Review

Correspondence
Mauro Costa-Mattioli
mauro.costa-mattioli@mail.mcgill.ca

Genetic variability of hepatitis A virus
Mauro Costa-Mattioli,1 Anna Di Napoli,2 Virginie Ferré,3 Sylviane Billaudel,3 Raul Perez-Bercoff2 and Juan Cristina4

1Department of Biochemistry, McGill University, McIntyre Medical Building, Montreal, Quebec, Canada H3G 1Y6
2Laboratoire de Virologie Moléculaire et Structurale, CNRS, 1 Avenue de la Terrasse, 91198 Gif-sur-Yvette, France
3Laboratoire de Virologie UPRES-EA1156, Institut de Biologie, Centre Hospitalier Regional Universitaire de Nantes, Rue Quai Moncousu 9, 44093 Nantes, France
4Departamento de Técnicas Nucleares Aplicadas, Centro de Investigaciones Nucleares, Facultad de Ciencias, Universidad de la República, Iigua 4225, 11400 Montevideo, Uruguay

Knowledge of the molecular biology of hepatitis A virus (HAV) has increased exponentially since its identification. HAV exploits all known mechanisms of genetic variation to ensure survival, including mutation and genetic recombination. HAV has been characterized by the emergence of different genotypes, three human antigenic variants and only one major serotype. This paper reviews the genetic variability and molecular epidemiology of HAV. Its evolutionary mechanisms are described with particular emphasis on genetic recombination and HAV mutation rate. Genotypic classification methods are also discussed.

Incidence, morbidity and mortality
Human hepatitis A, a widespread infectious disease that is hyperendemic in vast areas of the world, results from the infection of the liver by hepatitis A virus (HAV). In fully symptomatic cases the patient presents with jaundice, fever, anorexia, nausea, vomiting, headache and fatigue. In non-complicated cases recovery takes between 4 and 6 weeks, but fulminating cases may be fatal (Ross & Anderson, 1991).

Transmission occurs mainly through the faecal-oral route where insufficient sanitation or poor hygienic conditions favour the pollution of water and food, especially shellfish (Hadler et al., 1980). Moreover, 63% of the total number of viral hepatitis cases reported in 1998 were associated to hepatitis A (CDC, 1999), but the true incidence of hepatitis A is thought to be much higher as the disease very often evolves without specific symptoms (an-icteric) and consequently goes undetected. Residence in correctional facilities, employment in or attendance at day-care centres for children, male homosexuality and travel to countries with poor sanitation are all risk factors.

In Canada, between 1000 and 3000 cases of hepatitis A are reported every year. Hepatitis A accounts for 25–40% of all cases of acute hepatitis, with a variable incidence rate of 4–11·2 per 100 000 individuals. One in four cases requires hospitalization and the mortality rate among hospitalized adult cases has been estimated at about 0–15%.

Travellers account for 10–15% of all reported cases of hepatitis, with an annual estimated rate of 3·6–7·2 per 100 individuals. Epidemiological studies indicate that the morbidity and mortality rate of hepatitis A among travellers is 500 times higher than those of cholera, 150 times higher than poliomyelitis, 10 times those of typhoid fever and three times higher than hepatitis B (Lenfant, 1994).

Systematic immunological surveys indicate that in South America over 90% of the low-income population bears serological evidence of infection at the age of 18, and hepatitis A is responsible for more than 50% of all acute hepatitis cases reported to national reference centres (Abuzwaida et al., 1987; Gaspar et al., 1992; Montano, 2002).

In southern Europe and the Mediterranean, about 90% of the population bears antibodies against HAV, whereas less than 3% of Scandinavians appear to have had contact with the virus (Frosner et al., 1979; Flehmig, 1980; Daemer et al., 1981; Mele et al., 1986).

