Chapter 2

Review of Literature
B. REVIEW OF LITERATURE

In the initial studies on viral hepatitis conducted from 1940 to 1960, investigators identified two distinct forms of viral hepatitis: Infectious hepatitis, which was designated as hepatitis A, and serum hepatitis, designated as hepatitis B. The existence of a third type of hepatitis was convincingly proved by Stephen Feinstone et al (1975), who conclusively demonstrated that the residual cases of hepatitis B negative post-transfusion hepatitis were not due to hepatitis A. This disease was thus named “non-A, non-B hepatitis” (NANBH). The breakthrough in NANBH research occurred when advances in modern molecular technology, allowed Michael Houghton et al (1987) to identify and characterize the causative agent of hepatitis C using molecular tools. This also represented a changing paradigm in the field of infectious diseases, i.e., the identification of an important human pathogen without the ability to grow, visualize, or detect the organism. Early studies on its etiology described a blood-borne agent that could be transmitted to chimpanzees and caused ultrastructural alteration within infected hepatocytes (Yanagi et al, 1997). The agent was rendered inactive by treatment with organic solvents and could pass through an 80 nm filter. These observations led to the conclusion that the infectious agent was a small, enveloped virus, but there was no immunological or molecular evidence to support this conclusion. The term hepatitis C virus was first adopted in 1989 following the identification of an RNA viral genome in a random-primed cDNA library.
derived from a human plasma sample containing the putative NANBH agent (Choo et al 1991).

This agent was then shown to be serially transmitted in chimpanzees (Tabor et al, 1979). Further work showed that this agent was sensitive to chloroform suggesting the presence of an envelope (Bradly et al, 1983). The size of the putative agent was estimated to be 30-60nm, with a 37 nm inner core (Jacob et al, 1990).

Considering the hitherto observed facts about this putative NANBH agent, hepatitis C virus was classified as a member of the genus Hepacivirus within the family Flaviviridae. The seventh report of the international committee on taxonomy of viruses (ICTV) also confirmed this classification scheme (Heinz et al, 2000).

Viral Genome organization:

The HCV genome is a single stranded RNA molecule of positive polarity, approximately 9.6 Kb in length. It contains a single open reading frame of ~ 9 Kb that encodes a polyprotein of ~ 3,000 amino acids (Choo et al, 1991) (Fig.1a). The ORF is flanked at both ends by untranslated regions which are the most conserved regions of the genome.

The 5'UTR has characteristic secondary structures and plays a crucial role in translation initiation. The 5' UTR is a 341 nucleotide sequence closely resembling the 5'UTRs of pestiviruses in terms of length and putative secondary structure (Liang et al, 2000). It functions via an internal ribosomal entry site (IRES) to direct cap independent translation of the genome. This
Fig. 1a: Hepatitis C Virus Genome Organization
region of the genome is highly conserved in terms of nucleotide sequence and secondary structure, although there are a number of genotype specific variations (Liang et al, 2000). This feature has facilitated the development of PCR-based genotyping methods. The IRES was shown to encompass most of the 5' UTR, with the 5' end residing between nucleotides 28 and 69. Deletion of most of the 5' terminal hairpin (stem loop I) increased translational efficiency of the IRES in vitro, suggesting that this structure functions to suppress translation (Liang et al, 2000). Domains II and III of the 5' UTR were essential for its IRES activity. In addition, the regions just downstream of the initiator codon AUG was shown to be essential for efficient translation. The initiator AUG codon is located within the secondary structure stem loop IV (Liang et al, 2000).

The 3' UTR spreads over approximately 30 nucleotides downstream of the 3' terminus of the polyprotein coding region followed by a poly U-C stretch of variable length (Tanaka et al, 1996). The extreme 3' terminus, known as the X region, is predicted to form 3 stem-loop structures. This region is involved in the regulation of translation, probably by interacting directly or indirectly with the IRES element (Ito et al, 1998).

The nascent viral polyprotein is processed by a combination of host and viral proteases into the mature viral proteins. All the structural proteins have hydrophobic C termini. This feature is thought to be important for membrane association and cleavage from the polyprotein by host signal peptidases (Yasui et al, 1998). The viral structural proteins are core or capsid and envelope glycoproteins E1 and E2, which are major components of prototype vaccine studies for HCV. The core protein remains on the
cytoplasmic side of the endoplasmic reticulum and is cleaved from E1 and E2. Within the ER, E1 and E2 are modified by N-linked glycosylation and are thought to form the HCV envelop proteins. The folding and assembly of the envelope proteins within the ER are likely assisted by cellular chaperones (Choukhi et al, 1998).

The core protein is believed to be the main structural component of the viral capsid. It is highly conserved among all HCV strains with some genotype variations. Because of the high prevalence of anticore antibodies in infected individuals, the protein has been used extensively in serologic assays for detecting HCV infection (Liang et al, 2000). In addition to being a structural component of the HCV virion, the core protein has been implicated in a variety of other functions. These include inhibition or stimulation of apoptosis, regulation of cellular or viral promoters, and activation of transcriptional factors. Core protein has also been shown to cooperate with oncogenes to transform primary rat embryo fibroblast in vitro (Ray et al, 1996). Furthermore, the transgenic expression of core protein has been shown to result in hepatic steatosis and subsequent hepatocellular carcinoma in a transgenic mouse model (Moriya et al, 1998).

