Special Article

Variability of Hepatitis C Virus

PETER SIMMONDS

The discovery and characterization of the main causative agent of non-A, non-B hepatitis (NANBH) represent a major success in the application of cloning techniques in the identification of new infectious agents.[1] The technique that led to the characterization and the development of antibody tests for hepatitis C virus (HCV)[2] has since been applied successfully to determine the etiological agent of other diseases, such as hepatitis E virus, responsible for enterically transmitted NANBH.[3] HCV contains a positive sense RNA genome approximately 10,000 bases in length. In overall genome organization and presumed method of replication, it is most similar to members of the *flaviviridae* family particularly in the coding for a single polyprotein that is then cleaved into a series of structural proteins (nucleocapsid protein, two membrane glycoproteins) and nonstructural proteins, with presumed enzymatic roles in virus replication.

Without a cell culture system to investigate differences in neutralization and cytopathic properties of HCV, nucleotide sequence comparisons and typing assays developed from sequence data have become the principal techniques for characterizing different variants of HCV. This type of analysis is relatively easy to perform, especially because viral sequences can be amplified by the polymerase chain reaction (PCR) directly from clinical specimens. Sequence comparisons of HCV have provided information about the virus at several levels. It is possible to identify and classify HCV into a series of distinct "genotypes," which differ substantially in nucleotide sequence from one another and show varied geographical and epidemiological distributions. In particular, the inferred amino acid sequences of the envelope glycoproteins differ considerably, and it is likely that antibody elicited by infection with one genotype would fail to neutralize others. The extent of variation observed within HCV is comparable with that between serotypes of other RNA viruses, and may pose problems in the development of vaccines for HCV.

HCV differs from most RNA viruses in its ability to establish chronic infection and progressive disease in a large proportion of those exposed. The mechanism for HCV persistence remains uncertain; most explanations are based on the concept of "immune escape" in which rapid sequence change in the envelope proteins alters the antigenicity of the outer virus surface sufficiently to avoid neutralization by the host humoral immune system. Indeed, the relatively high rates of nucleotide mis-incorporation on replication of HCV might contribute to the genetic plasticity of envelope proteins and facilitate the generation of novel antigenic variants. However, it remains unclear why "immune escape" should occur specifically in HCV infection and not on infection with other RNA viruses where mutation rates are equally high, but where infection in immunocompetent individuals is almost always transient and rapidly eliminated after the appearance of a specific immune response.

In this review I will attempt to describe different areas of research concerning the sequence variability of HCV, including the rationale for a proposed classification into genotypes and the biological and serological consequences of variability. The review also describes possible selective pressures on the virus during persistent infection that may restrict sequence variability in certain regions of the genome, but perhaps favor it elsewhere. Whether the routine identification of genotypes of HCV infecting patients has any clinical utility is currently uncertain. However, there is convincing evidence for systematic differences in the course of HCV infection and response to antiviral treatment between genotypes, although the underlying mechanism for these differences and whether genotype determination is the most appropriate pretreatment variable to monitor in a patient are currently unknown.

SEQUENCE ANALYSIS OF HCV

Our present understanding of the genomic organization and method of replication of HCV has been largely inferred from analysis of its nucleotide sequence and comparison with other positive strand viruses. The genome contains a single potential open reading frame that would produce an extremely large protein (of more than 3,000 amino acids). By analogy with flaviviruses and picornaviruses, which show a similar genetic organization and where it is possible to study replication experimentally, it is thought that the HCV "polyprotein" is cleaved after translation to produce a series

Abbreviations: NANBH, non-A, non-B hepatitis; HCV, hepatitis C virus; PCR, polymerase chain reaction; NCR, noncoding region; PT, prototype; RFLP, restriction fragment length polymorphism.

From the Department of Medical Microbiology, University of Edinburgh, Edinburgh, United Kingdom.

Received April 23, 1994; accepted July 28, 1994.

Address reprint requests to: Peter Simmonds, BM, Department of Medical Microbiology, University of Edinburgh, Teviot Place, Edinburgh, EH8 9AG, United Kingdom.

0270-9139/95/2102-0042$3.00/0

TABLE 1. Properities of HCV Proteins

Protein	Nucleotide Position*	Size†	In Virion	Antigen‡	Function
Core	1-573	191	Yes	c22-3	Thought to form the virus nucleocapsid and shows RNA-binding activity. Highly conserved between HCV genotypes.
E1	574-1149	192	Yes		Sequence predicts a membrane anchored glycoprotein, with several potential *N*-linked glycosylation sites. Highly variable between genotypes.
E2/NS-1	1150-2430(?)	327(?)	?		Another membrane bound glycoprotein, with several potential *N*-linked glycosylation sites. Highly variable between genotypes. 3′ cleavage site uncertain.
NS-2	2431-3078	316	No		Metallo-proteinase
NS-3	3079-4971	631	No	c33c	Multifunctional protein with protease activity and sequence motifs suggesting a helicase activity
NS-4a	4972-5133	54	No	c100-3	Unknown
NS-4b	5134-5916	261	No		Unknown
NS-5a	5917-7260	448	No	NS-5	Unknown (? replicase function)
NS-5b	7261-9033	591	No		Probably the RNA-dependent RNA polymerase necessary for genome replication

* Nucleotide positions numbered as in reference 4.
† Size expressed in number of amino acid residues.
‡ Origin of HCV antigens used in current second- and third-generation assays for antibody to HCV.

of structural and nonstructural proteins necessary for replication and virion formation (Table 1).

The precise relationship between HCV and other positive-stranded RNA viruses remains unclear. The current proposal is to assign HCV as a separate genus within the *Flaviviridae,*[4,5] a group that contains the (insect- and tick-borne) flaviviruses as its prototype genus, and which has recently been expanded to accommodate the pestiviruses.[6] However, the proposed genera differ from each other in the number and arrangement of proteins and, more significantly, in method of replication. Translation of flavivirus genome proceeds as for eukaryotic messenger RNAs by ribosomal binding to the 5′ capped viral RNA and scanning to the first methionine codon.[7] In contrast, translation of the HCV and pestivirus polyproteins is thought to be initiated by cap-independent ribosomal binding to an acceptor site in the 5′ noncoding region (5′NCR) and translation from an internal methionine codon.

Structurally, HCV is also more similar to the pestiviruses than flaviviruses, showing an exceptionally low buoyant density in sucrose (1.08 to 1.11 g/cm^3),[8] similar to that reported for pestiviruses and attributable in both cases to heavily glycosylated external membrane glycoproteins in the virus envelope. By contrast, flavivirus envelope glycoproteins contain few sites for *N*-linked glycosylation, and the virion itself is relatively dense (1.2 g/cm^3). Unfortunately, the difficulty in visualizing HCV has so far prevented a precise morphological comparison between HCV and other viruses.

In common with other RNA viruses, variants of HCV show considerable sequence variability, many differing considerably from the prototype HCV (PT).[4] Differences of up to 34% have been found between the complete genomic sequences of the most extremely divergent variants analyzed to date, comparable with that observed between serotypes of other human positive-strand RNA viruses, such as polio, Coxsackie, and coronaviruses. Sequence variability is evenly distributed throughout all viral genes (Fig. 1), apart from the highly conserved nucleotide (and amino acid) sequences of the core (nucleocapsid) protein and 5′NCR, and the greater variability of the envelope gene (Fig. 1).

CLASSIFICATION OF HCV INTO GENOTYPES

No straightforward method for *in vitro* culture currently exists for HCV, and, therefore, many of the traditional virological methods for virus classification re-

A

Type	5′NCR	Core	E1	E2/NS-1	NS-2	NS-3	NS-4	NS-5	3′UTR	TOTAL
1a	0	1	6	7	7	5	4	4	4	5
1b	2	9	25	28	29	20	21	21	33	21
1c	1	10	23	23	28	21	20	20	36	20
2a	6	19	38	33	43	30	33	34	70	32
2b	6	19	44	33	43	30	35	34	74	33
3a	9	16	35	33	42	31	31	32	50	32

B

Type	Core	E1	E2/NS-1	NS-2	NS-3	NS-4	NS-5	TOTAL
1a	1	6	7	7	5	4	4	4
1b	2	21	22	23	9	13	16	15
1c	2	24	17	25	8	12	16	14
2a	10	40	29	42	19	27	30	28
2b	10	45	26	44	19	28	31	28
3a	9	34	28	41	19	20	28	25

FIG. 1. (A) Percentage nucleotide and (B) % amino acid sequence divergence between the complete genomic sequence of HCV-PT[10] with variants representing other HCV genotypes. Sources of sequences: 1a: HCV-H[104]; 1b: HCV-J[9]; 1c: HC-J9[24]; 2a: HC-J6[12]; 2b: HC-J8[11]; and 3a: NZL1.[105] In all comparisons, the 5′NCR is the most conserved subgenomic region (maximum 9% nucleotide sequence divergence), whereas highly variable regions are found in parts of the genome encoding E1 and NS-2 (35% to 44% nucleotide sequence and 34% to 45% amino acid sequence differences).

