, Cairns, Australia, p. 64.
- Yauk, C.L., Bois, P.R., Jeffreys, A.J., 2003. High-resolution sperm typing of meiotic recombination in the mouse MHC Ebeta gene. *EMBO J.* 22, 1389–1397.
- Zampino, R., Marrone, A., Mangoni, E.D., Santarpia, L., Sica, A., Tripodi, M.F., Utili, R., Ruggiero, G., Adinolfi, L.E., 2004. Anti-envelope 1 and 2 immune response in chronic hepatitis C patients: effects of hepatitis B virus co-infection and interferon treatment. *J. Med. Virol.* 73, 33–37.