For reprint orders, please contact: reprints@futuremedicine.com

Hepatitis A virus evolution and the potential emergence of new variants escaping the presently available vaccines

Rosa M Pintó*[1,2,3], Lucía D'Andrea[1,2,3], Francisco J Pérez-Rodriguez[1,2,3], M Isabel Costafreda[1,2,3], Enric Ribes[1,4], Susana Guix[1,2,3] & Albert Bosch[1,2,3]

[1]Enteric Virus Laboratory, School of Biology, University of Barcelona, Barcelona, Spain
[2]Institute of Nutrition & Food Safety, University of Barcelona, Barcelona, Spain
[3]Department of Microbiology, School of Biology, University of Barcelona, Barcelona, Spain
[4]Department of Cellular Biology, School of Biology, University of Barcelona, Barcelona, Spain
*Author for correspondence: Department of Microbiology, University of Barcelona, Diagonal 645, 08028 Barcelona, Spain ■ Tel.: +34 934 034 621 ■ Fax: +34 934 034 629 ■ rpinto@ub.edu

Review
Future Microbiology

Hepatitis A is the most common infection of the liver worldwide and is fecal–orally transmitted. Its incidence tends to decrease with improvements in hygiene conditions but at the same time its severity increases. Hepatitis A virus is the causative agent of acute hepatitis in humans and belongs to the *Hepatovirus* genus in the *Picornaviridae* family, and it has very unique characteristics. This article reviews some molecular and biological properties that allow the virus to live in a very quiescent way and to build an extremely stable capsid that is able to persist in and out of the body. Additionally, the relationship between the genomic composition and the structural and antigenic properties of the capsid is discussed, and the potential emergence of antigenic variants is evaluated from an evolutionary perspective.

Hepatitis A infection

Differential definition

Approximately 400 years BCE, Hippocrates described an illness characterized by episodes of jaundice that could probably correspond to a viral hepatitis. In ancient China, jaundice illnesses were well recognized as well. However, it is not until the middle of the 20th century that the expression 'infectious hepatitis' was defined and associated with a kind of infectious jaundice that could occur in epidemics. Nevertheless, to be historically accurate, it should be mentioned that the first accurate reference to epidemic jaundice could be that recorded by Cleghorn in *Epidemic Diseases of Minorca 1744 to 1749* [1]. In the 1940s, two separate entities were identified – 'infectious' and 'serum' hepatitis – and since 1965, the major etiological agents (hepatitis A, B, C, D and E viruses) of viral hepatitis have all been identified. While all hepatitis viruses are infectious, the previously used 'infectious' and 'serum' terms refer to the mode of transmission. The 'infectious' type corresponds to hepatitis transmitted through the fecal–oral route, or enteric hepatitis, and includes hepatitis A and E, while the 'serum' hepatitis corresponds to those that are parenterally transmitted, and include hepatitis B, C and D (see Box 1).

Clinical features

Hepatitis A is an acute infection of the liver produced by the hepatitis A virus (HAV). In children under 5 years of age the infection mostly develops asymptomatically or subclinically, while in older children and in adults, the infection usually presents with symptoms [2]. In the latter case, the clinical course of hepatitis A is indistinguishable from that of other types of acute viral hepatitis. The clinical case definition for hepatitis A is an acute illness with a moderate onset of symptoms including fever, malaise, anorexia, nausea, abdominal discomfort, dark urine and jaundice, with elevated serum bilirubin and aminotransferases levels later on [3]. The incubation period of hepatitis A ranges from 15 to 50 days and clinical illness usually does not last longer than 2 months, although 1.5–15% of patients have prolonged or relapsing signs and symptoms for up to 6 months [3–5]. In fact, with the advent of new highly sensitive techniques, a high and long-lasting viremia has been detected even in normal clinical courses [6], with the peak (up to 10^7 genome copies/ml of sera) occurring from just before the beginning to 2 weeks after, the onset of symptoms, and the viremia lasts up to an average of 6 weeks after the start of symptoms [6,7]. By contrast, fecal shedding of the virus reaches its maximum just before the onset of symptoms, at which point the individual is most infectious (see Figure 1 for evolution of viral titers and clinical signs of the disease). There is no evidence of chronicity of the infection. However, the infection may

Keywords
■ codon usage ■ fitness ■ hepatitis A virus ■ quasispecies ■ vaccine-escaping antigenic variant

Future Medicine part of fsg

Box 1. Hepatitis viruses.