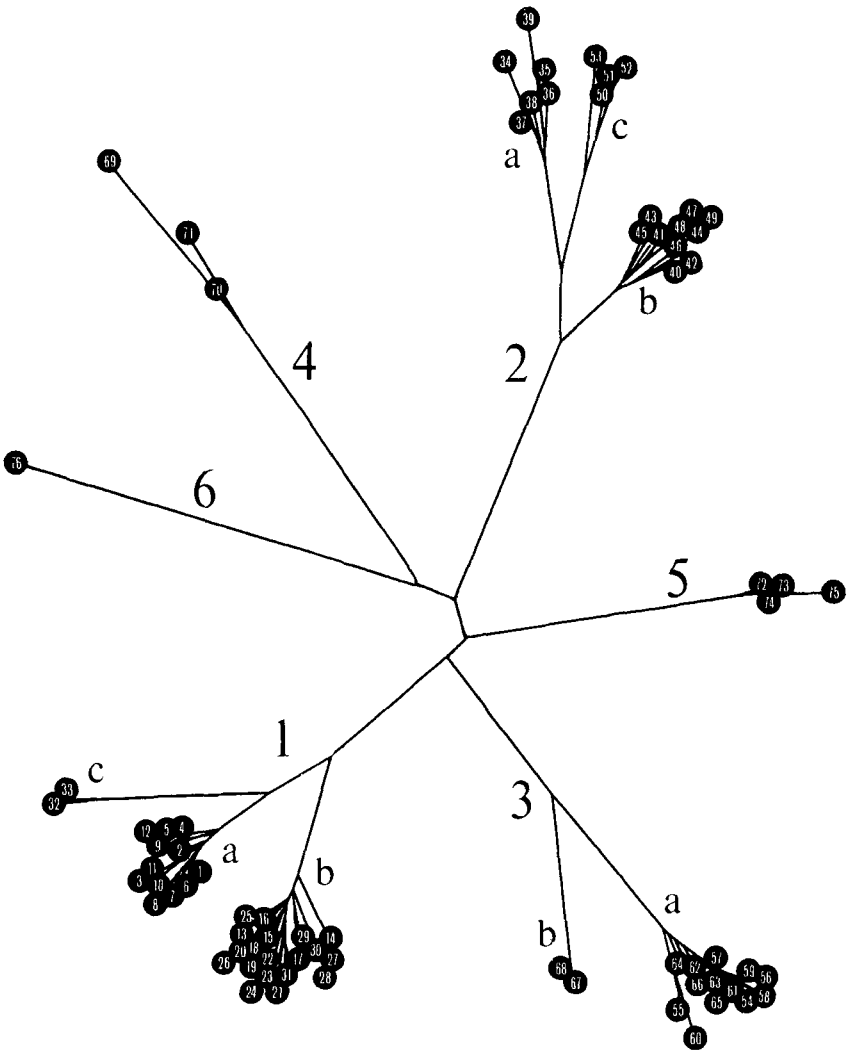

FIG. 2. Phylogenetic analysis of nucleotide sequences in part of the HCV NS-5 region. Individual sequences (numbered dots, identified in reference 16) are separated by branch lengths proportional to their reconstructed evolutionary distances. Six main groups of sequence variants are shown (1 to 6); those labeled 1, 2, and 3 each comprise a number of more closely related sequences. The numbering of the phylogenetic groups follows the scheme proposed in Table 2.

main untried. For example, there is no neutralization assay to show whether distinct serotypes of HCV exist among the known variants of HCV. Inevitably, classification of HCV has relied almost entirely on nucleotide sequence comparison of complete genomes or subgenomic fragments between variants. Whether this leads to a useful biological division of HCV currently remains an open question that will have to re-investigated when other methods of virus characterization become possible.

Variants of HCV obtained from Japan show substantial sequence differences from the prototype HCV variant obtained in the United States, HCV-PT.[4] Comparison of the complete genome sequence of HCV-J[9] and -BK[10] from Japan showed 92% sequence similarity to each other, but only 79% with HCV-PT. At that time, the former were referred to as the "Japanese" type (or type II), whereas those from the United States (HCV-PT and -H) were classified as type I. However, far more divergent variants of HCV have since been found in Japan[11,12] and elsewhere,[13,14] and inconsistencies in nomenclature of the variants of HCV rapidly appeared in the literature.

One way to investigate the underlying relationships between the known variants of HCV is to perform phylogenetic analysis of nucleotide sequences of complete genome or subgenomic regions. We have previously analyzed a 222-base fragment of the NS-5 gene amplified by PCR[15] from plasma samples of a series of HCV-infected individuals in Europe, North and South America, and the Far East.[16] The region was chosen because it was variable (up to 44% sequence divergence between some variants of HCV), could be readily amplified by PCR, and there exist a number of previously published sequences. A phylogenetic tree constructed from the sequence data showed that the variability of HCV was structured into six equally divergent main groups of sequences, many of which contained more closely related groups within them (Fig. 2).

In trying to place the nomenclature for HCV variants on a rational basis, we thought that it should reflect to a certain extent the hierarchical sequence relationships between different isolates. Therefore, based on previous suggestions,[14,15] we proposed a nomenclature of "types," corresponding to the major branches in the phylogenetic tree and "subtypes" corresponding to the

TABLE 2. Comparison of Different Classifications for HCV Genotypes

Genotype	Example	Reference No.	Chiron	Enomoto	Mori/ Okamoto
1a	HCV-1	4	I	K-PT	I
1b	HCV-J, -BK	9, 10	II	K-1	II
1c	EG-28	17	nc	nc	nc
"1c"*	HC-J9	24	nc	nc	1c
2a	HC-J6	12	III	K-2a	III
2b	HC-J8	11	III	K-2b	IV
2c	T0994	101	III	nc	nc
3a	E-b1, Ta	13, 14	IV	nc	V
3b	Tb, Tr(V)	13, 61	IV	nc	VI
4†	EG-16, 29, 33	17	nc	nc	nc
5a	SA-1, 7, 11	39, 40	V	nc	nc
6a	HK-1, 2, 3, 4	16, 39	nc	nc	nc

NOTE. Proposed nomenclature for published HCV sequences[14,17] and comparison with existing schemes (Chiron[40,102]; Enomoto[15]; Mori/Okamoto)[11,13]. The proposed system reproduces the type/subtype relationship between HCV sequences apparent upon phylogenetic analysis of nucleotide sequences in the NS-5 and other coding regions (see Fig. 2). Sequences not classified by originating authors are indicated as "nc."

* Complete genome of a further subtype of type 1 that is distinct from types 1a and 1b and from a previously described subtype of the same name.[16]

† Type 4a was the name proposed for genotypes found in Egypt and elsewhere in the Middle East on the basis of sequence comparisons in core[17] and NS-5 regions.[16] Its relationship to other type 4 variants found in Central Africa[18,50] remains unclear as different regions of the genome have been sequenced.

more closely related sequences within some of the major groups (Fig. 2; Table 2).[16] The types have been numbered 1 to 6 and the subtypes have been lettered a, b, and c, in both cases in order of discovery. Therefore, the sequence cloned by Chiron is assigned type 1a, HCV-J and -BK are 1b, HC-J6 is type 2a, and HC-J8 is 2b. This nomenclature closely follows the schemes originally described by Enomoto et al,[15] on the basis of sequence comparisons of the NS-5 region (PT [=1a], K1 [=1b], K2a [=2a], and K2b [=2b]) and by Chan et al,[14] who provided evidence for a further HCV type (type 3a) on the basis of phylogenetic analysis of sequences in the NS-5, NS-3, core, and 5'NCRs. This approach avoids the inconsistencies of earlier systems and is easier to extend when new genotypes are discovered. For example, the original Chiron scheme assigned two subtypes 1a and 1b as group I and group II, but the other groups corresponded to separate major genotypes (group 3 = type 2a, 2b, and 2c; group IV = type 3a and 3b; group V = type 5a). The Okamoto/Mori system assigned types I to VI to the first six genotypes discovered, irrespective of their sequence relationships, so that type I corresponds to 1a, II to 1b, III to 2a, IV to 2b, V to 3a, and VI to 3b. However, sequences corresponding to further subtypes of types 1 and 2 have since been found (types 1c and 2c). Although type 2c has been added into group III of the original Chiron classification, it would be difficult to justify adding 1c into either groups I or II or to assign it as a new group (group VI ??). In the Okamoto system, it would be misleading to call 1c and 2c as types VII and VIII, given their closer relationship to types I and II and type III/IV respectively. Neither of these two schemes assigns the recently described type 4a variants[17] or the provisionally assigned new genotypes, types 5a and 6a.[16,18]

Whether sequence relationships between HCV genotypes are equivalent in different regions of the genome has been a matter of controversy and remains currently unresolved. However, comparisons of an extensive set of sequences in the envelope region of HCV (E1), either at the amino acid level[18] or by phylogenetic analysis of nucleotide sequences,[19] have provided evidence for six major genotypes and several subtypes analogous to those observed in NS-5. Indeed, by matching variants through sequence comparisons of the 5'NCR, we found that each of the six genotypes and most of the subtypes described for E1 and NS-5 were equivalent.[19] It now appears that any coding region could be used for identification of genotypes and would produce equivalent relationships between genotypes.[19,20] At present, suitable comparative sequence data for the known genotypes exist for part of NS-5,[16] NS-4 (unpublished data), and E1.[18]

Sequences in the upstream 5'NCR are far more conserved (6% sequence divergence between type 1 and type 2 sequences compared with 33% over the whole genome). Relatively few sequence differences exist between different HCV types, whereas sequences of different subtypes may be identical. However, the 5'NCR sequences of genotypes 1 to 6 are distinct, and several virus typing methods have been developed that are based on amplified sequences from this region.[21-23] In

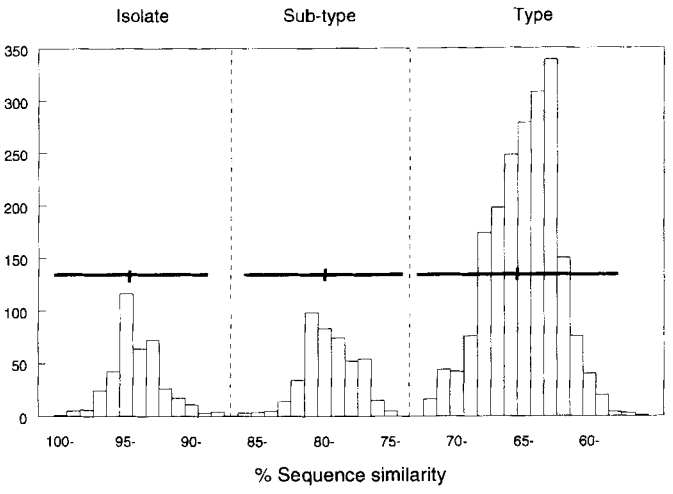

FIG. 3. Discontinuous distribution of sequence distances on pair-wise comparison of sequences of different genotypes of HCV shown as a histogram. A total of 2,850 individual pair-wise comparisons were made between 76 NS-5 nucleotide sequences of HCV variants from several geographic areas. Observed sequence similarities in increments of 1% recorded on y-axis. Vertical and horizontal bars show mean and range (±3 SDs) of sequence distances within each group. From this type of analysis, the range of sequence distances between variants of different major genotypes, subtypes, and within genotypes can be determined (see Table 3).