The cost of a non-complicated case of hepatitis A in France amounts to about € 300 (and there are between 30 000 and 50 000 new cases per year), whereas a fulminating case costs about € 6 000 and every year between 50 and 60 cases are reported. In Italy, after a few years of decline, hepatitis A re-emerged, with 11 000 cases reported in the Puglia region (Germinario et al., 2000).

The current trend of migration of large numbers of nationals from developing countries into the EU is bound to bring...
new variants of HAV and this should stress the need of simple, reliable and cheap methods for detecting antigenically unrelated variants both in the environment and in pathological specimens.

The virus

The aetiological agent of hepatitis A is a picornavirus of the genus *Hepatovirus* (Matthews, 1979, 1982; Melnick, 1992), but, in contrast to all other members of the *Picornaviridae*, HAV replicates very slowly in all tissue culture systems tested so far (days instead of hours), without interfering noticeably with host-cell macro-molecular synthesis and, consequently, without apparent cytopathic effect (Hollinger & Emerson, 2001). Adaptation to tissue culture conditions requires a series of blind passages.

HAV is a non-enveloped, 7.5 kb positive-stranded RNA virus (Hollinger & Emerson, 2001). Like all picornaviruses, the genome can be defined into three distinct regions (Fig. 1). The 5′ UTR contains an extensive secondary structure required for cap-independent translation and is covalently linked at the viral protein VPg. A single open reading frame encodes all of the viral proteins: P1 encodes the structural proteins (VP1, VP2 and the putative VP4) and the P2 and P3 regions encode the non-structural proteins associated with replication (Fig. 1). Finally, there is a short 3′ UTR terminating in a poly(A) tract.

Only one serotype of HAV has been identified thus far (see below), and the only naturally antigenic variants are HAV strains collected from Old World monkeys (Nainan *et al.*, 1991; Tsarev *et al.*, 1991). Although, studies with monoclonal antibodies (mAbs) suggest that there are a limited number of antigenic epitopes closely grouped at the surface of the virus (Stapleton & Lemon, 1987; Ping *et al.*, 1988; Ping & Lemon, 1992), three human antigenic variants were reported recently (Costa-Mattioli *et al.*, 2002; Sanchez *et al.*, 2002).

Genotype history and classification

The first comparative study of the nucleotide sequences of HAV strains, which employed RNA-fingerprinting, demonstrated a significant heterogeneity among the strains analysed (Weitz & Siegl, 1985). In the mid-1980s, several groups reported the first nucleotide sequences of different HAV strains isolated from hepatitis outbreaks of diverse origin (Ticehurst *et al.*, 1983; Baroudy *et al.*, 1985; Linemeyer *et al.*, 1985; Najarian *et al.*, 1985; Venuti *et al.*, 1985; Ovchinnikov Iu *et al.*, 1985; Divizia *et al.*, 1986; Cohen *et al.*, 1987a). Five HAV strains were adapted to growth in cell culture before molecular cloning (Linemeyer *et al.*, 1985; Najarian *et al.*, 1985; Ovchinnikov Iu *et al.*, 1985; Venuti *et al.*, 1985; Paul *et al.*, 1987). One of them was the first isolate reported to induce cytopathic effect in vitro (Venuti *et al.*, 1985). Strain HM-175, currently the HAV reference strain, was first isolated from an outbreak in Australia (Gust *et al.*, 1985) and was subsequently passaged three times in marmosets. Determination of the complete nucleotide sequence of wild-type HM-175 allowed comparison of this sequence to those from cell culture-adapted HAV strains as well as to sequences from other picornaviruses (Cohen *et al.*, 1987a, b), and the study of genomic heterogeneity among human and non-human strains of HAV using cDNA–RNA hybridization probes (Lemon *et al.*, 1987). In 1989, genetic comparison of human and simian HAV strains revealed that VP1 and the 5′ UTR are the most variable and conserved genomic regions, respectively (Brown *et al.*, 1989). Moreover, Ticehurst *et al.* (1988) reported that different human HAV strains of diverse geographic origin were remarkably closely related.