- Hepatitis A virus (HAV) is the causative agent of hepatitis A infection. It belongs to the *Picornaviridae* family (30-nm naked particles containing a single-stranded positive RNA genome of approximately 7 kb coding for a single open reading frame [ORF]) and is transmitted through the fecal–oral route
- HBV is the causative agent of hepatitis B infection. It belongs to the *Hepadnaviridae* family (42–47-nm enveloped particles containing a partially double-stranded circular DNA genome of approximately 3 kb coding for four ORFs) and is parenterally transmitted
- HCV is the causative agent of hepatitis C infection. It belongs to the *Flaviviridae* family (40–50-nm enveloped particles containing a single-stranded positive RNA genome of approximately 10 kb coding for a single ORF) and is parenterally transmitted
- HDV is the causative agent of hepatitis D infection. It belongs to the *Deltaviridae* family (30-nm enveloped particles containing a single-stranded covalently closed RNA genome of approximately 1.7 kb, which needs a helper virus to replicate, usually HBV) and is parenterally transmitted
- HEV is the causative agent of hepatitis E infection. It belongs to the *Hepeviridae* family (30–34-nm naked particles containing a single-stranded positive RNA genome of approximately 7 kb coding for three ORFs) and is transmitted through the fecal–oral route
- HFV is the causative agent of a putative type F hepatitis infection. It belongs to the *Flaviviridae* family (40–50-nm enveloped particles containing a single-stranded positive RNA genome of approximately 10 kb coding for a single ORF) and is parentally transmitted. Its existence requires further confirmation
- GB viruses belong to the *Flaviviridae* family (40–50-nm enveloped particles containing a single-stranded positive RNA genome of approximately 10 kb coding for a single ORF) and are parenterally transmitted. Their role in human hepatitis is still unclear
- Transfusion-transmitted viruses belong to the *Circoviridae* family (30–50-nm enveloped particles containing a single-stranded negative DNA genome of approximately 3.8 kb) and are, mainly, parenterally transmitted. Their role in human hepatitis is still unclear

occasionally proceed to a fulminant hepatitis, mainly among patients with underlying chronic liver diseases [2,3,8].

Hepatitis A epidemiology: lower prevalence correlates with increased disease severity

Although hepatitis A is a clinically moderate illness, it still remains the most important acute hepatitis in regard to the number of cases worldwide.

The distribution patterns of hepatitis A in different geographical areas of the world are closely related to their socioeconomic development [2,3]. Hepatitis A infection is highly endemic in developing regions while it is much less frequent in developed regions. This epidemiological pattern has important implications on the average age of exposure and hence on the severity of clinical disease. Since hepatitis A infection induces life-long immunity, severe infections among adults are rare in highly endemic regions where most children are infected early in life. By contrast, in low-endemic areas, the disease occurs mostly in adulthood, mainly as a consequence of traveling to endemic regions, having risky sexual practices or consuming contaminated water or food [3] and hence the likelihood of developing severe symptomatic illness is high.

An epidemiological shift, from intermediate to low prevalence, has been noticed in recent decades in many countries, particularly in southern Europe, including Spain, Italy and Greece [9–11]. Consequently, although the first well-documented hepatitis epidemic occurred in the Mediterranean island of Minorca 250 years ago, the Mediterranean basin as a whole should no longer be considered as an endemic area [2,3,12].

Additionally, some other countries from eastern Europe [13,14] have also described significant declines in the incidence of hepatitis A. Likewise, in several Asian and Latin American countries, a shift from highly to moderately endemic has been described as well [15–17].