TABLE 3. % Sequence Similarity Between Type, Subtype, and Isolate in E1, NS-4, and NS-5 Regions

	E1	NS-4	NS-5
Start*	574	4940	7975
End*	1149	5299	8196
Length	576	360	222
No. of sequences	64	37	76
Genotypes	1-6	1-6	1-6
Isolate			
Min‡	99.1	97.7	99.5
Max‡	88.0	87.6	87.8
Subtype			
Min	78.6	80.9	86.0
Max	67.5	73.9	74.8
Type			
Min	69.4	68.9	72.1
Max	53.5	54.6	56.2

* Sequences numbered as in reference 4.

† Number of sequences compared.

‡ Minimum and maximum isolate, subtype, and type sequence similarities (%).

practice, although existing genotypes can be recognized, it is problematic to assign new genotypes on the basis of sequence comparisons in this region alone. Therefore, it is essential that sequence comparisons of putative new genotypes are extended to the coding regions.

The validity of the distinction between type and subtype is more clearly apparent when the range of pair-wise distances between sequences are compared. For example, in the NS-5 region, pair-wise comparison of sequences from the six major genotypes and several subtypes of HCV show variability restricted to three nonoverlapping distributions (Fig. 3). Using this approach, we have calculated sequence divergences for genotype, subtype, and isolate in several regions of the genome (Table 3). These ranges provide a rapid method for provisionally assigning new variants as major genotypes or subtypes (see below). An even clearer distinction between type, subtype, and isolate is found on pair-wise comparisons of nucleotide sequences of complete genomes. Using sequences from types 1a, 1b, 2a, 2b, 3a, and a newly described subtype of type 1 from Indonesia,[24] sequence similarities between major genotypes are tightly confined to a range from 66% to 69%, from 77% to 80% between subtypes, and from 91% to 98% between epidemiologically unrelated variants within a genotype.

One potential problem of nomenclature based purely on genotypic classification is the possibility of hybrid viruses arising by recombination. It would be difficult to classify a variant that contained type 1a sequences at one end of the genome and type 2a sequences at the other. There is currently little evidence for the existence of such hybrids, as shown by the existing complete genome sequences and parallel analyses of samples in several regions of the genome (see below), although recombination does occur in RNA viruses, and more work is needed to show whether or not it occurs in HCV.

PROPOSAL FOR AN AGREED NOMENCLATURE FOR HCV GENOTYPES

If the nomenclature of HCV remains based on nucleotide sequence comparisons, then clearly the most definitive classification would be one based on sequences of complete genomes. However, as described above, it appears that sequence comparisons of subgenomic regions may accurately reproduce those that exist between complete sequences, providing a more practical method for detection and designation of new genotypes as they are discovered. Indeed, sequence data for the known genotypes of HCV exist for several subgenomic regions (NS-5, E1, and NS-4), whereas complete genomic sequences have been obtained only for types 1a, 1b (and a third subtype of type 1), 2a, 2b, and 3a.

The specific details of agreed proposals for assignment of new genotypes have been published elsewhere.[25] In brief, we have suggested that identification and classification of new genotypes should be divided into at least two stages. Provisional identification could initially be based on measurement of percentage sequence similarities between the new sequence and those previously classified, using the known ranges of sequence variability between type, subtype, and isolate (as shown in Table 3). For example, in the part of NS-5 amplified by primers described by Enomoto et al,[15] sequence similarities of less than 72% with any known sequence would be evidence for the existence of a new HCV type, whereas maximum sequence similarities of 75% to 86% with some members of the dataset would provide evidence of a new subtype (Table 3).

Confirmation of the existence of a new genotype should be made by phylogenetic analysis of the new sequence with those already obtained. The branching order within the resulting unrooted tree will indicate the relationship of the variants with existing genotypes. Secondly, until the issue of recombination is resolved (see above), we believe that it would also be necessary to show equivalent phylogenetic relationships using nucleotide sequences from at least one other coding region of the genome. Agreement between investigators is now needed over which parts of the genome should be compared and how long the sequences should be to reliably identify and classify new HCV genotypes.

In the long term, designation of names for new genotypes might be achieved by a committee (perhaps set up under the auspices of the International Committee for the Taxonomy of Viruses) to receive submissions from researchers with new sequence data. The delays inherent in publication and the unregulated practice of assigning genotype numbers and subtype letters to new variants by authors continues to cause confusion. For example, there are currently two different subtypes of type 1, both described as type "1c"[16,24] and two different type "4a" variants.[16,18]

HOW "OLD" ARE THE HCV GENOTYPES?

If one knew the rate of nucleotide sequence change of HCV with time, it ought to be possible in principal

to calculate the time of divergence between subtypes and types of HCV. Rates of sequence change over relatively short intervals of time are provided by a number of studies of HCV carriers or experimentally infected chimpanzees[26-28] or between individuals several years after transmission of HCV from one to the other.[29] Estimates of 0.144% (complete genome)[28] and 0.192% (5′ half of genome)[26] nucleotide changes per year have been reported. Using the former figure and assuming a constant rate of divergence, it can be calculated that variants of the same genotype, e.g., HCV-BK and HCV-J that differ by 9%, diverged approximately 30 years ago, whereas subtypes (20% to 23% sequence difference) diverged 70 to 80 years ago. Finally, the radiation of the major genotypes can be predicted to have occurred 100 to 120 years ago. This reasoning would suggest that HCV is a relatively recent infection of humans, with a relatively short evolutionary history compared with that of other human viruses.

However, when comparing very different sequences (such as between genotypes of HCV), it becomes increasingly difficult to reconstruct (chronological) times of divergence between lineages. Problems arise from the assumption that the rate of sequence change remains constant irrespective of the degree of divergence between variants. Although there is likely to be little overall change in the rate of frequency at which mutations occur over time, most of those that do occur would reduce the overall fitness of the progeny viruses and will be lost from the population by natural selection. Changes that are found on comparison of HCV sequences are usually "synonymous," i.e., do not alter the encoded amino acid sequence (usually by sequence changes at the third position of the codon). For example, over the 8 years that elapsed between the collection of variants HC-J4/83 and HC-J4/91 from an experimentally infected chimpanzee, 69 of the 111 nucleotide differences between the sequences were silent; approximately three times more than would be expected by chance. Clearly, without the functional constraints on protein sequence variability, the rate of sequence change would have been several times higher.

In fact, the distinction between synonymous and nonsynonymous (i.e., amino acid changing) substitutions is overly simplistic. Clearly, some amino acid replacements are functionally more significant than others; indeed, many may lead to the synthesis of functionally equivalent proteins in certain circumstances. Exactly which changes can be accommodated while retaining a "fit" phenotype remains an area of comparative ignorance, because little is known about the effect of amino acid sequence change on the structure and function of proteins. In some cases, a tendency for substitution of similar amino acids can be observed (for example, leucine for isoleucine, aspartate for glutamate), but it is difficult to generalize much further than this. The restraint on sequence change would presumably differ in each protein, and would depend on whether the substituted residues were functionally critical (i.e., part of an active enzymatic site) or not. Functional constraints on protein sequences may change through time. For example, HCV may have encountered different selection pressures on spread by different routes of transmission, and perhaps into different hosts, leading to increases or decreases in the degree of functional constraint acting on each amino acid residue. These could give rise to marked differences in the divergence rate.

The other factor that needs to be taken into account is the underestimation of sequence differences by the occurrence of multiple hits, in which some mutations will occur more than once at the same site, and lead to no net increase in sequence divergence (or even a decrease if the base changes back to the ancestral one). Correction for multiple hits allows the calculation of evolutionary distances between sequences, although there are several different algorithms for doing this that differ in the assumptions made about the underlying evolutionary processes.[30,31] A particular problem with analyzing sequences with high rates of synonymous substitutions (such as HCV) is the inequality in the rate of sequence change at different sites. For example, although HCV-PT and HC-J8 show 33% sequence divergence overall, most of the sequence differences are at silent sites while relatively few occur elsewhere. Therefore, most of the methods that correct for multiple substitution would greatly underestimate the true evolutionary distance between such sequences.

Taking all of these factors together, estimates for times of divergence are often much greater than those based on the "molecular clock." Although the time of divergence of different variants within a genotype (e.g., HCV-BK, HCV-J; 30 years) may be relatively accurate because sequence differences are relatively few, the time of divergence of sequences that differ considerably in sequence, e.g., PT and HC-J8, is likely to be many times greater than originally calculated. How much greater can only be guessed; 1,000 to 10,000 years are much more likely than the original estimates of 100 to 120 years.

SEQUENCE VARIABILITY IN THE 5′NCR

A different restraint on sequence change is the necessity for sequences in certain parts of the genome to internally base-pair to generate secondary structures. It is thought that the resulting stem/loop structures interact directly with viral and/or cellular proteins during virus RNA replication and translation. One example is the proposed interaction between ribosomes with sequences in the 5′NCR. By analogy with picornaviruses, it has been suggested that sequences in the 5′NCR direct ribosomal binding from the extreme 5′ end of the RNA (which may be uncapped) to an internal site, leading to initiation of protein synthesis from an internal methionine residue.[32-34] Although details of the interaction between RNA structures in the 5′NCR and the ribosome have yet to be determined, it is highly suggestive that secondary structure predictions for different variants of HCV are remarkably conserved. In particular, it has been possible to find a number of paired nucleotide changes that maintain base-pairing (covariance) across a proposed stem/loop structure within the 5′NCR (Fig. 4).[17] Significantly, the overall