HCV proteins E1 and E2 are membrane-associated glycolproteins, forming an integral part of the HCV virion envelope (Liang et al, 2000). E1 and E2 are thought to interact to form a heterodimeric noncovalent E1-E2 complex that ultimately constitutes the envelope (Duvet et al, 1998). The E1 and E2 regions of the HCV genome demonstrate the highest mutation rate at the nucleotide level as well as at the predicted amino acid level (Han, 1991). The finding of a rapidly evolving region within one of the envelope proteins
of HCV suggests that this region is under selective pressure by the host immune system. An extraordinarily high rate of nucleotide change that frequently resulted in codon changes was found in hypervariable region 1 of the E2 protein of the HCV genome and consists of 27 amino acids (Ogata et al, 1991). Studies in chimpanzees and in patients with acute and chronic hepatitis C have demonstrated that these infected hosts mount a humoral immune response to epitopes of hypervariable region 1 of the HCV genome (Shimizu et al, 1994; Zibert et al, 1995). The presence of this rapidly changing region may permit a mechanism by which HCV evades host immune surveillance and establishes and maintains persistent infection. Together with the core protein, envelope proteins play an important role in various stages of the viral life cycle, including cell entry, uncoating, and virion assembly (Liang et al, 2000). The function of the p7 protein has not been determined. Evidence from studies with pesti viruses suggests that p7 is probably not necessary for viral assembly (Elbers et al, 1996).

The non-structural proteins of HCV are NS2, NS3, NS4A, NS4B, NS5A and NS5B. Among these, NS2 is a transmembrane polypeptide and has some proteolytic function (Santolini et al, 1995). NS3 is a serine protease. The NS2 protein, together with the N-terminal domain of NS3, forms an NS2-NS3 protease, which is a zinc protease (Hijikata et al, 1993).

The NS3 protein possesses multiple enzyme activities. The N-terminal one-third (180 amino acids) of the protein has serine protease activity. The C-terminal region possesses RNA helicase and nucleotide triphosphatase activities. By virtue of it protease function, NS3 is an attractive target for antiviral drug development (Gallinari et al, 1998).
NS4A acts as a cofactor for the NS3 protease. It may act as a molecular chaperone to facilitate correct folding of the protease and stabilize its conformation. NS4B is crucial for HCV replication complex formation (Tanji et al., 1995).

NS5A is a phosphoprotein in nature, important for viral replication. NS5A is a critical factor in determining the susceptibility of the virus to interferon treatment. The C-terminal portion of NS5A has been shown to function as a transcriptional activator (Kato et al., 1997). NS5B codes for the RNA-dependent RNA polymerase, essential for virus replication. The protein has a molecular mass of 68 kDa and is cleaved from the C-terminus of the polyprotein by the NS3 protease (Hwang et al., 1997)

HCV markers:

HCV seropositivity is defined as the presence of antibodies in peripheral blood, specifically directed to HCV antigens as a result of past or ongoing infection. Antibodies directed to both structural (core or envelope) and nonstructural HCV antigens are present at various levels. HCV genetic variability may explain the apparent heterogeneity of anti HCV humoral responses. This phenomenon can be used for diagnostic purposes because specific antibodies are markers of the infective genotype (Liang et al., 2000).

Diagnosis for HCV:

Sero diagnosis

Multiple protein antigens encoded by the viral RNA seem to produce serologic responses in the host. Serologic responses to four of these protein
antigens are used in diagnostic laboratories for the detection of HCV infection. The first two proteins, termed 5-1-1 (155 bp clone), c100 (larger overlapping clone of 353 bp), and c100-3 (product of fusion of c100 and the gene for human superoxide dismutase, expressed in recombinant yeast) are derived from nonstructural regions of the HCV genome, specifically the NS3 and NS4 regions, and together they form the basis of the first-generation antibody assays (enzyme-linked immunosorbent assay 1 [ELISA-1] and strip immunoassay 1 [SIA-1]) (Feitelson, 2002). In addition to 5-1-1 and c100-3, proteins c33c (recombinant NS3) and c22-3 (recombinant core) are included in the second-generation antibody assays (ELISA-2 and SIA-2). Protein c33c is derived from the NS3 region, which is a nonconserved region of the viral genome. However, protein c22-3 is derived from the highly conserved nucleocapsid (C) region. A recombinant NS5 antigen and epitopes from c200 (NS3 and NS4) have been added to the above four antigens to improve the sensitivity of serologic assays for the detection of HCV antibodies (third-generation antibody assays) (Feitelson, 2002).

Diagram of HCV genome showing positions of recombinant antigens used in Anti-HCV ELISA
**Molecular Diagnosis**

Molecular diagnosis is performed by HCV-specific reverse transcription and polymerase chain reaction with primers from the most conserved regions of HCV genome, i.e., 5’UTR (Okamoto et al, 1992). Along with PCR, additional steps of hybridization are also often employed to increase the specificity of the procedure (Widell et al, 1994).

**Genetic variability:**

Genetic variability of HCV exists at several different levels. Most obvious is the substantial genetic divergence of the main genotypes of HCV which frequently show specific geographical distribution and associations with risk groups for infection. Sequence diversity observed between different strains within the same genotypes, reflects processes of neutral sequence drift over time (Liang et al, 2000).

**Classification into genotypes and subtypes:**

After the complete HCV genome was determined by Choo et al. in 1991, several HCV isolates from different parts of the world were obtained and sequenced (Chen et al, 1992; Delisse et al, 1991; Li et al, 1991; Li et al, 1991a). Comparison of the published sequences of HCV has led to the identification of several distinct types that may differ from each other by as much as 33% over the whole viral genome (Okamoto et al, 1992). Sequence
variability is distributed equally throughout the viral genome, apart from the highly conserved 5' UTR and core regions and the hypervariable envelope (E) region (Okamoto et al, 1994).

HCV is classified on the basis of the similarity of nucleotide sequence into major genetic groups designated genotypes. HCV genotypes are numbered (arabic numerals) in the order of their discovery. The more closely related HCV strains within some types are designated subtypes, which are assigned lowercase letters (in alphabetic order) in the order of their discovery (Tables 1a & 1b). The complex of genetic variants found within an individual isolate is termed as the quasispecies. The quasispecies composition of HCV results from the accumulation of mutations during viral replication in the host (Xavier et al, 1998).

