ORIGINAL ARTICLE

Hepatitis C virus quasispecies in chronically infected children subjected to interferon–ribavirin therapy

Magdalena Figlerowicz · Paulina Jackowiak · Piotr Formanowicz ·
Paweł Kędziora · Magdalena Alejska · Nelli Malinowska · Jacek Błażewicz ·
Marek Figlerowicz

Abstract Accumulating evidence suggests that certain features of hepatitis C virus (HCV), especially its high genetic variability, might be responsible for the low efficiency of anti-HCV treatment. Here, we present a bioinformatic analysis of HCV-1a populations isolated from 23 children with chronic hepatitis C (CHC) subjected to interferon–ribavirin therapy. The structures of the viral quasispecies were established based on a 132-amino-acid sequence derived from E1/E2 protein, including hypervariable region 1 (HVR1). Two types of HCV populations were identified. The first type, found in non-responders, contained a small number of closely related variants. The second type, characteristic for sustained responders, was composed of a large number of distantly associated equal-rank variants. Comparison of 445 HVR1 sequences showed that a significant number of variants present in non-responding patients are closely related, suggesting that certain, still unidentified properties of the pathogen may be key factors determining the result of CHC treatment.

M. Figlerowicz and P. Jackowiak equally contributed to these studies.

P. Jackowiak · P. Formanowicz · M. Alejska ·
N. Malinowska · J. Błażewicz · M. Figlerowicz (✉)
Institute of Bioorganic Chemistry Polish Academy of Sciences,
Noskowskiego 12/14, 61-704 Poznań, Poland
e-mail: marekf@ibch.poznan.pl

M. Figlerowicz
Department of Infectious Diseases and Child Neurology,
Karol Marcinkowski University of Medical Sciences,
Szpitalna 27/33, 60-572 Poznań, Poland

P. Formanowicz · P. Kędziora · J. Błażewicz
Institute of Computing Science,
Poznań University of Technology,
Piotrowo 3A, 60-965 Poznań, Poland

Introduction

At present, about 3% of the human population is infected with hepatitis C virus (HCV). The first, acute stage of HCV infection is usually asymptomatic [11]. However, only 15–30% of the infected eliminate the virus. The remaining patients develop chronic hepatitis C (CHC). Despite two decades of intensive studies, no obvious correlations have been found between the parameters describing CHC patients and the course of infection or the outcome of therapy [1]. Instead, HCV genotype and virus load were confirmed to be important factors affecting the final result of antiviral treatment. Individuals infected with genotype 2 or 3 and those with a low level of viremia respond better to the therapy [14, 18, 24]. This indicates that certain features of the virus, rather than those of the host, may actually be critical for CHC treatment.

Recently, it has become increasingly clear that one of the key factors making CHC therapy so difficult is the unusual genetic polymorphism of the virus [8, 27]. The major cause of this polymorphism is the effective, but error-prone replication of the RNA genome [8]. Consequently, HCV does not form a homogenous population but exists as a quasispecies [8, 27]. The latter is defined as a pool of phylogenetically related but genetically slightly distinct variants present in a single infected organism [3, 4]. All variants are subjected to continuous selection; therefore, only the most fit are able to survive and spread [8]. Several regions of the HCV genome demonstrate an especially high level of variability. Among them, hypervariable region 1 (HVR1),

encoding a 27-aa N-terminal fragment of the E2 glycoprotein, is thought to be the most heterogeneous one.

A number of attempts have been made in recent years to find a correlation between the complexity of HCV quasispecies and the outcome of CHC therapy in adults. The diversity of viral populations was usually evaluated based on two regions of the HCV genome: E1/E2 (including HVR1), and NS5A. As a result, it has been postulated that HCV quasispecies display decreased heterogeneity in patients who develop a sustained virological response [5, 20, 25, 30]. In contrast, little is currently known about the structure of the HCV population in children. This issue raises many important questions, especially since the natural history and clinical features of HCV infection in children and adults are different. A milder, often asymptomatic liver disease with a normal or slightly elevated level of alanine aminotransferase activity impedes the diagnosis of HCV infection in childhood [15]. Consequently, the percentage of CHC in pediatric patients is probably underestimated. In addition, available reports concerning HCV quasispecies in children are usually based on only a few cases. Some of these reports refer to viral populations in untreated children [6, 12], the others to HCV quasispecies in patients coinfected with human immunodeficiency virus (HIV) [2, 26].

In this paper, we present an analysis of HCV quasispecies isolated from 23 children with CHC, all infected with the same HCV subtype (1a). For each of them, the structure of the virus population was determined just before anti-HCV treatment and 2 weeks later.

Materials and methods

Patient qualification and treatment

The studies included two groups of children with chronic hepatitis C (all infected with HCV-1a). The first group included 10 patients subjected to combined interferon–ribavirin therapy in 2004 (patients designated P1-01 to P1-10). The second group comprised 13 patients who were treated in the same way, but 1 year later (patients P2-02 to P2-28). The study conforms to the ethical guidelines of the World Medical Association Declaration of Helsinki. In each case, both parents and children (12 years or older) provided informed consent for treatment and participation in the study. Before the commencement of the project, an appropriate approval from the Bioethical Commission of Karol Marcinkowski University of Medical Sciences was obtained. The inclusion criteria were as follows: (1) the detection of anti-HCV antibodies and HCV RNA in the serum for at least 6 months; (2) periodically or continuously elevated ALT activity; and (3) histopathological signs of chronic hepatitis [9]. The treatment was conducted according to the following schema: interferon alpha 2b was administered in 3 million unit doses subcutaneously 3 times a week; ribavirin, in doses of 15 mg/kg per day orally [9]. The level of HCV RNA accumulation was determined for each patient, just before therapy (T0) and then after 24 (T24), 48 (T48) and 72 (T72) weeks, using an Amplicor HCV Monitor Test v. 2.0 (Roche). The patients with a positive result (a detectable amount of HCV RNA) at T24 were excluded, whereas the others were treated for 24 more weeks. After 72 weeks of treatment and follow-up, patient response to combined interferon–ribavirin therapy was evaluated. It was considered: (1) sustained (SR) if HCV RNA was not detected at T24, T48 and T72; (2) transient (TR) if HCV RNA was not detected at T24 but appeared again at T48 and was still present at T72 or HCV RNA was not detected at T24 and T48 and appeared at T72; (3) no response (NR) if HCV RNA was detected at T24, T48 and T72. Anti-HCV antibodies were detected in sera by fourth-generation ELISA (ETI-AB-HCVK-4, DiaSorin).