Hepatitis A prevention: highly effective vaccines

Inactivated HAV vaccines have been available since the early 1990s and provide long-lasting immunity against hepatitis A infection. The immunity is largely related to the induction of high titers of specific antibodies. Thanks to the existence of a single serotype of HAV, these vaccines are highly efficacious [10,18]. They consist of viruses grown in cell culture, purified, inactivated with formalin and adsorbed into an aluminum hydroxide adjuvant, making their financial cost relatively high. This is the reason why many discrepancies already exist regarding their universal use in massive vaccination campaigns. However, the effectiveness of pediatric mass vaccination programs in reducing the incidence of hepatitis A has been evidenced in several countries [11,19–22]. As a general rule, in low- and inter-mediate endemic regions, where, paradoxically, the severity of the disease is high, vaccination against hepatitis A should be recommended in at least high-risk groups, including travelers to high endemic areas, MSM, drug users and patients receiving blood products. In

addition, the inclusion of hepatitis A vaccines in mass vaccination programs in those countries receiving high numbers of immigrants from endemic countries is also recommended.

Although several attenuated vaccine candidates have also been attempted, due to the successful use of inactivated vaccines, their development has not progressed.

Hepatitis A virus

Life cycle: adaptation at the edge of extinction

Cycle in the host

The only available animal models for HAV replication and disease induction are simians [23,24]. In experimental oral infections performed in monkeys, HAV antigen was detected both early and late after inoculation in cells from the stomach, the small and large intestines [25] and obviously in the liver, bile and stools [25,26], and led to the postulation of an 'enterohepatic cycle' model of replication [3]. This cycle involves an initial oral entry of viruses followed by their adsorption from the stomach/intestine into the bloodstream, to finally reach the liver. After replication in the liver, the virus is released through the biliary canaliculus into the intestine and out of the body in the feces or, alternatively, the enterohepatic cycle can be repeated again. However, the exact mechanism by which the virus enters the bloodstream is still not completely understood. Although a certain degree of intestinal primary replication does occur [27], virus release would probably be mostly into the intestinal lumen since it has been shown in *in vitro* infections of polarized human epithelial cell cultures that virus release is largely restricted to the apical membrane [28]. Transcytosis of the intestinal amplified virus inoculum through M cells, which surround the Peyer's patches in the distal ileum, could be one way to reach the blood vessels, as does happen with polioviruses [29]; however, it has never been demonstrated in HAV. A second model is based on reverse transcytosis of IgA-coated HAV particles through cells of the intestinal epithelium with missorted polymeric immunoglobulin receptor in the apical membrane [30]. This second model, although not proven in intestinal cells but in a kidney epithelium, is based on the fact that an important proportion of HAV particles in the intestine are IgA-coated and that the IgA–HAV complexes are infectious in hepatocytes [31]. Whatever may be the most significant way of crossing the intestinal epithelium, one should expect that viruses, which live at the edge of extinction, must explore and use as many entry ways as possible, and thus both models should be considered compatible rather than excluding each other.

For the development of an infection cycle, HAV has to overcome the challenges posed by the acidic pH of the stomach and the action of intestinal proteases and detergents (particularly biliary salts) during the entry phase, the decoy factors during the viremic phase and, again, the action of proteases and detergents during the exit phase. The selective pressure of pH, proteases and biliary salts calls for the need to shape a highly cohesive and stable capsid. Additionally, HAV replication in polarized hepatocytes with different entry and exit membranes, and thus relying on the blood connection of the enterohepatic cycle to initiate infection of new cells, requires a capsid that, as much as possible, escapes the blood clearance mechanisms of the host. In such a sense, the HAV capsid has evolved to avoid its interaction with the glycophorin A present in the erythrocyte membrane [32].

Cellular cycle

The main target cells for HAV replication are hepatocytes, although crypt cells of the intestine and Kupffer cells of the liver have also been