The genomic sequences of different HCV isolates vary by as much as 35%. The degrees of difference in nucleotide sequences among isolates vary from one genomic region to another, with more variability concentrated in regions such as E1 and E2, where as sequences of the core genes and some of the nonstructural genes, such as NS3, are more conserved. The lowest sequence variability between genotypes is found in the 5' UTR, where specific sequences and RNA secondary structures are required for replication and translation (Liang et al, 2000).

Despite the sequence diversity, all genotypes share an identical compliment of collinear genes of similar or identical types. However, in contrast to this general observation there is the marked variation in their
**Table 1a**: Terminology commonly used in studies related to HCV genomic heterogeneity

<table>
<thead>
<tr>
<th>Terminology</th>
<th>Definition</th>
<th>% Nucleotide similarity&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotype</td>
<td>Genetic heterogeneity among different HCV isolates</td>
<td>65.7–68.9</td>
</tr>
<tr>
<td>Subtype</td>
<td>Closely related isolates within each of the major genotypes</td>
<td>76.9–80.1</td>
</tr>
<tr>
<td>Quasispecies</td>
<td>Complex of genetic variants within individual isolates</td>
<td>90.8–99</td>
</tr>
</tbody>
</table>

<sup>a</sup> - Nucleotide similarity refers to the nucleotide sequence identities of the full-length sequences of the HCV genome.
Table 1b: Classification systems for HCV genotypes

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I PT 1a I 1a</td>
<td></td>
<td></td>
<td></td>
<td>HCV-1, HCV-H</td>
<td></td>
</tr>
<tr>
<td>II K1 1b II 1b</td>
<td></td>
<td></td>
<td></td>
<td>HCV-J, HCV-JT, HCV-BK</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1c</td>
<td></td>
<td>HC-G9, YS-117</td>
<td></td>
</tr>
<tr>
<td>III K2a 2a III 2a</td>
<td></td>
<td></td>
<td></td>
<td>HC-J6, HC-J5, HCV-K2a</td>
<td></td>
</tr>
<tr>
<td>IV K2b 2b III 2b</td>
<td></td>
<td></td>
<td></td>
<td>HC-J8, HC-J7, HCV-K2b</td>
<td></td>
</tr>
<tr>
<td></td>
<td>III 2c</td>
<td></td>
<td></td>
<td>S-83, T-983</td>
<td></td>
</tr>
<tr>
<td>V 3 IV 3a</td>
<td></td>
<td></td>
<td></td>
<td>HCV-K3a, T-1, T-7</td>
<td></td>
</tr>
<tr>
<td>VI IV 3b</td>
<td></td>
<td></td>
<td></td>
<td>HCV-TR, T-9, T-10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 4a</td>
<td></td>
<td></td>
<td>Z4, Z8, Z5, Syr1, Syr2, NS, Cam600, Z1, N1, N2, DK13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>V 5a</td>
<td></td>
<td></td>
<td>SA-1, SA-7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6a</td>
<td></td>
<td></td>
<td>HK-2</td>
<td></td>
</tr>
</tbody>
</table>

* It has been recommended that this classification be used in all future publications.
capability to express a protein that is generated by a translational frame shift mutation at codon 9 to 11 of the core gene (Xu et al, 2001). This contrast with the evolutionarily conserved nature of HCV replication supports the idea that this gene is more likely to be a computational artifact that has arisen from RNA structure- imposed constraints on third- codon position variability in the core gene (Tuplin et al, 2004).

Each major genotype of HCV contains a series of more closely related subtypes that differ from each other by 20-25% in nucleotide sequences (Table 1a & 1b).

Sequence variability within genotypes:

Several studies have described the rapid sequence drift of HCV over time, a process of diversification that leads ultimately to the existence of identifiably separate strains within human populations. By comparing HCV sequences from sequential samples from chronically infected individuals or from those infected from a common source, rates of sequence change were measured to be $1.44 \times 10^{-3}$ nucleotide changes per site per year over the whole genome (Okamoto et al, 1992, Smith et al, 1997).

The processes of neutral and adaptive evolution of HCV operate during the course of chronic infection within an individual, leading to continued nucleotide changes over time and the development of variable degrees of sequence diversity within the replicating population at a given
time point (Feitelson, 2002). The existence of a large and diverse population of viruses allows rapid adaptive changes in response to changes in the replication environment. This might take the form of evolving immune responses that select against viruses with specific T or B cell epitopes, and confer resistance to antiviral agents (Feitelson, 2002).

**Quasispecies:**

HCV, like many other RNA viruses, does not circulate in infected individuals as a homogenous population of identical viral particles, but as a pool of genetically distinct but closely related variants referred to collectively as a quasispecies. This nature of the virus confers a significant survival advantage, as the simultaneous presence of multiple variant genomes and the high rate at which new variants are generated allow rapid selection of mutants better suited to new environmental conditions (Weiner et al, 1991; Martell et al, 1992).

Viral heterogeneity results primarily from a high error rate of RNA dependent RNA polymerase, and enzyme encoded by the NS5B gene. In a given individual, the rate at which mutations accumulate during replication depends on the fidelity level of the viral RNA polymerase and on viral replication kinetics. Most mutant viral particles are replication deficient, but some propagate efficiently. The fittest infectious particles are selected continuously on the basis of their replication capacities and by the selective pressure resulting from immune responses (Weiner et al, 1992; Kato et al, 1993). During the chronic stage of infection, viral quasispecies are in a
state of equilibrium at any given time. Their composition maybe altered by any change in the environment. These changes can be spontaneous, related to complex metabolic interaction in the host, or triggered by external factors such as drug intake or antiviral treatments (Liang et al, 2000).

**Replication cycle:**

*Cell entry and uncoating:*

The HCV virion binds to the cell surface receptor, mediated through the viral envelope proteins. This interaction determines the host range and tissue tropism of the virus. Following uptake, the virus uncoats within acidic compartments such as endosomes, and releases the genome to begin the replication cycle. The genomic RNA then serves as a template for protein translation (Liang et al,2000).