The founding works on HAV genetic variability made use of discrete, selected partial HAV genomic regions, such as the C terminus of VP3 (Jansen *et al.*, 1990), the N terminus of VP1 (Robertson *et al.*, 1991) or the putative VP1–2A junction region (Fig. 1) (Jansen *et al.*, 1990; Robertson *et al.*, 1992). Since HAV and poliovirus share many genomic

![Fig. 1](image-url)
features, the different HAV strains were grouped by comparing the VP1–2A junction and using the method of Rico-Hesse and co-workers, a criterion used at the time for the genetic classification of poliovirus strains (Rico-Hesse et al., 1987). In 1992, using this approach, genetic analysis of 152 strains of HAV recovered around the world resulted in the designation of seven genotypes of HAV (I–VII) (Fig. 2A). A genotype is therefore defined as a group of viruses with >85% nucleotide sequence identity (Robertson et al., 1992). Viruses from four of the genotypes (I, II, III and VII) were recovered from HAV human cases, whereas viruses from the other three genotypes (IV, V and VI) were isolated only from unique simian species developing a hepatitis A-like illness during captivity (Nainan et al., 1991; Tsarev et al., 1991; Robertson et al., 1992). Genotypes I and III were subdivided further into two distinct groups (subgenotypes), which differed in sequence in no more than 7.5% of base positions.

This important work by Robertson and co-workers significantly influenced the direction of future research in the field (Robertson et al., 1992). Nevertheless, the vast majority of strains included in theses studies were isolated in the USA and Asia, whereas other regions of the world which have a hyperendemic pattern of HAV, such as South America, North and Central Africa and India, were conspicuously under represented. Moreover, by using the traditional method of genotyping, three HAV antigenic variants reported recently were not detected (Costa-Mattioli et al., 2002; Sanchez et al., 2002). This observation indicates that traditional genotyping methods may not always reflect antigenic variations.

To gain insight into the genetic variability of HAV and as an alternative to comparing the percentage of identity within short (168 nucleotides) selected regions of the genome, phylogenetic studies were carried out using full-length VP1 sequences (900 nucleotides) (Costa-Mattioli et al., 2002). Why was the full-length VP1 region chosen? Firstly, VP1 is the major surface-accessible protein in the mature picornavirus particle (Hughes et al., 1984; Hogle et al., 1985; Mateu et al., 1995). HAV escape mutants generated under mAb selection have shown that a number of amino acids within the VP1 protein contribute to the major immunodominant site of HAV (Stapleton et al., 1985; Stapleton & Lemon, 1987; Ping et al., 1988; Ping &

![Fig. 2. Phylogenetic analysis of different genetic regions of HAV. Neighbour-joining phylogenetic trees of the VP1–2A junction (A) and full-length VP1 (B), VP3 (C) and VP2 (D) regions using Kimura’s 2-parameter model. Genotypes and subgenotypes are indicated at nodes. Bars indicate genetic distances. The recombinant (9F94) and two putative parental strains (MBB and SLF88) are shown in bold and italics, respectively.](http://vir.sgmjournals.org)
Lemon, 1992). Secondly, molecular evolution studies have recently shown beyond doubt that sequence determination of the full-length VP1 was required to properly characterize human and simian strains of picornaviruses (Oberste et al., 1998, 1999, 2001; Brown et al., 1999; Haydon et al., 2001; McMinn et al., 2001; Norder et al., 2001). Thirdly, since the VP1 sequence of enterovirus correlates to some extent with serotype, the VP1 sequence can be used as a molecular surrogate for antigenic typing (Oberste et al., 1999, 2000, 2001, 2002; Brown et al., 2000; Caro et al., 2001; Casas et al., 2001; Costa-Mattioli et al., 2002; Norder et al., 2001).