Analysis of the HCV quasispecies structure

The structure of the HCV population was examined separately for each of 23 patients before therapy (T0) and 2 weeks after therapy had begun (T2). To this end, HCV RNA was isolated from the patients' sera, and a 462-nt portion of the E1/E2 coding sequence (a fragment located in the HCV genome between positions 1425 and 1886) was amplified by the RT-PCR method [19] using a OneStep RT-PCR Kit (QIAGEN). All nucleotide positions refer to the standard HCV-1a genome sequence, accession number AF009606.

The following pairs of primers were used for RT-PCR:

(1) Outer: F1: 5′CAG(CT)T(AG)CTCCGGATCCCACAAGC3′ and R1: 5′ACGTCCGTCTCATT(CT)(TG)C(AC)CCCCA3′;
(2) Inner: F2: 5′TCTGGATCCTATTCCATGGTGGGGAACTGG3′ and R2: 5′AATGAATTCTACAACAGGGCT(TG)GG(AG)GTGAA3′ (modified from ref. 19). Letters in parentheses indicate degenerate positions.

The RT-PCR products were cloned into the pUC vector, and at least 20 randomly selected clones were sequenced per population. The nucleotide sequences obtained were translated *in silico* into amino acid sequences. A bioinformatic analysis was conducted separately for each population (represented by 20 randomly selected clones at each time point; over 900 sequences in total). This was done using the predicted amino acid sequence of the entire amplicon (excluding primer hybridization sites) or its fragments: hypervariable region 1 (HVR1, located in the

HCV genome between positions 1491 and 1571) or the remaining portion of the amplified E1/E2 coding sequence (non-HVR1). The composition of the HCV quasispecies was established by direct comparison of the predicted amino acid sequences. As a result, the number of different variants within the population and the contribution of each individual variant to the examined population were established. Clones containing nonsense and frameshift mutations were excluded. Consequently, the number of variants analyzed was reduced for some populations. Genetic diversity was assessed based on the Hamming distance, which is defined as the number of amino acid differences between two sequences. The genetic diversity of a population was expressed as a mean Hamming distance (MHD), an average value taken for all sequence pairs. Statistical analysis was performed using the Statistica 8.0 package. The values of MHD determined for all patients were divided into two groups, (1) non-responders + transient responders and (2) sustained responders, and analyzed using the non-parametric U Mann–Whitney test (because the results of the Shapiro–Wilk test were negative). Phylogenetic trees were created using the maximum-likelihood method with 500 bootstrap replicates. WAG and HKY85 substitution models were employed for amino acid and nucleotide sequences, respectively. Selection tests were made using the internal fixed effects likelihood (IFEL) and single likelihood ancestor counting (SLAC) approaches, available at www.datamonkey.org [23]. To count the global (SLAC) and codon-specific (IFEL) dN/dS ratios, the HKY85 model of nucleotide substitution and neighbor-joining trees [17, 28] were utilized, with a significance level value of 0.1. All available sequences of viral variants were analyzed. A consensus amino acid sequence was produced using the Consensus tool available at the Los Alamos HCV Sequence Database http://hcv.lanl.gov/content/hcv-db/consensus/consensus.html [16].

Results

Complexity of HCV quasispecies

The complexity of HCV quasispecies was assessed twice for all 23 patients: just before treatment (T0) and 2 weeks later (T2). The number of different variants within each population and each population's genetic diversity were determined. The analysis was carried out for the entire amplified E1/E2 region, and then separately for its two fragments: HVR1 and the remaining portion of E1/E2 (non-HVR1). Finally, three sets of data describing the complexity of HCV quasispecies (Online Resources 1, 2, 3, 4) were examined in relation to parameters characterizing the patients' response to anti-HCV therapy, the level

of HCV RNA accumulation in their blood and the level of ALT activity (Table 1). As a result, a significant correlation was only found between the outcome of antiviral therapy and HCV quasispecies complexity if the latter was determined at T0 and based solely on HVR1. In this case, HCV populations isolated from non-responders and transient responders were composed of a markedly lower number of variants (mean: 6,462; SD: 3,406) than populations derived from sustained responders (mean: 11; SD: 3,496; Fig. 1a; Online Resource 3). The level of genetic diversity (expressed by the mean Hamming distance: MHD) differed significantly ($p < 0.05$) between sustained responders and children who failed to clear the virus (non-responders plus transient responders). The obtained p values were equal to 0.0011 and 0.0039 for T0 and T2, respectively. The MHD was lower for HCV populations derived from non-responders and transient responders (mean at T0: 1,899; SD: 2,043) and higher for quasispecies coming from sustained responders (mean at T0: 6,152; SD: 2,486; Fig. 1b). Only for one population isolated from a child classified as a non-responder did MHD reach a relatively high value of 4.93. For two quasispecies derived from sustained responders, the MHD value was considerably below average: 0.79 and 3.14. The analysis performed for the whole amplified E1/E2 region revealed tendencies similar to those observed for HVR1. However, they were much less pronounced (Online Resource 4).

The 2-week interferon–ribavirin therapy differentially influenced the composition of the viral populations. At T2, an increase in the number of variants was observed in the majority of populations present in non-responders. A decrease was observed in those from sustained responders. A closer analysis of individual variants revealed diverse patterns of quasispecies evolution over the 2 weeks of treatment. Only in some populations was the contribution of the dominating variant increased. For others, we observed one of the following: (1) a lack of any substantial changes in the population structure; (2) the formation of new co-dominants; (3) the extinction of the dominant variant or (4) the development of new dominating variants that had not been identified before treatment (Online Resource 3). The influence of the applied therapy on MHD was quite diverse. It was not related to the patients' response to antiviral treatment.