*Replication:*

Translation of the HCV genome yields a polyprotein that is cleaved to form structural and non structural viral proteins. Non structural proteins, along with the HCV RNA template and host cell factors, are thought to form a ribonucleoprotein complex. Within this complex a negative strand copy of the RNA genome is produced, which in turn serves as the template for the production of progeny positive strand RNA (Ishido et al,1998).

*Virion assembly:*

Virion assembly begins with the interaction of capsid proteins and genomic RNA to form a nucleocapsid. The nucleocapsid then acquires an envelope and the mature virion is released from the infected cell. Because
HCV glycoprotein complexes are mostly retained in the ER, it is thought that HCV budding may occur in the ER or ER-like structures (Duvet et al, 1998). It is not known how the envelope is added to the nucleocapsid, but presumably the envelope proteins have a binding affinity for the assembled capsid proteins as has been described for alpha viruses (Cheng et al, 1995).

Pathogenesis:

**Histological characteristic of Hepatitis C:**

The liver is considered to be the primary site of HCV replication, as negative strand RNA is only detected in the liver, not in PBMCs when highly specific techniques such as tagged polymerase chain reaction are employed (Lanford et al, 1995). HCV RNA has been detected in the cytoplasm of hepatocytes. It is interesting to note that most infected hepatocytes display little or no apparent hepatocellular damage, suggesting that HCV is not a cytopathic virus. Steatosis may in part be attributed to the HCV core protein. Lymphoid aggregates and follicles with B-cell germinal centers are seen frequently in Hepatitis C, especially in portal tracts in areas of focal or piecemeal necrosis. In regions of active inflammation, hepatocytes express Fas antigen, and the generation of acidophil Councilman bodies is regarded as a result of CTL-induced hepatocytes apoptosis.

**Immunopathogenesis:**

Indirect evidence for a protective role of the cellular immune response is provided by observations that HCV specific CTL as well as T helper cells
have been shown in the blood of HCV RNA negative persons exposed repeatedly to HCV and in people who resolved HCV infection spontaneously. Both the strength and the quality of the CD4 positive T helper and the CD8 positive CTL response differ between patients who recover from acute Hepatitis C and those who develop chronic infection.

**Transmission:**

With the discovery of HCV, and the development of specific serological tests for its detection, it became obvious that HCV was mainly a blood-borne disease (Alter *et al*, 1989). High risk factors for HCV transmission are multiple transfusion of blood especially from paid donors, and intravenous drug abuse. Long term hemodialysis (shared machines or cross contamination of supplies), transplantation of infected organs into uninfected recipients are also important risk factors. In developing countries, needle stick injuries, use of contaminated syringes, tattooing with contaminated instruments are more common risk factors.

**Methods for HCV Genotyping**

*Molecular Genotyping:*

Since differences in geographical distribution, disease outcome, and response to therapy among HCV genotypes have been suggested, reliable methods for determining the HCV genotype may become an important
clinical test. The reference standard and most definitive method for HCV genotyping is sequencing of a specific PCR-amplified portion of the HCV genome obtained from the patient, followed by phylogenetic analysis. Investigators of HCV genotyping have used sequence analysis of HCV NS5, core, E1, and 5’ UTRs. However, direct sequencing is impractical on a large scale because of the complexity of the procedure. Even with the introduction of automated sequencing methods that do not require radioactive isotopes, only a few laboratories are equipped to perform such procedures on a regular basis. Finally, sequencing of amplified DNA does not usually identify mixed infections with two different HCV genotypes.

Other methods that have been reported depend mainly on the amplification of HCV RNA from clinical specimens, followed by either reamplification with type-specific primers or hybridization with type-specific probes (Widell et al, 1994; Okamoto et al, 1992; Li et al, 1994) or by digestion of PCR products with restriction endonucleases that recognize genotype-specific cleavage site (Mac Omish et al, 1993). HCV genotyping by using type-specific primers was first introduced by Okamoto et al (1992) and used primers specific for the core region. This method lacked sensitivity and specificity. Without modification, this method was able to detect subtypes 1a, 1b, 1c, 2a, 2b, and 3a. However, modifications have been introduced to improve the sensitivity and specificity of this method (Widell et al, 1994; Okamoto et al, 1996), but more studies are required before the efficiency of this genotyping method can be compared with that of other methods. Several DNA hybridization assays for HCV genotyping have been
described. A commercial kit (InnoLipa) for HCV genotyping has been introduced in Europe by Innogenetics (Zwijndre, Belgium) and is based on hybridization of 5' UTR amplification products with genotype-specific probes (Stuyver et al, 1993). Although the initial version of InnoLipa had lower sensitivity, the newer version is capable of discriminating among HCV subtypes 1a, 1b, 2a to 2c, 3a to 3c, 4a to 4h, 5a, and 6a (Maertens et al, 1997). It has been shown that genotyping methods using 5' UTR, including InnoLipa, may not distinguish subtype 1a from 1b in 5 to 10% of cases and also may not distinguish between subtypes 2a and 2c (Smith et al, 1995).

Others have used restriction enzymes to determine a restriction fragment length polymorphism. In this method, a PCR-amplified DNA fragment is digested into fragments with different lengths by enzymes (restriction endonucleases) that recognize cleavage sites specific for each genotype (Xavier et al, 1998). Investigators have used different regions of the HCV genome for restriction fragment length polymorphism, including NS5 and the 5' UTR (Bukh et al, 1992; Chan et al, 1992).

A new genotyping system based on PCR of the core region with genotype-specific PCR primers for the determination of HCV genotypes 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a, and 6a was developed (Ohno et al, 1997).

Although all these methods are able to identify correctly the major genotypic groups, only direct nucleotide sequencing is efficient in discriminating among subtypes (Bukh et al, 1995; Simmonds et al, 1995). Moreover, all of these PCR-based methods have the shortcomings and
advantages of PCR. They are expensive and time-consuming and require specialized facilities to ensure accurate results and prevent contamination. Their reliability may further be compromised if viral RNA is lost in the serum or plasma through storage or improper laboratory handling or if it is absent from the circulation during sample collection. The advantages of PCR-based methods include reliability if performed accurately and the ability to obtain information relevant to the molecular pathogenesis of HCV.