Phylogenetic analysis of complete VP1 sequences of 86 HAV strains isolated worldwide unexpectedly revealed the presence of five distinct genetic groups (Fig. 2B), all of them supported by high bootstrap values (Costa-Mattioli et al., 2002). The only sequences not included in these studies were the one of strain JM-55 (genotype VI) and those representing genotype IIIB, because none of them were available in a public sequence database at that time. Based on these studies, a novel classification of HAV genotypes was proposed (Costa-Mattioli et al., 2002). Strikingly, the least observed variation was found between the previously established genotypes II and VII (Costa-Mattioli et al., 2002). These findings suggest that the previously defined genotypes II and VII (Robertson et al., 1992) may be one or two subgenotypes of the same type. A simple hypothesis is that a recombination event has taken place among strains of these two genotypes. To test this, phylogenetic analysis of the three major capsid proteins (VP1, VP2 and VP3) and the VP1–2A junction were carried out. Phylogenetic comparison of VP1 full-length protein (Fig. 2B) and the VP1–2A region (Fig. 2A) (Robertson et al., 1992), with the phylogeny based on the entire VP2 and VP3 sequences (Fig. 2C, D) (Costa-Mattioli et al., 2003), revealed that the phylogenetic trees were non-congruent, consistent with the notion that some of these strains might have arisen by recombination.

To determine the basis of this discrepancy, recombination analysis using RIP (recombination identification program) was carried out (Fig. 3A) (Costa-Mattioli et al., 2003). Analyses of these data sets showed that a recombination event occurred between the putative parental strains MBB (genotype IB) and SLF-88 (genotype VII) to generate the mosaic 9F94 strain (Fig. 3) (Costa-Mattioli et al., 2003). Indeed, a further characterization of the putative break-point site using the LARD method (Holmes et al., 1999) clearly shows the point of recombination on position 1461 of the sequences analysed, corresponding to position 51 of the VP1 protein (Costa-Mattioli et al., 2003).

Phylogenetic analysis may be severely biased if recombination events are ignored (Schierup & Hein, 2000; Posada & Crandall, 2001). Since only one other representative of type II has been identified to date, and its identification is

---

**Fig. 3.** (A) Analysis of recombination events occurring in the viral genome (top) using RIP. The x-axes represent the nucleotide position along the HAV capsid region. The recombination break point is indicated by an arrow. (B) Schematic representation of the putative parental and mosaic genome organization.
based on limited sequence analysis (Robertson et al., 1992), it is possible that type II strains may not be a distinct human genotype but rather an artefact of classification based on insufficient sequence analysis.

**Distribution of HAV genotypes**

Analysis of 152 strains of HAV revealed that genotype I was the most abundant type worldwide, particularly genotype IA, which included HAV strains from North America, China, Japan, the former USSR and Thailand (Jansen et al., 1990; Robertson et al., 1992).

It was found that HAV strains from Central and South America all belonged to subgenotype IA, suggesting that there is a circulating endemic population in these countries (Arauz-Ruiz et al., 2001; Costa-Mattioli et al., 2001a; Diaz et al., 2001; Mbayed et al., 2002). A recent report, however, has revealed the concomitant circulation of subgenotypes IA and IB in Brazil (de Paula et al., 2002, 2003). In addition, IB subgenotype HAV strains were isolated from shellfish imported from Peru (Sanchez et al., 2002).

Genotype IB contained strains from Jordan, North Africa, Australia, Europe and Japan (Robertson et al., 1992). Co-circulation of subgenotype IA and IB was reported in South Africa (Taylor, 1997).

In Europe, a more complex pattern has been observed because HAV isolates derive from multiple genotypes, probably representing viruses from other regions (Robertson et al., 1992; Apaire-Marchais et al., 1995; Bruisten et al., 2001; Costa-Mattioli et al., 2001b, 2002; Pina et al., 2001; Sanchez et al., 2002). For instance, in western Europe, HAV isolates were derived from multiple genotypes (Apaire-Marchais et al., 1995; Costa-Mattioli et al., 2001b, 2002; Pina et al., 2001; Sanchez et al., 2002; Chironna et al., 2003).