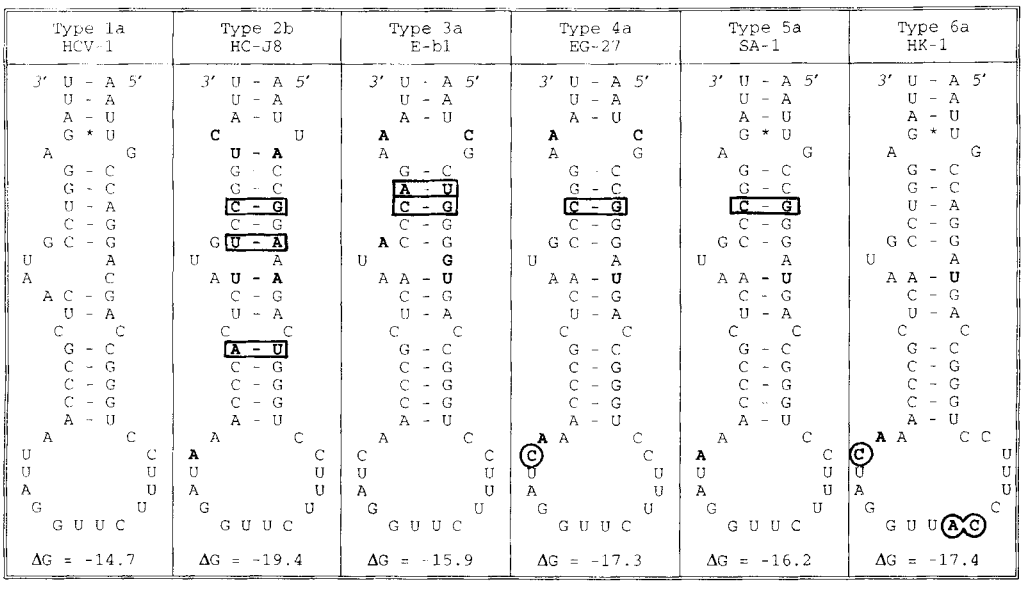

FIG. 4. Secondary structure predicted for representative 5′NCR sequences of types 1 to 6 between nucleotides −169 and −114. −, Watson-Crick base pairing between nucleotides within the stem. *, The noncanonical G-T base pairing that is tolerated in 2° structures. Differences from HCV-1 prototype sequence[4] are indicated in bold. Compensatory substitutions to maintain base pairing (shown in boxes) are found at several sites in the stem. At two positions, nucleotides (shown ringed) are inserted in the non–base-paired terminal loop. All structures show similar predicted stabilities in the conformations presented here; ΔG values in kcal/mol are shown under each sequence.

secondary structure for HCV closely matches that predicted for the related pestivirus genus (including the stem/loop structure described above), although there is almost no sequence similarity in the untranslated regions between the two virus groups.[5,33]

The proposed interactions between viral RNA sequences and cellular and viral proteins clearly constrain possible sequence variability in these regions. Many of the interactions may be absolutely dependent on the RNA sequence and, therefore, serve to eliminate naturally occurring variability at these sites. Almost as restrictive is the requirement for complementary bases in parts of the 5′NCR forming internally base-paired stems. A change in these regions may only produce viable viruses if accompanied by a compensating change in the opposite base in the stem to which it is bound to. The probability of two changes occurring within the same replication cycle are remote. It is likely that these two restraints on sequence variability are responsible for the highly conserved nature of sequences in the 5′NCR of HCV (6% sequence divergence between PT and HC-J8) compared with the coding regions (33%), where silent nucleotide changes may be tolerated.

Recent analysis of sequence variability in the core region has shown that the 5′ end shows far fewer amino acid replacement and silent changes than other parts of the coding region.[35] As sequences in this part of the core gene have been shown to be essential for the efficient *in vitro* expression of the HCV polyprotein, it is possible that restraints exerted by secondary structure requirements also operate in this part of the genome and perhaps elsewhere. Indeed, it is possible that there is a direct physical interaction between RNA from the core region with the ribosome or structures upstream in the 5′NCR during translation. In general, we are very ignorant about how the genetic material of single-stranded viruses may fold on replication and for packaging during virus assembly and release. Perhaps these processes further restrict sequence diversification throughout the viral genome in ways unseen by current methods of analysis.

GEOGRAPHICAL DISTRIBUTION OF HCV GENOTYPES

Some genotypes of HCV (types 1a, 1b, 2a, and 2b) show a broad worldwide distribution, whereas others, such as type 5a and 6a, are only found in specific geographical regions. Understanding the current distribution of HCV requires knowledge of routes of transmission, historical data on the prevalence and risk groups for HCV over the previous decades or centuries, and time of the divergence of the major genotypes and subtypes (see above). Unfortunately we know very little about any of these factors and, therefore, it is difficult to draw any definite conclusions from the existing information. The problem is compounded by the scarcity of information concerning genotype distributions worldwide; for example, very few countries have been specifically surveyed, and the genotype distributions of whole continents, such as Africa, remain largely unknown.

However, clearly discernible patterns of genotype distribution have been found in those countries that have been studied so far. For example, both blood donors and patients with chronic hepatitis from countries in Western Europe and the United States show frequent infection with genotypes 1a, 1b, 2a, 2b, and 3a, although the frequencies of each may vary.[16,18,22,23,36-46] There is a trend toward more frequent infection with type 1b in Southern and Eastern Europe. In many European countries genotype distributions vary with age of patients, reflecting rapid changes in genotype distribution with time within a single geographic area.

A striking geographic change in genotype distribution is apparent between South-East Europe and Turkey (mainly type 1b) and countries in the Middle East and parts of North and Central Africa where other ge-

notypes predominate. For example, a high frequency of HCV infection is found in Egypt (20% to 30%),[47-49] of which almost all correspond to the type 4a genotype.[17,41] HCV type 4 is also the principal genotype in countries such as Yemen, Kuwait, Iraq, and Saudi Arabia in the Middle East (unpublished information), and Zaire, Burundi, and Gabon in Central Africa.[18,39,50] Zaire, Gabon, and Gambia on the west coast of Africa have also been shown to be an area of high endemicity for HCV, with rates of infection approaching 20% to 30%[51] (Blumberg B, et al, Personal communication, April 1994). Although genotypes in the Middle East and Central Africa can be assigned as genotype 4 on the basis of sequence comparisons in the core, E1 and NS-5 regions,[16,18,50] it is apparent that this genotype actually comprises a bewildering number of subtypes. Although a specific genotype referred to as type 4a is the most frequently found genotype in Egypt[17] and elsewhere in the Middle East, we have detected at least three other subtypes of type 4 in this region alone. Similarly, E1 sequences from six infected individuals in Zaire yielded four subtypes, none of which correspond to those in the Middle East nor to many of those found in Gabon (Stuyver L, et al, VI International Symposium on Viral Hepatitis, Madrid, Spain, February 3-5, 1994).

HCV genotype 5a is frequently found among hepatitis C patients in South Africa[16,18,40,52] and is found in 30% to 50% of anti-HCV–positive donors in two areas of South Africa (Johannesburg and Durban; Davidson, et al, Submitted).[41] In contrast to type 4, relatively little sequence heterogeneity is found within this genotype, with all variants analyzed to date falling within a single subtype. Infection with this genotype is highly restricted geographically, being found only rarely in Europe or elsewhere.[18,42]

In the Far East, the genotype distributions are equally complex. Within Japan and Taiwan and probably parts of China, genotypes 1b, 2a, and 2b are the most frequently found types.[15,21,53-59] Infection with type 1a in Japan appears to be confined to hemophiliacs who have received commercial (United States–produced) blood products, such as factor VIII and XI clotting concentrates.[15,60] The distribution of type 3 shows marked geographic variability; it is only rarely found in Japan,[61] and is also infrequent in Taiwan, Hong Kong, and Macau. However, this genotype is found with increasing frequency in countries to the west, being frequently found in Singapore and accounting for the majority of infections in Thailand.[13] In a small sample, it was the only genotype found in Bangladesh and Eastern India (Mutimer D, Unpublished data, February 1994). As with type 4 in Africa, there is now evidence for considerable sequence diversity within the type 3 genotype. Six different subtypes of type 3 (types 3a to 3f) were found in infected patients in Nepal,[62] while we have found three more subtypes in Bangladesh and Thailand, making a current total of nine.

A genotype with a highly restricted geographic range is 6a. It was originally found in Hong Kong[17,39] and was shown to be a new major genotype by sequence comparisons in the NS-5 and E1 regions.[16,18] Approximately one third of anti-HCV–positive blood donors in Hong Kong are infected with this genotype, as are an equivalent proportion in neighboring Macau. A Vietnamese blood donor residing in Canada was also found to be infected with genotype 6.[43] There is currently insufficient data to show how extensively type 6a is distributed in South-East Asia. Unfortunately existing survey work performed in mainland China used an assay that could only detect genotypes 1a, 1b, 2a, and 2b[56,59]; had type 6a or possibly novel genotypes of HCV been present, they may not have been identified.

Knowledge of the geographic distributions of HCV genotypes not only provides information on virus origins and transmission, but will also be important for vaccine development and antiviral treatment. Almost nothing is known about whether genotypes 4, 5, and 6 are sensitive to interferon treatment, nor whether there is biological variation among the many subtypes of type 4 (see below). This data will be vital in addressing the major medical problem posed by high rates of HCV infection in countries such as Egypt and elsewhere in Africa. Similarly, a vaccine for HCV will probably have to be multivalent because the high degree of sequence divergence in the envelope gene makes cross-protection between genotypes unlikely.

BIOLOGICAL DIFFERENCES BETWEEN GENOTYPES OF HCV

An active area of research into HCV is the investigation of possible differences in the course of disease associated with different genotypes, such as the rate of development of cirrhosis and hepatocellular carcinoma, and whether certain genotypes are more or less likely to respond to interferon treatment. From a virological perspective, it is impossible to predict whether the degree of sequence difference between genotypes would influence the behavior of the virus. As discussed above, it can be inferred that most of the nucleotide sequence differences that exist between genotypes have accumulated through evolutionary drift, and would be expected to be approximately neutral in effect. Determinants of relative pathogenicity, should they exist between genotypes, might result from changes at only one or two key nucleotide or amino acid sites, and would be unlikely to be detected by simple sequence comparisons alone in relatively restricted regions of the genome.