Serological Genotyping:

More recently, investigators identified genotype-specific antibodies that could be used as indirect markers for the HCV genotype (serotyping or serologic genotyping) (Simmonds et al, 1993; Tsukiyama-Kohara et al, 1993; Lopez-Labrador et al, 1997). Serological genotyping has several advantages that make it suitable for large epidemiologic studies. These advantages include the low risk for contamination and the simplicity of the assay. However, serological typing seems to lack specificity and sensitivity, which limits its usefulness (Feitelson, 2002).

Two commercially available serological genotyping assays have been introduced over the past 3 to 4 years. The RIBA SIA was introduced by Chiron Corp. and contained five different serotype-specific peptide sequences taken from the NS4 region and two serotype-specific peptide sequences taken from the core region of the HCV genomes for genotypes 1, 2, and 3 (Dixit et al, 1995). The second serologic genotyping assay is the Murex HCV serotyping enzyme immune assay (Murex Diagnostics Ltd., Dartford, Kent, UK), which is based on the detection of genotype-specific antibodies directed
to epitopes encoded by the NS4 region of the genomes for genotypes 1 through 6. These two assays have been compared and showed a concordance rate of more than 96% for genotypes 1, 2, and 3 (Gish et al, 1997).

A recent study by Beld et al (1998) showed high reliability of HCV serotyping by the RIBA SIA (Chiron Corp., Emeryville, Calif., USA) in immunocompetent individuals infected with genotype 1a. However, the assay had low sensitivity in samples containing genotype 3a or in samples from patients coinfected with HIV. These findings suggest that the use of this assay may be limited at this time, particularly in geographic regions where genotype 1a is not prevalent. Similarly, Songsivilai et al (1998) showed that serotyping had poor sensitivity for samples from patients infected with HCV genotype 6. Unlike the two previous studies, a study conducted in the United States reported high concordance between serologic genotyping (by use of type specific antisera) and molecular genotyping assays (Gish et al, 1997). These findings suggest variation in the reliability of these assays based on the distribution of HCV genotypes in a specific geographic area.

The choice of typing method for HCV should be based on the expertise in a specific laboratory or institution and the goal of typing. To identify all subtypes and to identify novel sequences if present, PCR amplification followed by sequencing should be the method of choice. However, the goal in treatment modalities is frequently required to separate patients infected with genotype 1 from those infected with other genotypes — a task that could be done adequately by any of the methods mentioned.
Role of Genomic Heterogeneity in HCV Persistence and Vaccine Development

With the rarity of severe acute or fulminant HCV infections, the significance of this infection in humans is its tendency to become persistent and to induce chronic liver disease. The mechanisms of HCV persistence are not known. However, in most human viral infections, the interaction of several arms of the immune system is important in limiting viral replication and preventing persistence. These arms of the immune system include humoral and cellular immunity.

Antibody responses are often directed against several viral proteins, although it is the antibodies directed against the viral envelope proteins that usually serve as neutralizing antibodies (Nelson et al, 1998). Neutralizing antibodies are often specific for a particular serologic type of the virus, an issue that is particularly relevant in discussions of strategies for vaccination against highly variable viruses such as HIV or HCV (Feitelson, 2002).

Whether infection with HCV elicits protective immunity in the host remains unclear. Farci et al (1994) attempted to neutralize HCV in vitro with plasma obtained from a chronically infected patient. The source of HCV was the same patient during the acute phase of posttransfusion non-A, non-B hepatitis. The residual infectivity was evaluated by inoculation of seronegative chimpanzees. The authors showed that neutralization was achieved with plasma obtained 2 years after the initial exposure but not with
plasma obtained 11 years later. Analysis of viral isolates for the same patient showed significant genetic divergence of HCV over time. These data support the quasispecies nature of HCV and the selection of strains to avoid immune pressure. This experiment also emphasized the possible role of genetic heterogeneity of HCV in escaping the immune system. It has been suggested that these antibodies are likely to be directed against epitopes of hypervariable region 1 located in the E2 region (Farci et al, 1994).

Similarly, cellular immune responses, particularly those mediated by cytotoxic T lymphocytes (CTLs), are important components of protective immunity against many viral infections, including hepatitis B. In HCV infection, the role of CTLs in protecting against viral persistence is unknown. HCV-specific, HLA class I-restricted CTLs were demonstrated within the liver. Possible targets for HCV-specific CTL recognition within the conserved core protein and additional epitopes in the more highly variable region E2 protein were also identified (Koziel et al, 1993). HCV heterogeneity may also be important in escaping CTL-induced immunity. In a chronically infected chimpanzee, CTLs obtained from the liver were initially able to recognize an epitope in the NS3 protein. Over a period of several years, a new strain of the virus emerged with a mutation in the CTL epitope that was no longer recognized by the CTLs isolated earlier. Although direct evidence for the presence of CTL escape mutants in human HCV infection is lacking, it has been shown that single amino acid changes in CTL epitopes result in failure of recognition by HCV-specific CTLs (Koziel et al, 1996). These
single-amino-acid changes are found in natural isolates of HCV, hence the need to address the problem of type specificity of immune responses.

**Outcome of Acute HCV Infection:**

After initial exposure to HCV, the infection fails to resolve in the majority of patients (80%) who become chronically infected. The ability to evolve into chronic disease associated with liver damage is by far the most striking feature of HCV. The spontaneous clearance of HCV following acute infection in a small proportion of patients has been the focus of intense investigations. It has been proposed that differences in the host cellular (Missale *et al.*, 1996) or humoral (Zein *et al.*, 1999) immune responses to HCV are important in spontaneous clearance, but these hypotheses remain to be proved.