Most of the remaining human HAV strains are segregated in genotype III, which has two subgenotypes (IIIA and IIIB) (Robertson et al., 1992). The prototype virus strain of the genotype IIIA, PA21, a subgenotype linked previously to intravenous drug users in Sweden during the 1980s (Robertson et al., 1992) and in Norway at the end of the 1990s (Stene-Johansen et al., 1998), was originally isolated from captured Panamanian owl monkeys (Brown et al., 1989). Strains closely related to this genotype have been collected from humans with HAV in India, Sri Lanka, Nepal, Malaysia and USA (Jansen et al., 1990; Robertson et al., 1991; Khanna et al., 1992). Recently, genetic analysis of HAV strains isolated from environmental samples, such as a shellfish-associated outbreak in France, Spanish sewage samples and mussels imported to Italy, revealed for the first time the presence of strains closely related to genotype IIIA in those countries (Costa-Mattioli et al., 2001b; Pina et al., 2001; Chironna et al., 2003). Therefore, the molecular epidemiology of HAV infection in southern European countries (Italy, Spain and France) seems to confirm the presence of different HAV strain variants in western Europe and the Mediterranean. In general, genetic analysis of strains can provide valuable information with regard to the source of the virus in both sporadic and epidemic infection (De Serres et al., 1999; Costa-Mattioli et al., 2001b; Pina et al., 2001; Chironna et al., 2003). Recently, Tallo et al. (2003) reported that, although genotype IA was the predominant one in Estonia during the past years, the last HAV outbreak was associated to a genotype IIIA HAV strain. In fact, the replacement of one genotype with another one may explain, at least in part, the shift in HAV epidemiology from a disease in children to an infection in young adults (Ciacciara, 2000; Tanaka, 2000). Taken together, these data suggest that genotype IIIA is becoming significantly more prevalent among the HAV-infected population than previously appreciated.

Early studies on HAV isolates from cell culture have shown little (if any) genetic variation between the different strains, most likely associated to cell culture cross-contamination problems. However, more recent PCR studies based on strains isolated from clinical specimens (stools, liver suspension and serum) have shown more genetic heterogeneity. In regions of the world such as the USA, Japan and China, HAV-related isolates tend to cluster, suggesting an endemic spread. A high degree of genetic conservation was shown during the infection period of an individual (Robertson et al., 1992, 2000) or even among different isolates with a common source of infection (Grinde et al., 1997; Chudy et al., 1999; De Serres et al., 1999; Arauz-Ruiz et al., 2001; Diaz et al., 2001; Tallo et al., 2003). In contrast, a higher degree of heterogeneity than reported previously has been found in strains isolated in South America (Costa-Mattioli et al., 2001a, 2002; de Paula et al., 2002; Mbayed et al., 2002). Surprisingly, the N terminus of the VP1 region from strains isolated in South America was more variable than the VP1–2A junction region (Costa-Mattioli et al., 2001a; Mbayed et al., 2002). Moreover, these strains did not cluster according to geographic origin, as was reported for isolates from Europe (Costa-Mattioli et al., 2002) and from other regions of the world (Robertson et al., 1991, 1992, 2000; Robertson & Naiman, 1997; Taylor, 1997). Recent studies suggest a changing epidemiologic pattern in HAV infection throughout South America, which may result in more clinical cases in teenagers and adults and a greater potential for new outbreaks (Tapia-Conyer et al., 1999; Tanaka, 2000). Whether this changing pattern is related to a higher genetic variability of HAV in that particular geographic region than expected previously to changes in hygienic conditions or to a combination of these and other factors remains to be established.