Indeed, we have very little insight into which of the HCV proteins might influence the rate of replication, cytopathology, and resistance to antiviral treatment. There may be differences in the activity of viral proteinases in processing the HCV polyprotein after translation, or in the activities of proteins involved in RNA replication, leading to alteration in their ability to induce an endogenous interferon response. Variation in the envelope proteins may contribute to differences in cell tropism and the extent to which infection causes cytopathic changes in the cell during the latter stages of virus assembly and release. It is also possible that the characteristics of the virus infection *in vivo* are determined by noncoding parts of the genome; perhaps

the configuration of the 5′ or 3′NCR affects binding of viral and cellular proteins involved in replication and translation, as has been suggested for type 1 and type 2 RNA constructs *in vitro*.[32]

It is similarly difficult to infer from comparisons with other RNA viruses whether genotypes of HCV differ in their biological properties. Most positive-stranded RNA virus genera comprise variants that are equally or even more divergent than the major genotypes of HCV, yet in many cases each member behaves similarly on infection in its host (e.g., poliovirus types 1, 2, or 3, dengue fever virus, serotypes 1 to 4). However, other viruses, such as the 6 serotypes of Coxsackie B show considerable variability in the degree of disease caused and the main organs in the body targeted by infection. Pestiviruses are probably the most closely related viruses to HCV; this genus currently comprises three distinct variants, which show a degree of sequence divergence from each other similar to that between the major genotypes of HCV. However, each varies in host range (cow, sheep, or pig), method of transmission, and type of disease. These comparisons provide at least the basis for supposing that differences might exist between HCV genotypes.

CLINICAL UTILITY OF GENOTYPING

A large number of clinical investigations have documented severe and progressive liver disease on infection with each of the well-characterized genotypes (types 1a, 1b, 2a, 2b, 3a, and 4a, see below), so there is little evidence so far for variants of HCV that are completely nonpathogenic or perhaps nonhepatotropic. However, possible variation in the rate of disease progression, differences between genotypes in routes and frequency of person-to-person transmission, or in the probability of achieving a sustained response to antiviral treatment would indicate the potential utility for the identification of the infecting genotype in certain clinical situations. For example, if there were consistent differences between genotypes in response to interferon treatment, genotyping might play an important role in patient selection and in calculating the most effective duration and dose of interferon treatment to achieve a long-term response. The following sections review progress to date of studies investigating biological heterogeneity of HCV and describe what further information still remains to be determined.

How May Genotypes Be Identified? Although determination of the nucleotide sequences is the most reliable method for identifying different genotypes of HCV, this is not practical for large cohort studies. Many of the published methods for "genotyping" are based on amplification of viral sequences in clinical specimens, either using type-specific primers that selectively amplify different genotypes or using analysis of the PCR product by hybridization with genotype-specific probes or by restriction fragment length polymorphism (RFLP) (Table 4). The assays have different strengths and weaknesses. For example, methods based on amplification and analysis of 5′NCR sequences have advantages of sensitivity because this region is highly conserved and can be more frequently amplified from HCV-infected individuals than other parts of the genome. However, relatively few nucleotide differences are found between different genotypes. Although it is possible to reliably differentiate six major genotypes using RFLP or type-specific probes, the identification of subtypes is a matter of serendipity. For examples, type 2a and 2b consistently differ at position −124, allowing them to differentiated using the restriction enzyme *Scr*FI[22] or by probes 10-13 in the InnoLipa.[23] However, sequences of type 2c are indistinguishable from some of those of type 2a. A more important practical problem is the separate identification of types 1a and 1b. Although type 1b sequences generally have a "G" at position −99 and type 1a has an "A," exceptions exist. Approximately 5% of type 1a and 1b sequences have the wrong nucleotide at this position, and therefore, would be misidentified. Furthermore, as with type 2, sequences of other subtypes of type 1 cannot be differentiated. For example, the variant from Indonesia[24] described as type "1c" would appear as a type 1a using current assays.

In contrast, typing methods based on coding regions can reliably identify subtypes as well as major genotypes because the degree of sequence divergence is much greater (Fig. 1). However, it is generally more difficult to amplify sequences in coding regions of the genome, because of lack of conservation in the primer-binding sites, resulting in assays that in some cases are less reliable than those based on amplification of the 5′NCR. The use of type-specific primers is a technically elegant and simple method to identify genotypes; a currently widely used assay produces different length products from each genotype that can be resolved on an agarose gel.[63,64] However, the assay would become increasingly complex if it were extended to detect all of the recently described genotypes.

TABLE 4. Comparison of Typing Methods for HCV

Method	Region	Genotypes*	Reference No.
PCR-based methods			
Type-specific primers	Core	1a, 1b, 2a, 2b, 3a	63, 64
	NS-5	1a, 1b, 2a, 2b, 3b	61
RFLP	5′NCR	1†, 2a, 2b	21
	5′NCR	1†, 2a, 2b, 3, 4, 5, 6	22, 41
Type-specific probes	NS-5	1a, 1b, 2a, 2b	15
	5′NCR	1†, 2a, 2b, 3, 4, 5	23
Serological methods			
Peptide enzyme-linked immunosorbent assay	NS-4	1, 2, 3, 4, 5, 6	65‡
	Core	1, 2	103
Recombinant protein enzyme-linked immunosorbent assay	NS-4	1, 2	66

* Genotypes identified by assay.

† Sequences of type 1a and 1b may be identical in the 5′NCR, so in principal these two subtypes cannot be separately identified by any method based upon this region (see text).

‡ Now modified by addition of peptides from genotypes 4, 5, and 6 (Bhattachejee, et al, Submitted).

Serological typing methods have advantages over PCR-based methods in terms of speed and simplicity of sample preparation and in the use of simple equipment that would be found in any diagnostic virology laboratory. By careful optimization of reagents, such assays may show high sensitivity and reproducibility. For example, type-specific antibody to NS-4 peptides can be detected in approximately 89% of hepatitis C–infected patients and blood donors (Bhattachejee, et al, Submitted). Furthermore, the assays can be readily extended to detect new genotypes. Our NS-4–based assay reliably identifies type-specific antibody to six major genotypes (Table 4; Bhattachejee, et al, Submitted). However, the antigenic similarity between subtypes currently precludes the separate identification of type 1a from 1b and of 2a from 2b using the NS-4 peptides alone, although it is possible that the greater sequence differences in the envelope proteins may be exploited for this purpose in the future.

A crucial assumption of any PCR-based or serological genotyping method is that the region analyzed (5′NCR, core, NS-4, NS-5) is representative of the genome as a whole. The whole principle of "genotyping" would break down if recombination between HCV genotypes occurred during replication and led to the formation of hybrid viruses that contained segments of different genotypes in different parts of the genome. Therefore, it is particularly important to compare different assays on as wide a range of samples as possible. We have previously found almost perfect agreement between the results of serotyping 137 blood donor samples with NS-4 peptides with typing by RFLP in the 5′NCR.[65] In a larger collaborative study, remarkably consistent results were obtained from genotyping a total of 139 samples from American hepatitis C patients using type-specific primers in NS-5,[61] RFLP of the 5′NCR,[41] serotyping using NS-4 proteins[66] or peptides,[65] and using direct sequence comparison of part of NS-5.[46] These findings encourage the belief that genotyping is a valid procedure despite the theoretical objections (such as recombination), and show that many assays can reliably identify infection with at least those HCV genotypes likely to be encountered in clinical practice (1a, 1b, 2a, 2b, and 3a).

What Genotypes Should Be Compared? Although the list of genotypes is becoming ever longer as more regions of the world are examined, most countries show a relatively restricted range of variants in the population. For example, current surveys indicate that genotypes 1a, 1b, 2a, 2b, and 3a account for almost all hepatitis C patients (and HCV-infected blood donors) in many countries in Western Europe and Canada and the United States, whereas types 1b, 2a, and 2b are found in Japan and elsewhere in the Far East (see above). The adaptation of typing assays to detect only those genotypes that are likely to be encountered is one way in which typing could be simplified for clinical use.

At present it is not known whether biological differences exist primarily between the major genotypes of HCV or whether it would be important to also differentiate between subtypes (e.g., type 1a and 1b). If subtypes behaved similarly, then the use of assays to distinguish them would be unnecessary for clinical assessment. Until this issue is resolved, it would be difficult to make specific recommendations for the choice of genotyping assay.

FACTORS INFLUENCING DISEASE PROGRESSION AND RESPONSE TO INTERFERON

Several clinical studies have catalogued a variety of factors (including genotype) that correlate with severity of liver disease and which show predictive value for response to antiviral treatment. Factors that have been frequently shown to influence response to interferon treatment include age and duration of infection,[67-69] presence of cirrhosis before treatment,[68-74] genotype,[44,54,67,68,72,74-80] and pretreatment level of circulating viral RNA in plasma.[68,74,76,79-81]

A consistent finding reported by several different groups using a variety of typing assays has been the greatly increased rate of long-term response found on treating patients infected with genotypes 2a, 2b, and 3a compared with type 1b.[44,54,67,68,72,74,75,77-80] For example, in our recent study we found long-term normalization of alanine transaminase levels (more than 12 months) was infrequently achieved in those infected with type 1 variants (29%), compared with 52% with type 2 and 74% for type 3.[78] These findings have since been repeated in a much larger collaborative study of 610 patients from several centers in Europe (Dusheiko GM, et al, Submitted).

One of the major problems with comparisons of rates of disease development and interferon response with genotype is controlling for other factors that might also alter the outcome. For example, the prevalence of different genotypes may change with age; in Europe, type 3a is more common in young individuals than 1b, which accounts for most infections in those 50 years of age or older. Thus, one might plausibly attribute the observed greater response rate to interferon in those infected with type 3a to the relatively younger age, shorter duration of infection, or absence of cirrhosis compared with those individuals infected with type 1b. Therefore, it is essential that large enough groups of patients are studied to perform multivariate analysis to separately analyse the contribution of different possible influences on the outcome.

Where this has been performed, it has been reported that the better response found in those infected with genotype 2 is independent of liver cirrhosis, although this emerges as a separate independent risk factor. In a more recent study, infection with type 1b, presence of cirrhosis, and high pretreatment virus load were each independently associated with a reduced chance of response (relative risks of 16, 5, and 4, respectively).[74] In our recent European study of 610 patients, we have found that absence of cirrhosis and infection with genotypes 2 or 3 are each independent predictive factors for response.

The mechanism by which different genotypes might differ in responsiveness to treatment remains obscure. For treatments such as interferon it is not even clear

whether the effect of the drug is directly antiviral or whether the inhibition of virus replication is secondary to increased expression of major histocompatibility complex class I antigens on the surface of hepatocytes and greater cytotoxic T-cell activity against virally infected cells. Elucidating the mechanism of action of interferon and whether there are virological differences between genotypes in sensitivity to antivirals awaits a cell culture model for HCV infection.