Amoroso *et al.* (1998) specifically investigated the role of HCV genotypes in persistence of HCV infection following an acute exposure. The rate of evolution to chronicity after acute exposure to HCV was 92% in patients exposed to HCV genotype 1b infection, compared with 33% to 50% in patients exposed to other genotypes (Amoroso *et al.*, 1998). These data provided evidence that viral factors, including the HCV genotype, may potentially play an important role in the development of chronic infection following acute exposure to HCV.

**Progression of Liver Disease:**

The role of HCV genotypes in the progression of liver disease is one of the most controversial areas of HCV research. There appears to be
significant biologic variation in HCV disease expression in the host over the length of the infection (typically the life of the patient). In a retrospective analysis of patients with chronic HCV infection whose time of HCV acquisition was known, Zein *et al* (1996) detected variation among the infected persons. The mean times from exposure to HCV to the diagnosis of chronic active hepatitis, to compensated liver cirrhosis, to decompensated cirrhosis, and to hepatocellular carcinoma were 11, 18, 23, and 29 years, respectively (*Zein et al*, 1996). What is striking is that severe complications such as cirrhosis and hepatocellular carcinoma can occur over a short period in some persons whereas others have no complication despite a much longer period of infection (*Liang et al*, 2000). Therefore, it is likely that viral or host factors, including the infecting HCV genotype, contribute to these variations in the natural history among infected patients.

Currently, investigators are divided into those who strongly believe in differences in pathogenicity among genotypes and those who do not. Conclusions have been derived from indirect evidence, because conducting accurate investigations to answer these questions has been difficult. Frequently, the role of genotypes as an independent factor in the progression of liver disease cannot be separated from the roles of other cofactors such as viral load, alcohol intake, and length of time of HCV infection. Patients may not provide accurate information about drug use or the amount of alcohol intake; therefore, the time of HCV acquisition often is not known. Because of
the overall slow progression of liver disease in HCV-infected patients, prospective studies often are not possible.

In patients with chronic HCV, infection with genotype 1b is reportedly associated with a more severe liver disease and a more aggressive course than is infection with other HCV genotypes (Nousbaum et al., 1995). Similar observations were also made by others (Pozzato et al., 1995; Silini et al., 1995; Zein et al., 1996), who found that HCV genotype 1b was significantly more prevalent among patients with liver cirrhosis and those with decompensated liver disease requiring liver transplantation than among those with chronic active hepatitis C infection. Although this is an indirect evidence, but it suggests that an association exists between HCV genotype 1b and the development of these complications. Furthermore, a possible link to hepatocellular carcinoma has been proposed for HCV genotype 1b. There is compelling evidence that hepatocellular carcinoma occurs more frequently or emerges earlier among HCV-infected Japanese patients (Yano et al., 1993) than among HCV carriers in western countries (Di Bisceglie et al., 1991). Because HCV genotype 1b is more common in Japan than in Europe or in the United States, the hypothesis relating to genotype is attractive and appears to explain these differences. Furthermore, HCV genotype 1b was present in most of the patients with HCV-associated hepatocellular carcinoma studied by Zein et al (1996). Similarly, Reid et al (1994) determined the HCV genotypes in 28 patients with hepatocellular carcinoma and found that 19 (68%) were infected with HCV genotype 1b and the rest were infected with a
mixture of HCV genotypes that always included genotype 1b (Reid et al, 1994).

Some reports refuted the associations as described above (Han et al, 1997; Naoumov et al, 1997; Simmonds et al, 1996). A possible and simple explanation may reconcile these reported discrepancies. Zein et al (1996) found that patients infected with HCV genotype 1b were older than those infected with other genotypes and that genotype 1b may have been present before the other genotypes. Thus, patients infected with genotype 1b may have been infected for a longer time. Similar observations have been made in France and Spain (Lopez-Labrador et al, 1997). According to this explanation, HCV genotype 1b is a marker for more severe HCV-associated liver disease, because it reflects a longer time of infection rather than a more aggressive form of hepatitis C infection. Future studies are still needed to rule out other host, viral, or environmental factors that may contribute to these differences.

Zein et al (1995) and Gordon et al (1997) reported that in liver transplant recipients, HCV genotype 1b is associated with earlier recurrence and more severe hepatitis than are other genotypes. The difference in the duration of infection that may have been a factor in non-transplant-associated HCV patients is not likely to explain the discrepancies in the literature about posttransplantation HCV. However, more studies are needed to delineate the above observation.
Response to Interferon Therapy

Since the discovery of HCV, considerable effort has been devoted to defining the factors that may be important in predicting the long-term response to interferon therapy (Zein, 2000). The interferon dose, duration of treatment, viral RNA level, and liver histology all seem to play a role in predicting response (Zein, 2000). It has been suggested that patients infected with HCV genotypes 1b and, to a lesser degree, 1a are less likely to have a favorable response to interferon treatment than are those infected with genotype 2 or 3 (Zein, 2000). Zein et al (1996) reported a complete biochemical response at the end of 6 months of treatment with interferon in 60 to 70% of patients infected with HCV genotype 2 and in 10 to 15% of those infected with genotype 1. This difference was also present for sustained response and was independent of liver histologic features or the pretreatment HCV RNA levels. This may partly explain the higher rates of long-term response to interferon treatment that have been reported in Europe, where HCV genotype 2 is more prevalent than in the United States or Japan. However, a meta-analysis of the most relevant studies was performed recently (Davis et al, 1997), and although there was a difference, the predictive value of the non-1 HCV genotype for response to treatment was low (58% for response at the end of treatment and 55% for sustained response). It has been suggested that the introduction of more effective therapies such as the combination of interferon and ribavirin may make the
value of predictive factors for response to therapy less important and the differential response of HCV genotypes less obvious (Davis et al, 1997). However, more recent treatment trials using interferon plus ribavirin in interferon-naive patients with chronic HCV (McHutchinson et al, 1998) or in patients in whom previous interferon treatment failed (Davis et al, 1998) showed higher rates of sustained response to therapy in patients with HCV genotypes other than 1. Among patients with HCV genotype 1, 48 weeks of treatment was required to achieve a response similar to that of patients infected with other genotypes treated for 24 weeks (McHutchinson et al, 1998).