Genotypes II and VII were also described using the traditional approach (Robertson et al., 1992) and each has only a single strain. Genotype VII contains a strain isolated in Sierra Leona in 1988 (SLF88), whereas genotype II is represented by a strain isolated in France in 1979 (CF-53). Although the complete sequence of strain SLF88 was
The mode of HAV evolution was recently studied by two findings. Studies on HAV rates will be required to confirm these clusters that spanned a period of at least 10 years, further strain Cy145) (Nainan et al., 1992) and the other from an African green monkey (Cercopithecus aethiops) imported from Kenya (genotype V, strain AGM27) (Tsarev et al., 1991). The other simian strain was recovered in Atlanta, GA, USA, from a cynomolgus macaque imported from the Philippines (genotype IV, strain Cy145) (Nainan et al., 1991). These simian strains tend to differ genetically as much between themselves as they do from human viruses. Strains AGM27 and JM55 appear to be more closely related to each other than to the Cy145 strain even though they were recovered from different species of primate.

In summary, molecular epidemiologic studies tend to suggest that the degree of genetic heterogeneity among HAV isolates from different areas of the world is higher than expected.

**HAV evolution**

RNA viruses exploit all known mechanisms of genetic variation to ensure their survival (Domingo & Holland, 1997), including mutation and genetic recombination. Their high rates of mutation and replication allow them to move through sequence space at a rate that often makes their DNA-based host’s evolution look glacial by comparison (Worobey et al., 1999).

**Mutation rate and mode of evolution**

An accurate knowledge of the HAV mutation rate could be of assistance in estimating when the different viruses and their genotypes branched off from each other.

The first study attempting to estimate the number of synonymous mutations per synonymous site (Ks) of HAV was reported recently by Sanchez et al. (2003). Ks values from HAV strains isolated from a clam-associated outbreak were 7-5 times lower in the VP1–2A region than the Ks values for GenBank sequences representing geographically and temporally divergent HAV strains. The estimated Ks values varied from 0·038 for VP0 to 0·29 for VP1 (Sanchez et al., 2003) and the Ks value of the P1 region was reported to be close to that obtained for foot-and-mouth disease virus (FMDV).

Since Ks values were not calculated based on monophyletic clusters that spanned a period of at least 10 years, further studies on HAV rates will be required to confirm these findings.

The mode of HAV evolution was recently studied by two groups. Genetic studies using full-length VP1 and capsid sequences have shown the presence of different patterns in the intragenic distributions of synonymous substitutions in the VP1 protein, suggesting that synonymous divergence could be random in the VP1 gene (Costa-Mattioli et al., 2002, 2003). Nevertheless, the distribution of non-synonymous substitutions along the VP1 protein shows a completely different situation, with extremely low rates of substitutions compared to those of synonymous substitutions. Thus, the pattern of divergence observed for HAV VP1 is probably due to selective forces that do not allow amino acid replacements, despite the relative high rates of synonymous substitutions observed all over the gene. Consequently, negative selection appears to be the main force shaping the pattern of non-synonymous substitutions, selecting against most replacements in all protein regions and resulting in a well-conserved protein. These data were confirmed in an extended comparison of the whole HAV capsid region (Sanchez et al., 2003). This is in contrast with the situation found in multiple serotype viruses, such as the haemagglutinin gene of influenza virus (Ina & Gojobori, 1994), the complete capsid region of serotypes A and C of FMDV (Haydon et al., 2001) and the VP3 region of human immunodeficiency virus (Seibert et al., 1995), which are all subjected to positive selection.

Thus, the mode of evolution of HAV appears, at least in part, to explain the presence of only one known serological group of HAV.

**Recombination**

Over the last two decades it has become increasingly clear that many RNA viruses add the capacity to exchange genetic material with one another. Genetic exchange by homologous and non-homologous recombination is a phenomenon that is common among RNA viruses and may lead to hybrid or defective interfering RNA molecules (Lai, 1992; Nagy & Simon, 1997). Despite this general assumption, non-segmented, negative-stranded viruses are not known to recombine.

Genetic recombination was first described in poliovirus (Hirst, 1962; Ledinko, 1963). Later, a number of recombination events was reported between strains that are genetically related (intratypic recombination) and between those that are more distantly related (interotypic recombination).