Phylogenetic analysis

To determine not only the diversity of variants constituting each individual HCV population but also the relationships between them, phylogenetic trees were constructed for all of the quasispecies (Fig. 2; Online Resources 5, 6, 7). Again, the most significant correlation between the outcome of antiviral therapy and the character of HCV phylogeny was found for trees produced on the basis of HVR1

Table 1 Clinical, biochemical and virological characteristics of children with chronic hepatitis C subjected to interferon–ribavirin therapy

Patient no.	Age (years)	Sex	Pretreatment viremia (copies/mL)	Pretreatment ALT (IU/L)	Pretreatment anti-HCV antibodies (fourth-generation ELISA)	Treatment response
P1-01	8	F[a]	7,803,000	44	+	NR[c]
P1-02	8	M[b]	14,148,000	35	+	NR
P1-03	12	F	804,600	59	+	NR
P1-04	17	M	9,072,000	19	+	TR[d]
P1-05	14	M	9,018,000	29	+	NR
P1-06	11	F	337,500	32	+	SR[e]
P1-07	9	M	631,800	87	+	TR
P1-08	10	F	858,600	30	+	SR
P1-09	10	M	804,600	20	+	SR
P1-10	10	F	46,440	48	+	SR
P2-02	12	F	768,600	17	+	SR
P2-04	13	M	11,280,000	315	+	NR
P2-05*	11	F	2,170,000	53	−	NR
P2-08	16	F	1,870,000	40	+	SR
P2-10*	11	M	2,670,000	45	−	NR
P2-17	8	M	56,300,000	19	+	SR
P2-18	15	F	696,000	54	+	SR
P2-19	10	F	726,000	35	+	NR
P2-20	15	F	152,000	37	+	SR
P2-22	11	M	713,000	35	+	SR
P2-23	10	M	840,000	63	+	NR
P2-24	16	F	286,000	28	+	TR
P2-28*	14	F	365,000	51	−	NR

[a] Female, [b] Male, [c] No response, [d] Transient response, [e] Sustained response, *Patient previously subjected to immunosuppressive treatment of an oncologic disease, + Anti-HCV positive, − Anti-HCV negative

identified at T0. Trees depicting populations derived from non-responding children were compact, almost monophyletic. Interestingly, the only quasispecies from this group for which a visibly higher MHD value was obtained (P2-23) was found to be composed of two intrinsically homogeneous subpopulations. In contrast, trees based on populations isolated from sustained responders were very extended, without any clear clusters.

After the 2 weeks of treatment (T2), the topology of the phylogenetic trees remained mostly unaffected. Some changes could be observed almost exclusively for populations isolated from sustained responders. In these cases, a reduction or an increase in the degree of tree branching was observed. An analogous analysis based on the entire E1/E2 region or its non-HVR1 fragment did not yield conclusive results (Online Resources 5 and 6). This again confirmed that HVR1 best reflects differences between quasispecies obtained from children with different responses to anti-HCV therapy.

Inter-quasispecies HVR1 conservation

To evaluate the relationships between HVR1 variants present in different quasispecies at T0, an integrated phylogenetic tree was constructed. It was obtained based on all (445) variants identified in 23 HCV populations. While variants from sustained responders were noticeably dispersed, those derived from patients responding transiently and from non-responders tended to concentrate in clusters (Online Resource 8). Astonishingly, the variants isolated from ten non-responding children clustered into only seven groups. A more detailed analysis revealed that one of them, cluster 1, encompassed variants from five patients. Together with the closely related cluster 2, they constituted a pair grouping variants derived from six epidemiologically distinct viral populations (P1-01, P1-02, P1-03, P1-05, P2-05 and P2-19) (Online Resource 8; Fig. 3a). This observation suggested that HVR1 variants formed in different non-responding patients were often identical or quite similar. All of the analyses presented above were based on comparisons of amino acid sequences. To verify the results concerning inter-quasispecies conservation of HVR1, nucleotide sequences of variants grouping in clusters 1 and 2 were compared. Both phylogenetic analysis (Fig. 3b) and sequence alignment revealed population-specific differences occurring at the nucleotide level. Consequently, consensus sequences unique to each quasispecies could be generated (Fig. 4). This provided
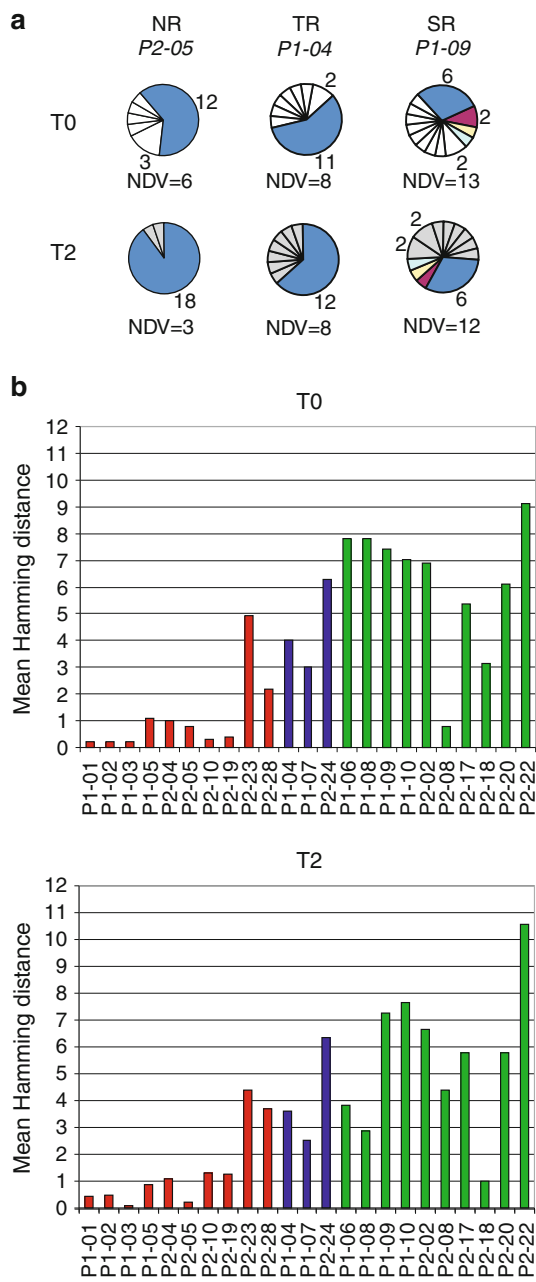
Fig. 1 Complexity of HCV quasispecies (determined based on the HVR1 amino acid sequence). **a** Number of different variants. Circular diagrams show the composition of the individual HCV populations derived from the representative: non-responder (NR), transient responder (TR) and sustained responder (SR). Each individual variant is represented by one sector. Sectors representing variants identified exclusively at T0 or T2 are *white* and *gray*, respectively. Sectors representing identical variants found at both T0 and T2 in one patient (not in different patients) are indicated with other *colors*. *Numbers* accompanying the diagrams indicate how frequently each variant was identified (the lack of a number indicates that the sequence was found only once). The number of different variants (NDV) found in each population is given below the diagram. **b** Genetic diversity. The mean Hamming distance was calculated separately for each viral population based on 20 HVR1 clones randomly selected at T0 and T2. The colors of the bars correspond to the three types of responses of the patients to the treatment: *red* no response; *blue* transient response; and *green* sustained response

additional evidence that the similarity/identity between HVR1 variants was not caused by cross-contamination. These results suggest that the similar/identical HVR1 variants that were identified were already optimal: their appearance was strictly connected with the lack of patients' response to the antiviral therapy. None of them were found in populations isolated from patients with transient and sustained response (Online Resource 8). These variants were present in viral populations before treatment and persisted at least through the first 2 weeks of its duration (Online Resource 3: identity of dominating variants between T0 and T2).