An "interferon-sensitive" region in the nonstructural portion of the HCV genome has been identified in Japanese patients infected with genotype 1b (Enomoto et al, 1995). However, studies from the United States and Europe failed to confirm these findings (Odeberg et al, 1998). The significance of these findings is not known, and the clinical application of such an expensive and labour-intensive procedure to predict the response to treatment is impractical.

**Treatment response of Different genotypes:**

Typically 10-20 and 40-50% of individuals infected chronically with genotype 1 HCV on mono therapy and combination therapy respectively, exhibit complete and permanent clearance of virus infection. This long term response rate is much lower than the rates of 50 and 70-80% that are observed
on treatment of HCV genotype 2 or 3 infections (Pawlotsky et al, 2003, Zeuzem et al, 2004). This difference has proved to be highly significant in patient management in order to achieve acceptable efficacy of the treatment. Very recently, it was observed that genotype 2 and 3 infected individuals were non responsive or weakly responsive to short term treatment with the BILN 2061 protease inhibitor (Reiser et al, 2004), in contrast to its efficacy in genotype 1 infected individuals (Lamarre et al, 2003). Genotype specific differences in response to the new generation of antiviral agents will be a major research priority in the future.

Molecular Epidemiology:

Geographical Distribution of HCV Genotypes

Because of geographic clustering of distinct HCV genotypes, genotyping may be a useful tool for tracing the source of an HCV outbreak in a given population as it has been shown by tracing the source of HCV infection in a group of Irish women to contaminated anti-D immunoglobulins (Power et al, 1995). All of these women were infected with HCV genotype 1b, a genotype identical to the isolate obtained from the implicated batch of anti-D immunoglobulin. Similar observation has been made by Hohne et al (1994) who used genotyping to trace the sources of outbreaks in Germany (Hohne et al, 1994). More recently, genotyping and molecular characterization of HCV isolates provided evidence for a patient-to-patient transmission of HCV during colonoscopy (Bronowicki et al, 1997). The index case as well as the two other infected patients had HCV genotype 1b.
Nucleotide sequencing of the NS3 region showed that the three patients had the same isolate (100% homology), strongly suggesting a common source of infection.

Suspected nonconventional routes of HCV transmission could also be investigated by molecular analysis of HCV strains from different persons. These include the vertical and sexual routes. Weiner et al. (1993) showed that a single predominant HCV variant was transmitted to an infant born to a mother infected with multiple variants. Gish et al. (1996) reported similar findings in their study. A specific 12-nucleotide insertion in the E2 hypervariable region of the HCV genome was noted in the vertically transmitted sequence of an infant born to a mother infected with two different genotypes, each composed of multiple heterogeneous sequences (Aizaki et al., 1996). These data may suggest a potential role of HCV heterogeneity and genotypes in mother-to-infant transmission of HCV (Zein, 1997).

Reports on the sexual transmission of HCV infection are conflicting. The detection of anti-HCV positivity ranged from 0% in partners of transfusion-associated hepatitis patients (Everhart et al., 1990) to 8% in male homosexuals (Esteban et al., 1989) and 5% in household contacts (Ideo et al., 1990). A possible explanation is that sexual transmission occurs only in association with specific HCV genotypes or in the presence of specific mutations along the HCV genome. As with vertical transmission, samples from patients with suspected sexual transmission of HCV have undergone nucleotide sequence analysis to confirm the similarity of sequences obtained from sexual partners and thus the common origin of these HCV strains.
Although Zein et al. (1996) observed that there was no association between HCV genotypes and the mode of HCV acquisition in their population, others have provided evidence for such an association (Watson et al., 1996; Pawlotsky et al., 1995; Berg et al., 1997). It has been suggested that genotypes 3a and 1a are closely associated with intravenous drug use and that genotype 1b is seen more often in patients who acquired HCV through blood transfusion. This information may be useful in tracing sources of HCV epidemics.

At least six major genotypes of HCV, each comprising multiple subtypes, have been identified worldwide. Substantial regional differences appear to exist in the distribution of HCV genotypes (Zein, 2000) (Fig. 1b). Although HCV genotypes 1, 2, and 3 appear to have a worldwide distribution, their relative prevalence varies from one geographic area to another (Zein, 2000).

HCV subtypes 1a and 1b are the most common genotypes in the United States (Zein, 2000). These subtypes also are predominant in Europe (Nousbaum et al., 1995; Mac Omish et al., 1994). In Japan, subtype 1b is responsible for up to 73% of cases of HCV infection (Takada et al., 1993). Although HCV subtypes 2a and 2b are relatively common in North America, Europe, and Japan, subtype 2c is found commonly in northern Italy. HCV genotype 3a is particularly prevalent in intravenous drug abusers in Europe.
Fig. 1b: Worldwide distribution of HCV genotypes. Genotype 1a in USA, Europe; 1b in USA, Europe, Japan, India; 2a & 2b in North America, Europe, Japan; 2c in Northern Italy; 3a in IDUs in Europe & USA, and otherwise in general in India, Pakistan, Nepal, Bangladesh; 3b in South Asian (India, Pakistan, Bangladesh, Nepal, Srilanka) countries; Genotype 4 in North Africa & Middle East; 5 in South Africa & Hong Kong; 6 in South Africa, Hong Kong & other South East Asian countries.
HCV genotypes 4 appears to be prevalent in North Africa and the Middle East (Chamberlain et al, 1997; Abdulkarim et al, 1998), and genotypes 5 and 6 seem to be confined to South Africa and Hong Kong, respectively (Cha et al, 1992; Simmonds et al, 1993). HCV genotypes 7, 8, and 9 have been identified only in Vietnamese patients (Tsukiyama-Kohara et al, 1993), and genotypes 10 and 11 were identified in patients from Indonesia (Tokita et al, 1996). There has been disagreement about the number of genotypes into which HCV isolates should be classified. Investigators have proposed that genotypes 7 through 11 should be regarded as variants of the same group and classified as a single genotype, type 6 (Tokita et al, 1998; Simmonds et al, 1996; Mellor et al, 1996; de Lamballerie et al, 1997).