It is generally accepted that the most biologically relevant mechanism of recombination, in which RNA molecules are joined by an actively copying RNA polymerase that switches from one template to another during negative-strand synthesis is copy-choice (Kirkegaard & Baltimore, 1986; Jarvis & Kirkegaard, 1992). Regardless, however, recombination has also been observed with non-replicating genomes (Gmyl et al., 1999) and, recently, Pierangeli et al. (1999) produced experimental evidence supporting a ‘primer alignment-and-extension’ model in which prematurely terminated RNA molecules hybridize to homologous
sequences in the opposite strand of a replication complex and serve as primers that are subsequently extended.

Most of the recombinant strains described previously were isolated from experimental infection in cell culture under an artificial selective pressure. Kew & Nottay (1984) first reported the isolation of a naturally occurring recombinant poliovirus that contained sequences derived from all three serotypes of poliovirus vaccine strains as a result of two crossovers. Recombination among vaccine and wild-type polioviruses has been reported as a natural means of evolution of poliovirus (Guillot et al., 2000). This has also been observed in enterovirus (Santti et al., 1999). In the case of HAV, genetic exchange among strains had been observed in cell culture (Lemon et al., 1991; Beard et al., 2001; Gauss-Muller & Kusov, 2002) but for many years it was supposed not to occur in nature. This view was subsequently challenged by the report of a case of dual infection of a young childcare provider (AUX-23) with HAV strains belonging to different subgenotypes (de Paula et al., 2003). It is noteworthy that AUX-23 was hired at a childcare centre in which HAV IA and IB were circulating. These particular conditions may have facilitated the double infection. However, the incidence of common HAV mixed infection in single individuals is unknown (C. Niel, Department of Virology, Oswaldo Cruz Institute, Rio de Janeiro, Brazil, personal communication). Indeed, the first HAV recombinant isolated from an infected patient was reported recently (Fig. 3) (Costa-Mattioli et al., 2003). The recombinant strain 9F94 comes from a little girl who was hospitalized in France after a 3 month holiday in Morocco. Accordingly, the putative parental strains SLF88 (genotype VII) and MBB (genotype IB) were also originally isolated in North Africa, a region of high endemicity for HAV infection and one in which multiple genotypes co-circulate (Melnick, 1995). Recombination requires double infection of a single cell. Two conditions were therefore necessary for this to occur: co-circulation of different genotypes in a given geographic area and in a short period of time and infection of a susceptible subject with at least two sources of infection.

Evidence of recombination in polio- and enteroviruses was confined to areas far removed from the capsid region (Kew & Nottay, 1984; Lipskaya et al., 1991; Guillot et al., 2000; Cuervo et al., 2001; Kew et al., 2002). Recently, however, poliovirus (Liu et al., 2000; Martin et al., 2002) and FMDV recombinants in the capsid region (Tosh et al., 2002) have been reported. Studies using HAV capsid sequences revealed intertypic recombination (genotypes IB and VII) in the VP1 capsid protein of a human HAV strain 9F94 (Fig. 3A, B) (Costa-Mattioli et al., 2003). These findings indicate that capsid recombination may play a significant role in shaping the genetic diversity of HAV and, as such, can have important implications for its evolution, biology and control. HAV recombination may serve two opposite purposes: exploration of new combinations of genomic regions from different origins or rescuing of viable genomes from debilitated parental genomes (Lai, 1992; Domingo & Holland, 1997; Domingo et al., 1997; Worobey & Holmes, 1999). Evidence for recombination is most easily demonstrated when the genomic sequences of both parental viruses are known. The frequency and possible implications of HAV capsid recombination events for the generation of pathogenic HAV strains are not clear at present.