Selection pressure as a factor shaping the HCV population

A viral quasispecies is subjected to natural selection throughout its existence within the host. Positive selection acts towards introducing nonsynonymous mutations and increasing the diversity of the population. In contrast, negative selection favors synonymous changes and restricts the process of diversification, keeping the population more homogenous. In order to determine what kind of selection pressure shaped the viral population in the group of patients studied, the nucleotide sequences of all HVR1 variants identified at T0 and T2 were compared. The analysis included only the HVR1 region because it best reflected differences between populations derived from patients with different responses to antiviral therapy. Initially, the nucleotide sequences of 904 HVR1 variants derived from 23 patients were divided (according to the sampling time and type of response to the antiviral therapy) into six groups: NR-T0, NR-T2, TR-T0, TR-T2, SR-T0, SR-T2 (NR—no response, TR—transient response, SR—sustained response). First, the number of synonymous and nonsynonymous substitutions per site (dS and dN, respectively) was established. Then, the global and codon-specific dN/dS ratio (parameter characterizing selection pressure) was determined (Table 2). Values lower than 1 are characteristic of negative selection; values higher than 1, of positive selection. Under positive selection, only 2 and 4 codons were identified within the sequences of variants forming the NR-T0 and NR-T2 groups, respectively; under negative selection, 4 and 5 codons were identified in the same sequences. A similar trend was detected among HVR1 sequences from the TR-T0 and TR-T2 groups. In contrast, as many as 10 codons were subjected to positive selection, and only 2 were subjected to negative selection in variants belonging to the SR-T0 group. In SR-T2, 9 codons were found to be under positive selection and 4 codons were under negative selection. The global character of selection within the entire HVR1 was negative in non-responders and transient responders. It was
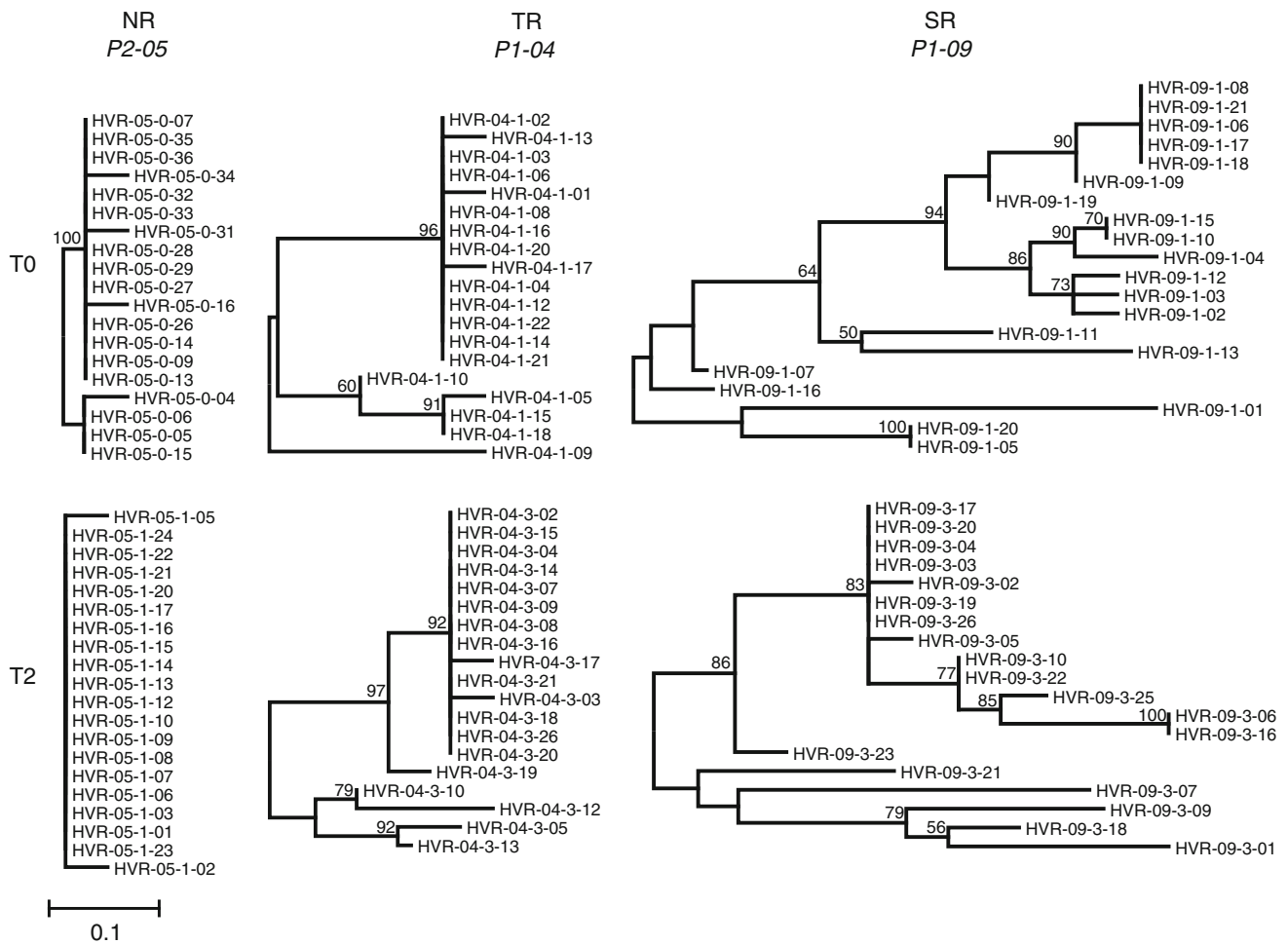
Fig. 2 Phylogenetic analysis of individual HCV quasispecies. Representative phylogenetic trees obtained for HCV populations isolated at T0 and T2 from non-responders (NR), transient responders (TR) and sustained responders (SR). The trees were constructed based on HVR1 amino acid sequences. The percentage of trees in which the variants clustered together is indicated above the branches (values lower than 50% were removed). Branch lengths are proportional to the number of substitutions per site