Among individuals infected with the same genotype, there is close correlation between the level of circulating viral sequences and the response to interferon.[74,76,79,81] Whether the previously reported differences in responsiveness of genotypes can be entirely accounted for by differences in circulating virus load remains unclear; if this hypothesis were true, then it would follow that virus loads in the genotypes that are responsive to interferon would be consistently lower than those found in those infected with genotype 1b. Very recent evidence suggests that this may not be the case; for example, type 2– or type 3–infected patients responded well to interferon yet levels of replicating virus were reported to be similar to those infected with other genotypes[74] (Brechot C, et al, Trepo et al, VI International Symposium on Viral Hepatitis, Madrid, Spain, February 3-5, 1994). Because of this uncertainty, it is not currently clear whether pretreatment assessment of patients should be performed for virus load, for genotype, or for both. The recent finding that genotype and load are independent factors influencing response[74] suggests a value for both types of assay in clinical practice.

IMMUNE ESCAPE?

In contrast to the highly restricted sequence diversity of the 5′NCR and adjacent core region, the two putative envelope genes are highly divergent between different variants of HCV[82-85] (Fig. 1), and show a 3 to 4 times higher rate of sequence change with time in persistently infected individuals.[86] Because these proteins are likely to lie on the outside of the virus, they would be the principal targets of the humoral immune response to HCV elicited on infection. An attractive theory that explains both the high degree of envelope sequence variability and the persistent nature of HCV infection is that changes in the E1 and E2 genes alter the antigenicity of the virus to allow "immune escape" from neutralizing antibodies.[83] In this model, continuing virus replication is a race between diversification of HCV and the efforts of the host immune system to respond to changes by developing neutralizing antibody with an ever greater range of reactivity. Supporting this model is the observation that much of the variability in the E1 and E2 genes is concentrated in discrete "hypervariable" regions.[83-85] By analogy with influenza virus hemagglutinin, it has been suggested that the particularly variable region at the 5′ end of the E2 gene encodes sequences surrounding the part of the protein involved in virus/cellular receptor interactions and that become a target for neutralizing antibody.

Experimental confirmation that immune pressure drives the sequence variability of the envelope gene has proved to be notoriously difficult to obtain, even for other viruses, notably human immunodeficiency virus type 1, for which *in vitro* neutralization assays have been developed. One of the main testable predictions from the theory is the appearance of novel variants of HCV with changes in the E1 and E2 genes that are antigenically distinct from previous variants. In many cases, the *in vivo* appearance of variants with different sequences in the hypervariable region is followed by the development of antibodies that specifically recognize the new variants.[83,87-89] However, this situation would be expected regardless of the selective forces involved and does not prove that immunologic recognition forced the appearance of putative "escape" mutants. For example, it is possible that individuals may be infected with different variants at the time of transmission and that may reappear as different foci of infection become active or as infection spreads to different tissues. Alternatively, sequence change in the envelope region may occur more rapidly because there are fewer restraints on the amino acid sequence (see above). Therefore, the high degree of variability could be attributed to more rapid random drift.

There are other observations that argue directly against the hypothesis of immune escape. First, envelope sequences obtained sequentially from persistently infected individuals sometimes show no significant change,[90-92] whereas in others, variants coexist with antibodies that recognize the corresponding hypervariable region peptides.[87] Some hypervariable region variants carry mutations that increase their recognition by antibodies present at the time of their detection in plasma.[89] Contrary to what would be expected, antibody to envelope proteins can be found at similar or greater frequency in individuals who remain chronic carriers of HCV compared with those who clear the virus infection and become nonviremic.[93,94]

The weak or absent humoral response to envelope proteins is perhaps related to their extensive glycosylation *in vivo* that may serve to shield amino acid epitopes from antibody. It may be that cytotoxic T cell reactivity plays a more important role in virus clearance,[95,96] in which case amino acid sequence change anywhere in the genome would contribute equally to immune escape because viral proteins are processed before presentation to lymphocytes. However, a lack of any effective immune response to HCV can be inferred from the chimpanzee model of infection, where it has been repeatedly shown that experimentally infected animals can be easily re-infected with the same strain of HCV.[97,98] Similarly, multiple episodes of hepatitis have been observed in thalassemic children repeatedly exposed to HCV.[99] Thus, infection in itself may not always induce protective immunity from rechallenge, making it less likely that the weak antibody responses induced in those persistently infected with HCV would play any role in shaping viral sequence change. Finally, virus persistence may be achieved by quite different mechanisms. For example, it has been observed that

HCV associates with lipoproteins in the circulation,[8,100] and, therefore, may persist because it is shielded from circulating antibody and other immune effector mechanisms. The association with lipid would also provide a route by which HCV might gain entry into the liver because lipid micelles are specifically taken up by hepatocytes as part of normal cellular metabolism.

In summary, although there is circumstantial evidence to support the theory of immune escape, a number of observations will need to be taken into account if it is to be proposed as a plausible model of virus persistence. Many of the current uncertainties may be resolved on the development of a satisfactory *in vitro* neutralization assay for HCV to enable the effect of amino acid changes in the envelope gene to be directly investigated. We also need to know more about the relative importance of humoral and cell-mediated immunity to HCV, and to discover which is the more important in virus clearance and protection from re-infection.

In this review I have attempted to summarize and discuss current findings in a highly active area of research. Answers to many of the major remaining questions concerning HCV variability will undoubtedly be provided soon by experiments using *in vitro* culture of HCV. Similarly, the origins, spread, and current genotype distributions will become more clearly understood when more is known about routes of virus transmission in different geographic areas.

Acknowledgment: I am very grateful to colleagues in Edinburgh for contributing to the preparation of this manuscript, in particular, I would like to thank Fiona Davidson for making available unpublished data on genotype distributions and Eddie Holmes and Donald Smith for discussion and careful reading of the manuscript.

REFERENCES

1. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA derived from a blood-borne non-A, non-B hepatitis genome. Science 1989;244:359-362.
2. Kuo G, Choo QL, Alter HJ, Gitnick GL, Redeker AG, Purcell RH, Miyamura T, et al. An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. Science 1989;244:362-364.
3. Reyes GR, Purdy M, Kim JS, Luk KC, Young LM, Fry KE, Bradley DW. Isolation of cDNA from the virus responsible for enterically transmitted non-A, non-B hepatitis. Science 1990;247:1335-1339.
4. Choo QL, Richman KH, Han JH, Berger K, Lee C, Dong C, Gallegos C, et al. Genetic organization and diversity of the hepatitis C virus. Proc Natl Acad Sci U S A 1991;88:2451-2455.
5. Han JH, Shyamala V, Richman KH, Brauer MJ, Irvine B, Urdea MS, Tekamp Olson P, et al. Characterization of the terminal regions of hepatitis C viral RNA: identification of conserved sequences in the 5′ untranslated region and poly(A) tails at the 3′ end. Proc Natl Acad Sci U S A 1991;88:1711-1715.
6. Collett MS, Anderson DK, Retzel E. Comparisons of the pestivirus bovine viral diarrhoea virus with members of the flaviviridae. J Gen Virol 1988;69:2637-2643.
7. Chambers TJ, Hahn CS, Galler R, Rice CM. Flavivirus genome organisation, expression and replication. Ann Rev Microbiol 1990;44:649-688.
8. Miyamoto H, Okamoto H, Sato K, Tanaka T, Mishiro S. Extraordinarily low density of hepatitis C virus estimated by sucrose density gradient centrifugation and the polymerase chain reaction. J Gen Virol 1992;73:715-718.
9. 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 U S A 1990;87:9524-9528.
10. Takamizawa A, Mori C, Fuke I, Manabe S, Murakami S, Fujita J, Onishi E, et al. Structure and organization of the hepatitis C virus genome isolated from human carriers. J Virol 1991;65:1105-1113.
11. Okamoto H, Kurai K, Okada S, Yamamoto K, Lizuka H, Tanaka T, Fukuda S, et al. Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct genotypes. Virology 1992;188:331-341.
12. Okamoto H, Okada S, Sugiyama Y, Kurai K, Iizuka H, Machida A, Miyakawa Y, et al. Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions. J Gen Virol 1991;72:2697-2704.
13. Mori S, Kato N, Yagyu A, Tanaka T, Ikeda Y, Petchclai B, Chiewsilp P, et al. A new type of hepatitis C virus in patients in Thailand. Biochem Biophys Res Commun 1992;183:334-342.
14. Chan S-W, McOmish F, Holmes EC, Dow B, Peutherer JF, Follett E, Yap PL, et al. Analysis of a new hepatitis C virus type and its phylogenetic relationship to existing variants. J Gen Virol 1992;73:1131-1141.
15. Enomoto N, Takada A, Nakao T, Date T. There are two major types of hepatitis C virus in Japan. Biochem Biophys Res Commun 1990;170:1021-1025.
16. Simmonds P, Holmes EC, Cha TA, Chan S-W, McOmish F, Irvine B, Beall E, et al. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. J Gen Virol 1993;74:2391-2399.
17. Simmonds P, McOmish F, Yap PL, Chan SW, Lin CK, Dusheiko G, Saeed AA, et al. Sequence variability in the 5′ non coding region of hepatitis C virus: identification of a new virus type and restrictions on sequence diversity. J Gen Virol 1993;74:661-668.
18. Bukh J, Purcell RH, Miller RH. At least 12 genotypes of hepatitis C virus predicted by sequence analysis of the putative E1 gene of isolates collected worldwide. Proc Natl Acad Sci U S A 1993;90:8234-8238.
19. Simmonds P, Smith DB, McOmish F, Yap PL, Kolberg J, Urdea MS, Holmes EC. Identification of genotypes of hepatitis C virus by sequence comparisons in the core, E1 and NS-5 regions. J Gen Virol 1994;75:1053-1061.
20. Holmes EC, Simmonds P, Cha T-A, Chan S-W, McOmish F, Irvine B, Beall E, et al. Derivation of a rational nomenclature for hepatitis C virus by phylogenetic analysis of the NS-5 region. In: Nishioka K, Suzuki H, Mishiro S, Oda T, eds. Viral hepatitis and liver disease. Tokyo: Springer-Verlag, 1994:57-62.
21. Nakao T, Enomoto N, Takada N, Takada A, Date T. Typing of hepatitis C virus (HCV) genomes by restriction fragment length polymorphisms. J Gen Virol 1991;72:2105-2112.
22. McOmish F, Chan S-W, Dow BC, Gillon J, Frame WD, Crawford RJ, Yap PL, et al. Detection of three types of hepatitis C virus in blood donors: Investigation of type-specific differences in serological reactivity and rate of alanine aminotransferase abnormalities. Transfusion 1993;33:7-13.
23. Stuyver L, Rossau R, Wyseur A, Duhamel M, Vanderborght B, Van Heuverswyn H, Maertens G. Typing of hepatitis C virus isolates and characterisation of new subtypes using a line probe assay. J Gen Virol 1993;74:1093-1102.
24. Okamoto H, Kojima M, Sakamoto M, Iizuka H, Hadiwandowo S, Suwignyo S, Miyakawa Y, et al. The entire nucleotide sequence and classification of a hepatitis C virus isolate of a novel genotype from an Indonesian patient with chronic liver disease. J Gen Virol 1994;75:629-635.
25. Simmonds P, Alberti A, Alter HJ, Bonino F, Bradley DW, Brechot C, Brouwer JT, et al. A proposed system for the nomenclature of hepatitis C viral genotypes. HEPATOLOGY 1994;19(suppl):1321-1324.
26. Ogata N, Alter HJ, Miller RH, Purcell RH. Nucleotide sequence