**Antigenic divergence**

HAV antigenic studies carried out during the last 20 years have revealed a low antigenic variability (Hollinger & Emerson, 2001). Consequently, mAbs raised to various strains of human HAV have failed to differentiate more than one antigenic type (Crevat et al., 1990). Owing to this high degree of antigenic conservation among human HAV isolates, infection with HAV is likely to confer lifelong immunity that protects against subsequent symptomatic re-infection by the same variant.

Cell culture studies revealed that mutants selected by passage in the presence of a mAb became resistant to neutralization by that antibody (Stapleton & Lemon, 1987; Ping et al., 1988; Nainan et al., 1992). Neutralization escape mutations for the human HAV strain HM-175 were identified at Asp70 and Gln74 of the VP3 protein and Ser102, Val171 and Lys222 of the VP1 protein (Ping et al., 1988; Ping & Lemon, 1992); however, those for strain HAS15 were identified at Pro45, Asp70 and Ser71 of the VP3 protein and Asn104, Lys105 and Gln232 of the VP1 protein (Nainan et al., 1992). The deleted region found in strain Uru-3 contains three amino acids (Ser102, Asn104 and Asn105) reported to be able to induce an escape response in neutralization experiments (Costa-Mattioli et al., 2002). Moreover, these residues align with recognized immunogenic sites in human rhinovirus type 14 (HRV-14) (Rossmann et al., 1985) and poliovirus type 3 (PV-3) (Fig. 4) (Hogle et al., 1985; Minor et al., 1986). This suggests that these residues are part of an immunogenic site that is analogous to neutralization immunogenic sites found in other picornaviruses (HRV-14 and PV-3). Therefore, it is possible that the deletion found in strain Uru-3 may alter the antigenic structure of the virus and suggests that this strain may be the first antigenic variant of HAV found in humans (Costa-Mattioli et al., 2002).

Recently, two other antigenic variants were isolated from an HAV outbreak associated with imported frozen cockles from South America (Sanchez et al., 2002). These naturally isolated mutants were mapped on a discontinuous epitope defined by mAb K3-4C8 and in a linear VP1 epitope of the virus, respectively.

Since HAV strains isolated from South America were reported to contain a highly genetic (Costa-Mattioli et al., 2001a, 2002; de Paula et al., 2002, 2003; Mbayed et al., 2002) and antigenic (Costa-Mattioli et al., 2002; Sanchez et al., 2003) variability, it is possible to speculate that a new serological type could emerge from this geographic region.
Conclusions

Over the last 10 years the molecular epidemiology of HAV has progressed based on comparing partial, discrete sequences of the genome. However, with the advent of PCR, larger regions of the genome can be analysed. Recently, a RT-PCR method able to amplify the complete VP1-encoding regions of HAV strains isolated worldwide was reported. In addition, two analyses have recently been undertaken to examine the relationship between HAV strains by comparing sequences from the entire capsid-encoding region (Costa-Mattioli et al., 2003; Sanchez et al., 2003). The study of other genome areas, and indeed complete genome sequences, may be very useful in determining the frequency of intra- and intertypic recombination in the field and the emergence of new genetic or antigenic variants. In addition, studies on HAV mutation rates would facilitate the understanding of HAV evolution.

ACKNOWLEDGEMENTS

HAV research studies discussed in this review have been supported by the Commission of the European Communities through contract no. IC18-CT98-0378 (DG12-CEOR) and UNESCO through project 00URU606. We thank S.U. Emerson and M. Bidiniosti for critical reading of the paper. M.C.-M. thanks Colin Lister for excellent technical assistance and Nahum Sonenberg for permitting the time necessary to write this review. M.C.-M. is supported by a post-doctoral award from the Canadian Institutes of Health Research (CIHR). A.D.N. was supported by a grant from the Suri D. Dan Memorial Research Foundation.

REFERENCES


Fig. 4. Amino acid alignment of the immunodominant sites of some members of the family Picornaviridae. Neutralization escape mutations are indicated in bold. The common immunodominant site is enclosed in yellow.


Geographic distribution of wild poliovirus type 1 genotypes. Virolgy 160, 311–322.