strongly positive in sustained responders. After 2 weeks of therapy, the number of selected sites increased in all three groups, but the character of the pressure operating globally in HVR1 remained unchanged. This observation clearly indicates that selection pressure is the key modulator of HCV quasispecies diversity. Strikingly, the commonly described process of continuous selection and adaptation was apparently ongoing only in sustained responders. In contrast, the evolution of the quasispecies in non-responding patients seemed to be rigorously confined, most probably by the necessity for maintaining defined HVR1 properties.
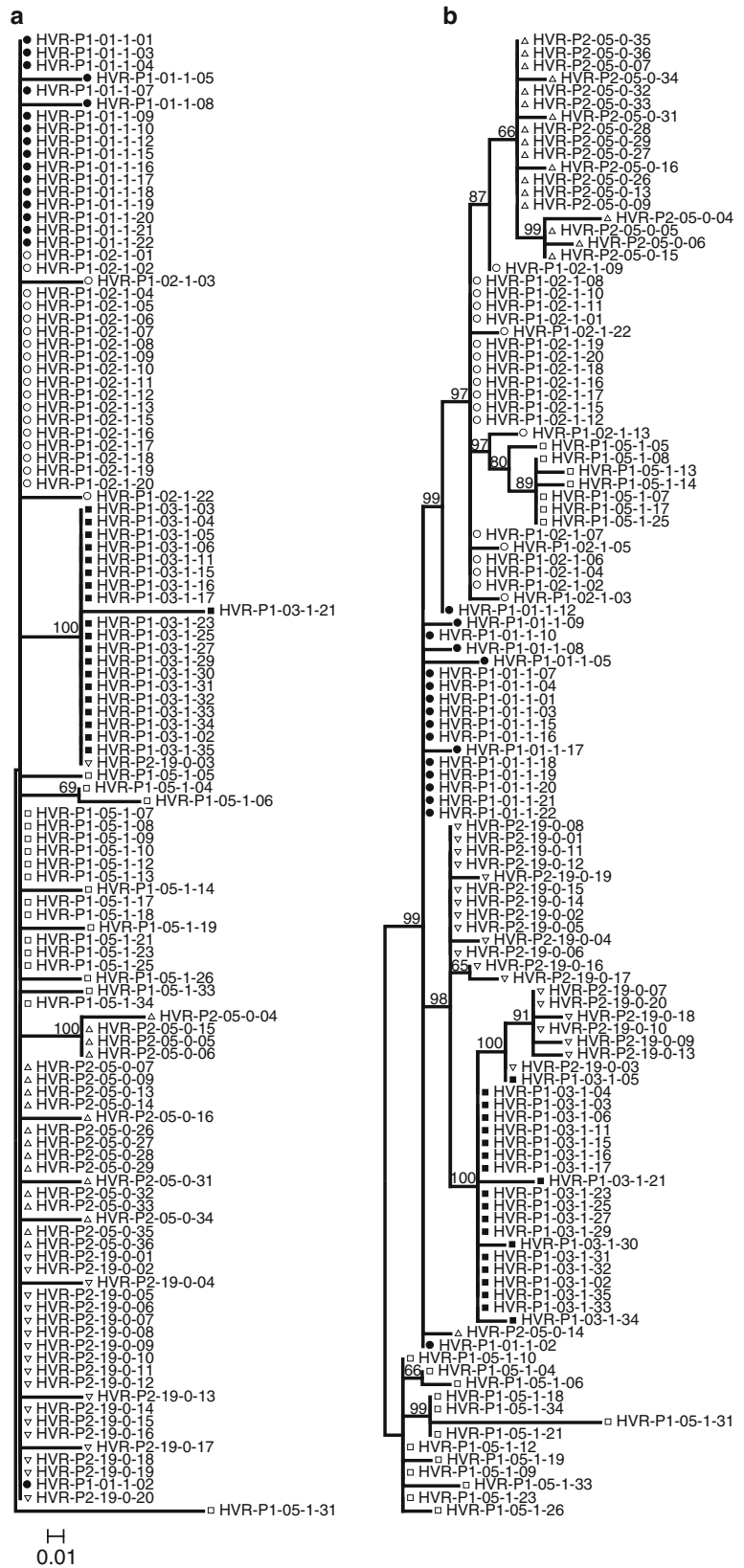
Discussion

In spite of ample evidence that high genetic variability of RNA viruses affects host-virus interactions, the question

concerning the correlation between particular parameters characterizing HCV quasispecies and the course of CHC or results of its treatment remains open. Considering earlier studies, we decided to focus our investigations on the 462-nt fragment of the E1/E2 region of the HCV genome, comprising HVR1. Previous analyses of the HCV quasispecies have usually involved several genotypes (often all six), although it was known that HCV-2 and -3 are eliminated more frequently than HCV-1 [7]. The latter observation suggest that the results obtained for one genotype cannot be extrapolated to the others, and that is why our analysis involved a group of 23 children with CHC, all of whom were infected with HCV-1a and qualified for combined interferon–ribavirin therapy.

The collected data revealed an apparent correlation between the final effect of CHC therapy and the HVR1 polymorphism observed just before treatment (at T0). Accordingly, we distinguished two types of HCV