and mutation rate of the H strain of hepatitis C virus. Proc Natl Acad Sci U S A 1991;88:3392-3396.
27. Abe K, Inchauspe G, Fujisawa K. Genomic characterisation and mutation rate of hepatitis C virus isolated from a patient who contracted hepatitis during an epidemic on non-A, non-B hepatitis in Japan. J Gen Virol 1992;73:2725-2729.
28. Okamoto H, Kojima M, Okada S-I, Yoshizawa H, Iizuka H, Tanaka T, Muchmore EE, et al. Genetic drift of hepatitis C virus during an 8.2 year infection in a chimpanzee: variability and stability. Virology 1992;190:894-899.
29. Cuypers HTM, Winkel IN, van der Poel CL, Reesink HW, Lelie PN, Houghton M, Weiner A. Analysis of genomic variability of hepatitis C virus. J Hepatol 1991;13:S15-S19.
30. Li W-H, Graur D. Fundamentals of molecular evolution. New York: Sinaur Associates, Inc, 1991.
31. Nei M. Molecular evolutionary genetics. Sunderland, MA: Columbia University Press, 1987.
32. Tsukiyama Kohara K, Iizuka N, Kohara M, Nomoto A. Internal ribosome entry site within hepatitis C virus RNA. J Virol 1992;66:1476-1483.
33. Brown EA, Zhang H, Ping H-L, Lemon SM. Secondary structure of the 5′ nontranslated region of hepatitis C virus and pestivirus genomic RNAs. Nucleic Acids Res 1992;20:5041-5045.
34. Wang CY, Sarnow P, Siddiqui A. Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. J Virol 1993;67:3338-3344.
35. Ina Y, Mizokami M, Ohba K, Gojobori T. Reduction of synonymous substitutions in the core protein gene of hepatitis C virus. J Mol Evol 1994;38:50-56.
36. Chan S-W, Simmonds P, McOmish F, Yap P-L, Mitchell R, Dow B, Follett E. Serological reactivity of blood donors infected with three different types of hepatitis C virus. Lancet 1991;338: 1391.
37. Li J, Tong S, Vitvitski L, Lepot D, Trepo C. Two French genotypes of hepatitis C virus: homology of the predominant genotype with the prototype American strain. Gene 1991;105:167-172.
38. Lee C, Cheng C, Wang J, Lumeng L. Identification of hepatitis C viruses with a nonconserved sequence of the 5′ untranslated region. J Clin Microbiol 1992;30:1602-1604.
39. Bukh J, Purcell RH, Miller RH. Sequence analysis of the 5′ noncoding region of hepatitis C virus. Proc Natl Acad Sci U S A 1992;89:4942-4946.
40. Cha TA, Beall E, Irvine B, Kolberg J, Chien D, Kuo G, Urdea MS. At least five related, but distinct, hepatitis C viral genotypes exist. Proc Natl Acad Sci U S A 1992;89:7144-7148.
41. McOmish F, Yap PL, Dow BC, Follett EAC, Seed C, Keller AJ, Cobain TJ, et al. Geographical distribution of different hepatitis C virus genotypes in blood donors: an international collaborative survey. J Clin Microbiol 1994;32:884-892.
42. Kleter GEM, Van Doorn LJ, Brouwer JT, Schalm SW, Heijtink RA, Quint WGV. Sequence analysis of the 5′ untranslated region in isolates of at least four genotypes of hepatitis C virus in the Netherlands. J Clin Microbiol 1994;33:306-310.
43. Murphy D, Willems B, Delage G. Use of the 5′ non-coding region for genotyping hepatitis C virus. J Infect Dis 1994;169:473-475.
44. Tisminetzky SG, Gerotto M, Pontisso P, Chemello L, Ruvoletto MG, Baralle F, Alberti A. Genotypes of hepatitis C virus in italian patients with chronic hepatitis C. Int Hepatol Commun 1994;2:105-112.
45. Qu D, Hantz O, Gouy M, Vitvitski L, Li JS, Berby F, Tong SP, et al. Heterogeneity of hepatitis C virus genotypes in France. J Gen Virol 1994;75:1063-1070.
46. Lau JYN, Mizokami M, Kolberg J, Davis GL, Prescott LE, Ohno T, Perrillo RP, et al. Application of hepatitis C virus genotyping systems in United States patients with chronic hepatitis C. J Infect Dis 1995 (in press).
47. Saeed AA, al Admawi AM, al Rasheed A, Fairclough D, Bacchus R, Ring C, Garson J. Hepatitis C virus infection in Egyptian volunteer blood donors in Riyadh. Lancet 1991;338:459-460.
48. Darwish MA, Raouf TA, Rushdy P, Constantine NT, Rao MR, Edelman R. Risk factors associated with a high seroprevalence of hepatitis C virus infection in Egyptian blood donors. Am J Trop Med Hyg 1993;49:440-447.
49. Kamel MA, Ghaffar YA, Wasef MA, Wright M, Clark LC, Miller FD. High HCV prevalence in egyptian blood donors. Lancet 1992;340:427.
50. Stuyver L, Rossau R, Wyseur A, Duhamel M, Vanderborght B, Van Heuverswyn H, Maertens G. Typing of hepatitis C virus isolates and characterization of new subtypes using a line probe assay. J Gen Virol 1993;74:1093-1102.
51. Delaporte E, Thiers V, Dazza MC, Romeo R, Mlikacabanne N, Aptel I, Schrijvers D, et al. High level of hepatitis C endemicity in Gabon, Equatorial Africa. Trans R Soc Trop Med Hyg 1993;87:636-637.
52. Ohno T, Mizokami M, Tibbs CJ, Ohba K, Suzuki K, Wu RR, Nouriaria KT, et al. New genotype of hepatitis C virus in South-Africa. J Med Virol 1994;42:409-413.
53. Kato N, Ootsuyama Y, Ohkoshi S, Nakazawa T, Mori S, Hijikata M, Shimotohno K. Distribution of plural HCV types in Japan. Biochem Biophys Res Commun 1991;181:279-285.
54. Takada N, Takase S, Enomoto N, Takada A, Date T. Clinical backgrounds of the patients having different types of hepatitis C virus genomes. J Hepatol 1992;14:35-40.
55. Liu K, Hu Z, Li H, Prince AM, Inchauspe G. Genomic typing of hepatitis C viruses present in China. Gene 1992;114:245-250.
56. Wang Y, Okamoto H, Mishiro S. HCV genotypes in China. Lancet 1992;339:1168.
57. Takada N, Takase S, Takada A, Date T. HCV genotypes in different countries. Lancet 1992;339:808.
58. Takada N, Takase S, Takada A, Date T. Differences in the hepatitis C virus genotypes in different countries. J Hepatol 1993;17:277-283.
59. Wang Y, Okamoto H, Tsuda F, Nagayama R, Tao Q, Mishiro S. Prevalence, genotypes, and an isolate (HC-C2) of hepatitis C virus in Chinese patients with liver disease. J Med Virol 1993;40:254-260.
60. Kinoshita T, Miyake K, Okamoto H, Mishiro S. Imported hepatitis C virus genotypes in Japanese hemophiliacs. J Infect Dis 1993;168:249-250.
61. Chayama K, Tsubota A, Arase Y, Saitoh S, Koida I, Ikeda K, Matsumoto T, et al. Genotypic subtyping of hepatitis C virus. J Gastroenterol Hepatol 1993;8:150-156.
62. Tokita H, Shrestha SM, Okamoto H, Sakamoto M, Hirokita M, Iizuka H, Shrestha S, et al. Hepatitis C virus variants from Nepal with novel genotypes and their classification into the third major group. J Gen Virol 1994;75:931-936.
63. Okamoto H, Sugiyama Y, Okada S, Kurai K, Akahane Y, Sugai Y, Tanaka T, et al. Typing hepatitis C virus by polymerase chain reaction with type-specific primers: application to clinical surveys and tracing infectious sources. J Gen Virol 1992;73: 673-679.
64. Okamoto H, Tokita H, Sakamoto M, Horikita M, Kojima M, Iizuka H, Mishiro S. Characterization of the genomic sequence of type V (or 3a) hepatitis C virus isolates and PCR primers for specific detection. J Gen Virol 1993;74:2385-2390.
65. Simmonds P, Rose KA, Graham S, Chan SW, McOmish F, Dow BC, Follett EAC, et al. Mapping of serotype-specific, immunodominant epitopes in the NS-4 region of hepatitis C virus (HCV)—use of type-specific peptides to serologically differentiate infections with HCV type 1, type 2, and type 3. J Clin Microbiol 1993;31:1493-1503.
66. Tanaka T, Tsukiyama-Kohara K, Yamaguchi K, Yagi S, Tanaka S, Hasegawa A, Ohta Y, et al. Significance of specific antibody assay for genotyping of hepatitis C virus. HEPATOLOGY 1994;19:1347-1353.
67. Yoshioka K, Kakumu S, Wakita T, Ishikawa T, Itoh Y, Takayanagi M, Higashi Y, et al. Detection of hepatitis C virus by polymerase chain reaction and response to interferon-alpha therapy: relationship to genotypes of hepatitis C virus. HEPATOLOGY 1992;16:293-299.
68. Hino K, Sainokami S, Shimoda K, Iino S, Wang Y, Okamoto H, Miyakawa Y, et al. Genotypes and titers of hepatitis C virus for predicting response to interferon in patients with chronic hepatitis C. J Med Virol 1994;42:299-305.
69. Piccinino F, Felaco FM, Aprea L, Messina V, Pasquale G, Sagnelli E. Non responders to interferon therapy among chronic