The geographic distribution and diversity of HCV genotypes may provide clues about the historical origin of HCV (Smith et al, 1997). The presence of numerous subtypes of each HCV genotype in some regions of the world, such as Africa and Southeast Asia, may suggest that HCV has been endemic for a long time. Conversely, the limited diversity of subtypes observed in the United States and Europe could be related to the recent introduction of these viruses from areas of endemic infection.

**Genotype distribution in India**

A second generation assay for antibody to hepatitis C virus (anti-HCV) was used to screen 78 southern Indian individuals with a high risk of infection (Valliammai et al, 1995). RT-PCR targeted at the 5' end
untranslated region (5'UTR) of the HCV genome was used to evaluate evidence of viraemia in 32 anti-HCV positive sera. The PCR products amplified from the 5'UTR of the HCV genome from 24 patients were sequenced, revealing the existence of two distinct groups of sequences: 21 corresponded to HCV type 1 while the other three sequences had 95% to 99% identity to HCV type 3. Two of these three isolates had more than 90% nucleotide identity in the NS5 region to established 3b sequences whereas the other had less than 74% nucleotide identity to any of the published genotype 3 (3a, 3b, 3c, 3d, 3e and 3f) sequences. However, a search of the EMBL nucleotide database revealed 91% identity to the unpublished sequence of an isolate of HCV from Indonesia. This study provide evidence that these two isolates may represent a novel subtype within genotype 3. This data also suggest that HCV genotype 1 predominates over HCV genotype 3 in southern India. (Valliammai et al, 1995).

In another study, 11 isolates from India were genotyped using sequence comparison for part of the non-structural (NS5) and structural (core) regions. Parts of the genome covering 451 bp (nt 9-459) of the core gene and a 249 bp fragment (nt 7959-8207) of the NS5 gene were reverse transcribed and amplified using nested polymerase chain reaction (RT-PCR). The amplified fragments were cloned and sequenced. The classification into genotypes was done on the basis of phylogenetic analysis. Four isolates showed sequence homology to type 1b. Two of the isolates were classified as type 3a. One isolate was classified as type 3b and the remaining four isolates were found to be variants of type 3 but did not belong to any designated
subtype. On the basis of phylogenetic analysis two of the unclassified isolates were clubbed together as a new subtype of 3 named as 3g. In one of these variants, parts of a 5'-noncoding (5' NCR; 204 bp), envelope-E1 (435 bp), and NS3 (502 bp) regions were also amplified, cloned, and sequenced. This study demonstrated the type 3 variants including a new subtype (3g) to be the major cause of HCV infection in India. (Panigrahi et al, 1996)

In another study, prevalence of genotypes were as follows: type I – 21%, type II – 25%, type III – 54%. In cirrhosis patients, 77% of genotype I, 33% of genotype II and III. Genotype I (35%), genotype II (22%) and genotype III (43%) were reported among CRF and renal transplant patient. (Amarapurkar et al, 2001).

Genotype 3 was detected more frequently in patients from eastern India (62.2%), as compared with its detection in patients from southern India (P=0.004). Similarly, genotype 1 was detected with greater frequency in individuals from southern India as compared with patients from eastern India (P=0.004). HCV genotypes 1 and 3 accounted for 81% of HCV infections in patients from this geographical region. (Raghuraman et al, 2004).

Healthy blood donors at the All India Institute of Medical Sciences (AIIMS), New Delhi, India, were screened for anti-HCV antibody. 1.85% of these donors were positive for HCV infection. The screening was also used to determine how many patients with acute hepatitis and chronic liver diseases were positive for anti-HCV antibody among them. About 13.83% chronic cases were found positive for HCV infection. Among sporadic acute viral
hepatitis patients, HCV was associated with 9% of the cases. (Panigrahi et al, 1997).

**Treatment and Prevention**

Early studies indicate that interferon and ribavirin are effective in cases of acute and chronic hepatitis C. A combination of interferon and ribavirin may be useful (Bonkovsky et al, 1997). There is more experience in the use of interferon for the treatment of hepatitis C. The current recommendation is that interferon treatment may be considered in those cases with chronic active hepatitis who are at risk of progression to cirrhosis and hepatocellular carcinoma. The recommended regimen is 3 MU of interferon thrice daily by subcutaneous or intramuscular route for six months (Liang et al, 2000). The response rate is around 50%. However, approximately 50% of responders relapse upon cessation of treatment (Liang et al, 2000). At present, it is not clear what factors predict response to interferon therapy. There is some data to suggest that older patients and those with established cirrhosis respond less well. There is also growing evidence that the genotype of the infecting HCV determines the response to interferon therapy. Type 1b, in particular is associated with poor response to IFN. It also appears that the level of viraemia at the commencement of treatment predicts the response (Liang et al, 2000). At present, it is not clear what factors predict relapse after treatment. For those who relapse after treatment, they may be offered a
The introduction of HCV screening in the blood transfusion service should reduce significantly the number of cases of post-transfusion hepatitis C. However, the extra cost is high as the current screening ELISA is 5 times more expensive than the anti-HIV screen test. Inactivation procedures should eliminate the risk of HCV spread via blood products. However, the above mentioned measures would only contribute to the prevention of a small proportion of cases of hepatitis C. Therefore, further preventive measures would depend on a better understanding of the epidemiology of hepatitis C infection. The efficacy of immunoglobulin for pre- or post-exposure prophylaxis is still controversial.