Fig. 3 Inter-quasispecies conservation of HVR1 variants isolated from non-responders. Phylogenetic analysis was carried out with selected HCV quasispecies isolated from non-responders at T0. Phylogenetic trees were constructed for HVR1 amino acid (**a**) and nucleotide (**b**) sequences. The individual sequences are named according to the following schema: *HVR-patient-number-clone-number*; e.g. HVR-P1-03-1-21, where 'P1-03' indicates the patient number, while '1-21' refers to the clone number. In addition, the names are accompanied by geometrical symbols. All sequences representing the same quasispecies are marked with the same symbol. The percentage of trees in which the variants clustered together is indicated above the branches (values lower than 50% were removed). Branch lengths are proportional to the number of substitutions per site
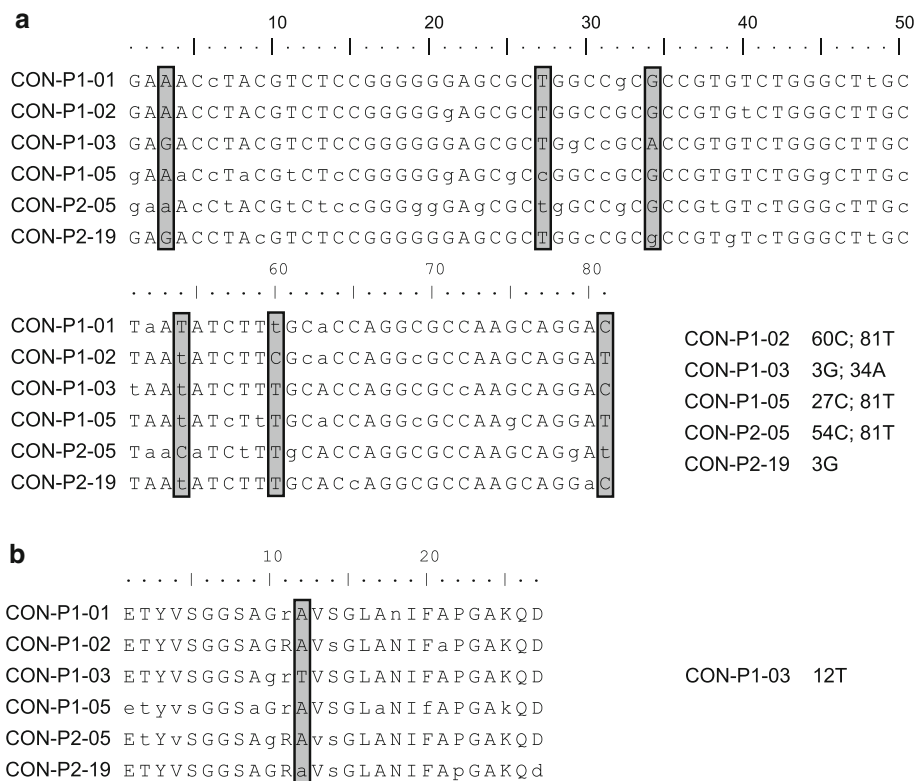
a

```
                    10        20        30        40        50
           ....|....|....|....|....|....|....|....|....|....|
CON-P1-01  GAAACcTACGTCTCCGGGGGGAGCGCTGGCCgCGCCGTGTCTGGGCTtGC
CON-P1-02  GAAACCTACGTCTCCGGGGGgAGCGCTGGCCGCGCCGTGtCTGGGCTTGC
CON-P1-03  GAGACCTACGTCTCCGGGGGGAGCGCTGGCcGCACCGTGTCTGGGCTTGC
CON-P1-05  gAAAccTaCGtCTcCGGGGGgAGCGcCGGCcGCGCCGTGTCTGGGCTTGc
CON-P2-05  gaaAcCtACGtCtccGGGggGAgCGCtgGCCgCGCCGtGTcTGGGCTTGc
CON-P2-19  GAGACCTAcGTCTCCGGGGGGAGCGCTGGccGCGCCGTgTcTGGGCTTGC
```

```
                    60        70        80
           ....|....|....|....|....|....|....|.
CON-P1-01  TaATATCTTtGCaCCAGGCGCCAAGCAGGAC
CON-P1-02  TAAtATCTTCGCaCCAGGcGCCAAGCAGGAT
CON-P1-03  tAAtATCTTTGCACCAGGCGCCaAGCAGGAC
CON-P1-05  TAAtATCTtTGCaCCAGGcGCCAAgCAGGAT
CON-P2-05  TaaCaTCtTTgCACCAGGCGCCAAGCAGgAt
CON-P2-19  TAAtATCTTTGCACcAGGCGCCAAGCAGGaC
```

CON-P1-02   60C; 81T
CON-P1-03   3G; 34A
CON-P1-05   27C; 81T
CON-P2-05   54C; 81T
CON-P2-19   3G

b

```
                    10        20
           ....|....|....|....|....|..
CON-P1-01  ETYVSGGSAGrAVSGLAnIFAPGAKQD
CON-P1-02  ETYVSGGSAGRAVsGLANIFaPGAKQD
CON-P1-03  ETYVSGGSAgrTVSGLANIFAPGAKQD
CON-P1-05  etyvsGGSaGrAVSGLaNIfAPGAKQD
CON-P2-05  EtYvSGGSAgRAvsGLANIFAPGAKQD
CON-P2-19  ETYVSGGSAGRaVsGLANIFApGAKQd
```

CON-P1-03   12T

Fig. 4 Consensus HVR1 nucleotide and amino acid sequences. The consensus sequences were determined for populations forming phylogenetic clusters 1 and 2 (the populations isolated from 6 non-responders at T0; see Online Resource 8). The sequences are named according to patients' numbers. *Uppercase letters* indicate positions with 100% identity within one population. **a** Consensus nucleotide sequences. Quasispecies-specific, single-nucleotide variations, determined in relation to the CON-P1-01 sequence, were observed within all other consensus sequences (*shaded*) and indicated next to the alignment. **b** Consensus amino acid sequences. Consensus sequences of the quasispecies grouping in cluster 1 are identical (P1-01, P1-02, P1-05, P2-05, P2-19). Quasispecies-specific single-amino-acid variation (12T) was only found in one consensus sequence obtained for the population forming cluster 2 (P1-03) (*shaded*)

Table 2 Selection pressures acting on the HVR1 coding region in HCV populations present in patients with diverse responses to treatment

Dataset	Global dN/dS ratio		Codons under positive selection (numbers refer to HVR1)		Codons under negative selection (numbers refer to HVR1)	
	T_0	T_2	T_0	T_2	T_0	T_2
NR	0.672	0.707	16, 22	8, 16, 21, 22	2, 7, 20, 23	1, 2, 7, 20, 26
TR	0.993	0.935	12, 17	11, 12, 16, 17, 21	1, 2, 15, 23	1, 2, 15, 20, 23
SR	1.792	1.614	10, 11, 12, 13, 14, 16, 17, 18, 21, 22	5, 11, 12, 13, 14, 16, 17, 21, 22	1, 23	1, 2, 23, 26