hepatitis patients infected with hepatitis C virus. Arch Virol 1993;8(suppl):257-263.
70. Uchida T, Taira M, Shikata T, Moriyama M, Tanaka N, Okubo H, Arakawa Y. Histological difference between complete responders and non-responders to interferon therapy of the livers of patients with chronic hepatitis C. Acta Pathol Jpn 1993;43:230-236.
71. Perez R, Pravia R, Linares A, Rodriguez M, Lombrana JLS, Suarez A, Riestra S, et al. Response related factors in recombinant interferon-alfa-2b treatment of chronic hepatitis C. Gut 1993;34(suppl):S139-S140.
72. Tsubota A, Chayama K, Arase Y, Koida I, Saitoh S, Ikeda K, Iwasaki S, et al. Factors useful in predicting the response to interferon therapy in chronic hepatitis C. J Gastroenterol Hepatol 1993;8:535-539.
73. Camps J, Crisostomo S, Garciagranero M, Riezuboj JI, Civeira MP, Prieto J. Prediction of the response of chronic hepatitis C to interferon-alfa—a statistical analysis of pretreatment variables. Gut 1993;34:1714-1717.
74. Tsubota A, Chayama K, Ikeda K, Yasuji A, Koida I, Saitoh S, Hashimoto M, et al. Factors predictive of response to interferon-alpha therapy in hepatitis C virus infection. HEPATOLOGY 1994;19:1088-1094.
75. Kanai K, Kako M, Okamoto H. HCV genotypes in chronic hepatitis C and response to interferon. Lancet 1992;339:1543.
76. Kobayashi Y, Watanabe S, Konishi M, Yokoi M, Kakehashi R, Kaito M, Kondo M, et al. Quantitation and typing of serum hepatitis C virus RNA in patients with chronic hepatitis C treated with interferon-beta. HEPATOLOGY 1993;18:1319-1325.
77. Dusheiko G, Schmilovitzweiss H, Brown D, McOmish F, Yap PL, Sherlock S, McIntyre N, et al. Hepatitis C virus genotypes—an investigation of type-specific differences in geographic origin and disease. HEPATOLOGY 1994;19:13-18.
78. Chemello L, Alberti A, Rose K, Simmonds P. Hepatitis C serotype and response to interferon therapy. N Engl J Med 1994;330:143.
79. Yamada G, Takahashi M, Miyamoto R, Tsuji T, Yoshizawa H, Okamoto H. Prediction of interferon effect in chronic hepatitis C by both quantification and genotyping of HCV-RNA. Dig Dis Sci 1994;39:441-444.
80. Chayama K, Arase Y, Koida I, Tsubota A, Saitoh S, Ikeda K, Matsumoto T, et al. Antiviral effect of lymphoblastoid interferon-alpha on hepatitis C virus in patients with chronic hepatitis type C. J Gastroenterol Hepatol 1994;9:128-133.
81. Lau JYN, Davis GL, Kniffen J, Qian KP, Urdea MS, Chan CS, Mizokami M, et al. Significance of serum hepatitis C virus RNA levels in chronic hepatitis C. Lancet 1993;341:1501-1504.
82. Kato N, Ootsuyama Y, Ohkoshi S, Nakazawa T, Sekiya H, Hijikata M, Shimotohno K. Characterization of hypervariable regions in the putative envelope protein of hepatitis C virus. Biochem Biophys Res Commun 1992;189:119-127.
83. Weiner AJ, Geysen HM, Christopherson C, Hall JE, Mason TJ, Saracco G, Bonino F, et al. Evidence for immune selection of hepatitis C virus (HCV) putative envelope glycoprotein variants: potential role in chronic HCV infections. Proc Natl Acad Sci U S A 1992;89:3468-3472.
84. Kato N, Ootsuyama Y, Tanaka T, Nakagawa M, Nakazawa T, Muraiso K, Ohkoshi S, et al. Marked sequence diversity in the putative envelope proteins of hepatitis C viruses. Virus Res 1992;22:107-123.
85. Hijikawa M, Kato N, Ootsuyama Y, Nakagawa M, Ohkoshi S, Shimotohno K. Hypervariable regions in the putative glycoprotein of hepatitis C virus. Biochem Biophys Res Commun 1991;175:220-228.
86. Okamoto H, Kojima M, Okada SI, Yoshizawa H, Iizuka H, Tanaka T, Muchmore EE, et al. Genetic drift of hepatitis-c virus during an 8.2-year infection in a chimpanzee—variability and stability. Virology 1992;190:894-899.
87. Kato N, Sekiya H, Ootsuyama Y, Nakazawa T, Hijikata M, Ohkoshi S, Shimotohno K. Humoral immune response to hypervariable region-1 of the putative envelope glycoprotein (gp70) of hepatitis C virus. J Virol 1993;67:3923-3930.
88. Lesniewski RR, Boardway KM, Casey JM, Desai SM, Devare SG, Leung TK, Mushahwar IK. Hypervariable 5'-terminus of hepatitis C virus E2/NS1 encodes antigenically distinct variants. J Med Virol 1993;40:150-156.
89. Taniguchi S, Okamoto H, Sakamoto M, Kojima M, Tsuda F, Tanaka T, Munekata E, et al. A structurally flexible and antigenically variable n-terminal domain of the hepatitis C virus e2/NS1 protein—implication for an escape from antibody. Virology 1993;195:297-301.
90. Okada S-I, Akahane Y, Suzuki H, Okamoto H, Mishiro S. The degree of variability in the amino terminal region of the E2/NS1 protein of hepatitis C virus correlates with responsiveness to interferon therapy in viraemic patients. HEPATOLOGY 1992;16:619-624.
91. Kumar U, Brown J, Monjardino J, Thomas HC. Sequence variation in the large envelope glycoprotein (E2/NS1) of hepatitis C virus during chronic infection. J Infect Dis 1993;167:726-730.
92. Higashi Y, Kakumu S, Yoshioka K, Wakita T, Mizokami M, Ohba K, Ito Y, et al. Dynamics of genome change in the E2/NS1 region of hepatitis C virus in vivo. Virology 1993;197:659-668.
93. Chien DY, Choo QL, Ralston R, Spaete R, Tong M, Houghton M, Kuo G. Persistence of HCV despite antibodies to both putative envelope glycoproteins. Lancet 1993;342:933.
94. Hada H, Koide N, Hanafusa T, Sakaguchi K, Shinji T, Sasaki S, Oka T, et al. Detection by western blotting of an antibody to the hepatitis C virus E1 envelope protein in sera of patients with chronic liver disease. Acta Med Okayama 1992;46:365-370.
95. Koziel MJ, Dudley D, Wong JT, Dienstag J, Houghton M, Ralston R, Walker BD. Intrahepatic cytotoxic T lymphocytes specific for hepatitis-c virus in persons with chronic hepatitis. J Immunol 1992;149:3339-3344.
96. Koziel MJ, Dudley D, Afdhal N, Choo QL, Houghton M, Ralston R, Walker BD. Hepatitis C virus (HCV)-specific cytotoxic T lymphocytes recognize epitopes in the core and envelope proteins of HCV. J Virol 1993;67:7522-7532.
97. Farci P, Alter HJ, Govindarajan S, Wong DC, Engle R, Lesniewski RR, Mushahwar IK, et al. Lack of protective immunity against reinfection with hepatitis C virus. Science 1992;258:135-140.
98. Prince AM, Brotman B, Huima T, Pascual D, Jaffery M, Inchauspe G. Immunity in hepatitis C infection. J Infect Dis 1992;165:438-443.
99. Lai ME, Mazzoleni AP, Argiolu F, De Virgilis S, Balestriesi A, Purcell RH, Cao A, et al. Hepatitis C virus in multiple episodes of acute hepatitis in polytransfused thalassaemic children. Lancet 1994;343:388-390.
100. Thomssen R, Bonk S, Thiele A. Density heterogeneities of hepatitis C virus in human sera due to the binding of beta-lipoproteins and immunoglobulins. Med Microbiol Immunol 1993;182:329-334.
101. Chan S-W, Holmes EC, McOmish F, Follett E, Yap PL, Simmonds P. Phylogenetic analysis of a new, highly divergent HCV type (type 3): effect of sequence variability on serological responses to infection. Hepatitis C and related viruses [Abstract]. Molecular Virology and Pathogenesis, First Meeting. Venice, Italy, July 1992;Abstract D5:73.
102. Houghton M, Weiner A, Han J, Kuo G, Choo QL. Molecular biology of the hepatitis C viruses: implications for diagnosis, development and control of viral disease. HEPATOLOGY 1991;14:381-388.
103. Machida A, Ohnuma H, Tsuda F, Munekata E, Tanaka T, Akahane Y, Okamoto H, et al. Two distinct subtypes of hepatitis C virus defined by antibodies directed to the putative core protein. HEPATOLOGY 1992;16:886-891.
104. Inchauspe G, Zebedee S, Lee DH, Sugitani M, Nasoff M, Prince AM. Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates. Proc Natl Acad Sci U S A 1991;88:10292-10296.
105. Sakamoto M, Akahane Y, Tsuda F, Tanaka T, Woodfield DG, Okamoto H. The entire nucleotide sequence and characterization of a hepatitis C virus of genotype V/3a. J Gen Virol 1994;75:1761-1768.