NR no response; *TR* transient response; *SR* sustained response

quasispecies. The first type, formed by one dominating variant accompanied by a small number of closely related ones, was characteristic of the non-responding patients. The second type, composed of a large number of distantly associated equally ranked variants was typical for sustained responders. In both groups of patients, the level of HCV RNA varied over a wide range (Table 1), and we did not find a statistically significant correlation between the accumulation of viral RNA and HCV quasispecies variability or therapy outcome. Unfortunately, the amount of information concerning the correlation between the complexity of HCV quasispecies and the efficacy of anti-HCV treatment in children is still very limited. The existence of two patterns of HCV evolution in a group of 12 untreated neonates with HCV infection was reported previously by Farci et al. [6]. They observed that mono- or oligoclonal HCV populations were associated with a high level of ALT activity and hepatic injury, while heterogenous ones were associated with a lower ALT activity and mild or no liver damage [6]. In the case of children with CHC, we did not

observe such a correlation between ALT activity and HCV population structure. Interestingly, our findings are also inconsistent with previous observations made for adults with CHC undergoing anti-HCV therapy. In that study, it was suggested that homogeneity of HCV quasispecies at the baseline [21, 29, 30] or dramatic reduction of HCV genetic diversity during the first weeks of the therapy [5] is associated with viral clearance. In contrast, we found that patients with an almost homogenous HCV quasispecies at T0 did not eliminate the virus, while the ones with heterogeneous populations developed a sustained response. Similarly to Farci et al. [5], we observed a reduction of HCV quasispecies diversity in some sustained responders. Taking into account differences in the course and manifestation of chronic hepatitis C in adults and children, one can speculate that the mechanism underlying resistance to anti-HCV treatment may also vary in these two groups of patients.

HVR1 is the main target of host antibodies and is the most variable region of the HCV polyprotein. Hence, the stability of this region reported here seems to be a surprising discovery. HVR1 homogeneity was reported earlier in immunocompromised children and adults [10]. In the group we classified as non-responders, there were three patients who had previously undergone immunosuppressive treatment due to an oncological disease who were anti-HCV negative by fourth-generation ELISA (P2-05, P2-10, P2-28; Table 1). However, at T0, the level of HVR1 diversity in these children was similar to that in anti-HCV-positive non-responders. These observations suggest that the conservation of HVR1 in non-responders is not caused by defects in their humoral immune response.

An analysis undertaken by Gerotto and coworkers showed that, in children with CHC, the degree of HVR1 variability increases continuously in parallel with the natural development of the host immune system [12]. As a consequence, a more constant HVR1 is expected in younger children. Since the mean age of non-responders was almost the same as that of sustained responders (11.1 years, range 8–14, vs. 11.8 years, range 8–16, respectively), the observed changes in HVR1 variability could not be explained by differences in immune system development. Statistical analysis also did not reveal any correlation between the patients' age and the level of HVR1 variability within the examined group of patients or within two subgroups: NR and SR patients.

Only recently, a lack of HVR1 variability over a long period of time has been reported in two immunocompetent perinatally HCV-1a/c-infected children [13]. Both of them were subjected to interferon alpha monotherapy, and both failed to eliminate the virus. In addition, Gismondi and coworkers detected almost complete conservation of the HVR1 amino acid sequence, accompanied by a strong negative selection operating at certain positions. A similar observation was made with our non-responders. Gismondi's group proposed two possible explanations for this HVR1 homogeneity: immunotolerance and adaptation of the virus to the specific conditions within each host. The results presented here favor latter hypothesis. In addition to intra-quasispecies homogeneity, we identified the surprising phenomenon of inter-quasispecies conservation of HVR1 variants in children who failed to clear the virus. It is plausible that these variants demonstrate the highest fitness, resulting from optimal structural and/or functional properties. Accordingly, selection would favor their conservation. However, a hypothetical mechanism of immunotolerance in NR patients cannot be ruled out. It is probably a combination of the viral factors and the immunological aspects of a given patient that determines the HCV persistence. Our findings support earlier observations that variants that are inherently resistant to therapy may be present before its onset. It is also worth mentioning that non-responder-specific conservative HVR1 variants were never found in populations isolated from sustained responders. Based on the data provided by Gismondi et al. [13], one can hypothesize that these variants remain unchanged throughout treatment. In contrast, in children demonstrating a sustained response, positive selection seems to accelerate viral evolution to generate a more fit population. The high variability in these quasispecies probably reflects an ongoing process of adaptation of suboptimal variants—apparently an unsuccessful one, at least in terms of viral survival, since these populations were eliminated during treatment.

At present, it would be highly tentative to speculate what specific features of the conserved HVR1 variants underlie their superiority. It is certainly beyond the scope of this study. Gismondi et al. predicted *in silico* that the identified variants are antigenic. This suggests that the invariability of HVR1 is not due to the lack of an adequate humoral immune response [13]. Recently, we have expressed the E2 gene in *E. coli*. HVR1 of the resulting protein was identical to the variant most prevalent in cluster 1. During the preliminary experiments, we were able to detect the partially purified recombinant E2 protein by ELISA with homologous and heterologous patient sera [P.J., M.F., M.F., unpublished results]. Even if HVR1 alone was not targeted by the antibodies (obviously, this could not be concluded from the initial studies with patient sera), other regions of E2 were, calling into question the need to keep HVR1 unchanged. Consequently, it is presumably not immune evasion that causes HVR1 homogeneity. For this reason, to gain insight into the mechanisms driving HVR1 variability, attention should be focused on other processes involving the E2 protein during the host-virus interaction.

The same conclusion can be drawn from the distribution of codons subjected to positive selection in our sustained responders. Previously, two antigenic stretches (spanning residues 3-13 and 19-25) were identified within HVR1 [22]. Together, they constitute the majority of HVR1. Surprisingly, our results demonstrate that positively selected codons prevail in the central, non-antigenic region (10–18), and these are probably involved in maintaining HVR1 conformation and interactions with other molecules [22]. This observation indicates that adaptation to host proteins is at least as an important factor driving HVR1 evolution in children as diversification caused by host antibodies.

The results presented in this paper suggest that the degree of baseline HCV quasispecies diversity correlates with the outcome of chronic hepatitis C treatment in children. If confirmed in a larger group of patients in the future, this observation would be of great importance for further optimization of the anti-HCV therapy. We also provide another line of evidence that HVR1 diversification and evolution may be driven differently in pediatric patients than in adults. This matter should be considered during prospective clinical and virological studies.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Alberti A, Benvegnu L (2003) Management of hepatitis C. J Hepatol 38(Suppl 1):104–118
2. Canobio S, Guilbert CM, Troesch M, Samson J, Lemay M, Pelletier VA, Bernard-Bonnin AC, Kozielski R, Lapointe N, Martin SR, Soudeyns H (2004) Differing patterns of liver disease progression and hepatitis C virus (HCV) quasispecies evolution in children vertically coinfected with HCV and human immunodeficiency virus type 1. J Clin Microbiol 42:4365–4369
3. Domingo E, Escarmis C, Sevilla N, Moya A, Elena SF, Quer J, Novella IS, Holland JJ (1996) Basic concepts in RNA virus evolution. Faseb J 10:859–864
4. Domingo E, Escarmis C, Lazaro E, Manrubia SC (2005) Quasispecies dynamics and RNA virus extinction. Virus Res 107:129–139
5. Farci P, Strazzera R, Alter HJ, Farci S, Degioannis D, Coiana A, Peddis G, Usai F, Serra G, Chessa L, Diaz G, Balestrieri A, Purcell RH (2002) Early changes in hepatitis C viral quasispecies during interferon therapy predict the therapeutic outcome. Proc Natl Acad Sci USA 99:3081–3086
6. Farci P, Quinti I, Farci S, Alter HJ, Strazzera R, Palomba E, Coiana A, Cao D, Casadei AM, Ledda R, Iorio R, Vegnente A, Diaz G, Tovo PA (2006) Evolution of hepatitis C viral quasispecies and hepatic injury in perinatally infected children followed prospectively. Proc Natl Acad Sci USA 103:8475–8480
7. Ferenci P (2004) Predictors of response to therapy for chronic hepatitis C. Semin Liver Dis 24(Suppl 2):25–31
8. Figlerowicz M, Alejska M, Kurzyńska-Kokorniak A, Figlerowicz M (2003) Genetic variability: the key problem in the prevention and therapy of RNA-based virus infections. Med Res Rev 23:488–518
9. Figlerowicz M, Służewski W, Kowala-Piaskowska A, Mozer-Lisewska I (2004) Interferon alpha and ribavirin in the treatment of children with chronic hepatitis C. Eur J Pediatr 163:265–267
10. Gaud U, Langer B, Petropoulou T, Thomas HC, Karayiannis P (2003) Changes in hypervariable region 1 of the envelope 2 glycoprotein of hepatitis C virus in children and adults with humoral immune defects. J Med Virol 69:350–356
11. Gerlach TJ, Zachoval R, Gruener N, Jung M-C, Ulsenheimer A, Schraut W, Schirren A, Waechtler M, Backmund M, Diepolder HGP (2001) Acute hepatitis C: natural course and response to antiviral treatment. Hepatology 34:A341
12. Gerotto M, Resti M, Dal Pero F, Migliorato I, Alberti A, Bortolotti F (2006) Evolution of hepatitis C virus quasispecies in children with chronic hepatitis C. Infection 34:62–65
13. Gismondi MI, Becker PD, Diaz Carrasco JM, Guzman CA, Campos RH, Preciado MV (2009) Evolution of hepatitis C virus hypervariable region 1 in immunocompetent children born to HCV-infected mothers. J Viral Hepat 16:332–339
14. Hadziyannis S, Cheinquer JH, Morgan T, Diago M, Jensen DM, Sette H Jr, Ramadori G (2002) Peginterferon alfa-2a (40kD) (PEGASYS) in combination with ribavirin (RBV): efficacy and safety results from a phase III, randomised, double-blind multicentre study examining effect of duration of treatment and RBV dose. J Hepatol 36(Supp 1):3
15. Jonas MM (2002) Children with hepatitis C. Hepatology 36:S173–S178
16. Kuiken C, Yusim K, Boykin L, Richardson R (2005) The Los Alamos hepatitis C sequence database. Bioinformatics 21:379–384
17. Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinform 9:299–306
18. McHutchison JG, Gordon SC, Schiff ER, Shiffman ML, Lee WM, Rustgi VK, Goodman ZD, Ling MH, Cort S, Albrecht JK (1998) Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. N Engl J Med 339:1485–1492
19. Mizokami M, Ohno T (1998) Determination of HCV quasispecies by cloning and sequencing. In: Lau JYN (ed) Hepatitis C protocols. Humana Press, New Jersey, USA, pp 207–211
20. Moreau I, Levis J, Crosbie O, Kenny-Walsh E, Fanning LJ (2008) Correlation between pre-treatment quasispecies complexity and treatment outcome in chronic HCV genotype 3a. Virol J 5:78
21. Pawlotsky JM, Pellerin M, Bouvier M, Roudot-Thoraval F, Germanidis G, Bastie A, Darthuy F, Remire J, Soussy CJ, Dhumeaux D (1998) Genetic complexity of the hypervariable region 1 (HVR1) of hepatitis C virus (HCV): influence on the characteristics of the infection and responses to interferon alfa therapy in patients with chronic hepatitis C. J Med Virol 54:256–264
22. Penin F, Combet C, Germanidis G, Frainais PO, Deleage G, Pawlotsky JM (2001) Conservation of the conformation and positive charges of hepatitis C virus E2 envelope glycoprotein hypervariable region 1 points to a role in cell attachment. J Virol 75:5703–5710

23. Pond SL, Frost SD (2005) Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21:2531–2533

24. Poynard T, Marcellin P, Lee SS, Niederau C, Minuk GS, Ideo G, Bain V, Heathcote J, Zeuzem S, Trepo C, Albrecht J (1998) Randomised trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. international hepatitis interventional therapy group (IHIT). Lancet 352:1426–1432

25. Puig-Basagoiti F, Forns X, Furcic I, Ampurdanes S, Gimenez-Barcons M, Franco S, Sanchez-Tapias JM, Saiz JC (2005) Dynamics of hepatitis C virus NS5A quasispecies during interferon and ribavirin therapy in responder and non-responder patients with genotype 1b chronic hepatitis C. J Gen Virol 86:1067–1075

26. Quesnel-Vallieres M, Lemay M, Lapointe N, Martin SR, Soudeyns H (2008) HCV quasispecies evolution during treatment with interferon alfa-2b and ribavirin in two children coinfected with HCV and HIV-1. J Clin Virol 43:236–240

27. Simmonds P (2004) Genetic diversity and evolution of hepatitis C virus—15 years on. J Gen Virol 85:3173–3188

28. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

29. Toyoda H, Kumada T, Nakano S, Takeda I, Sugiyama K, Osada T, Kiriyama S, Sone Y, Kinoshita M, Hadama T (1997) Quasispecies nature of hepatitis C virus and response to alpha interferon: significance as a predictor of direct response to interferon. J Hepatol 26:6–13

30. Ueda E, Enomoto N, Sakamoto N, Hamano K, Sato C, Izumi N, Watanabe M (2004) Changes of HCV quasispecies during combination therapy with interferon and ribavirin. Hepatol Res 29:89–96