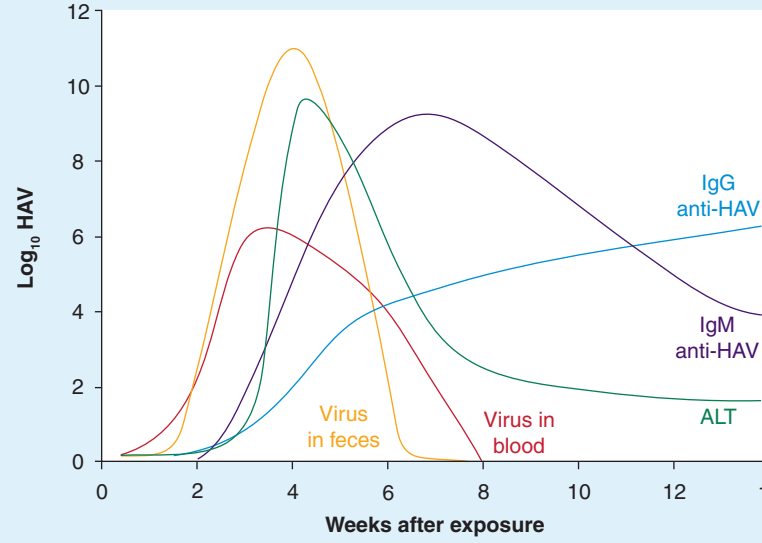

Figure 1. Evolution of viral titers and clinical signs during hepatitis A infection. In those cases of symptomatic infection, the onset of symptoms usually starts 2 weeks after exposure and clinical signs last over 5 weeks. Among clinical signs, increases in serum ALT activity (green) are remarkable. The immunological response includes anti-HAV IgM (purple) and IgG (blue) responses. While the IgM response is usually short-lived, the IgG response is usually life-long. Regarding the evolution of HAV viremia titers (red) develop from before the onset of symptoms to more than 3 weeks after the onset, reaching peaks of approximately 10^6 genome copies/ml of sera. In the same way, shedding of viruses in feces (orange) occurs from before the onset of symptoms to approximately 3 weeks after the onset, with peaks of approximately 10^{11} genome copies/g of feces.
ALT: Alanine aminotransferase; HAV: Hepatitis A virus.

proven positive for HAV antigen [25]. Two different cell entry mechanisms have been proposed for the initiation of the HAV cell cycle. On the one hand, the HAV cell receptor 1 (HAVCR1), which belongs to the T-cell immunoglobulin mucin family [33], has been shown to be a HAV receptor. On the other, the asialoglycoprotein receptor, which binds and internalizes IgA molecules, has also been proposed as a receptor for infectious IgA-coated HAV complexes [31]. Additionally, the IgAλ chain is a specific ligand of HAVCR1, and binding of the IgAλ chain of the IgA–HAV complex has a synergistic effect on the interaction of HAV with the receptor [34]. Thus, free HAV particles may use HAVCR1, while HAV particles in the form of IgA-coated complexes may use both HAVCR1 and the asialoglycoprotein receptor, showing once again the need for viruses to be as adaptable as possible.

The general scheme of the replicative cycle of HAV is very similar to that of the rest of picornaviruses [35]. After interaction with the receptor(s), the uncoating of the positive-sense RNA viral genome contained in the capsid takes place. This process is extremely slow in HAV, at least *in vitro* with cell-adapted strains, with described times of several hours compared with the most common 30-min period in most picornaviruses [36]. Once the RNA is in the cytoplasm, a cap-independent translation of the viral genome (see Figure 2 for the genomic organization) occurs through an internal ribosome entry site within the 5′ noncoding region. The polyprotein is co- and post-translationally processed by the viral protease, and the newly generated RNA-dependent RNA polymerase, as well as several membrane-interacting proteins, assemble with the 3′ end of the genomic RNA to start the synthesis of a negative-strand copy of the viral genome. The negative-strand copy of the genome is used as template for synthesis of multiple new copies of genomic positive-strand RNA, which in turn will be recycled for further RNA synthesis or translated into new proteins. After the assembly of the structural proteins into capsid particles, the positive-strand RNA molecules are packaged, and the newly synthesized virions are secreted across the apical membrane of the hepatocyte into the biliary canaliculus, from which they are passed into the bile and small intestine.

Immune response induced by the virus infection

IgM antibodies are usually detected by the onset of clinical symptoms and, later on, IgA and IgG are also synthesized [37,38]. The anti-HAV IgM response is usually limited to the initial infection and is used as a marker of acute disease. IgA is also induced for a limited period of time and is present in serum and feces, but the role of the secretory immunity in protection against HAV infection appears to be very limited. By contrast, the IgG response is delayed compared with IgM and IgA responses but is long-lasting and confers resistance to reinfection. All of these antibodies are mostly directed against discontinuous epitopes of the immunodominant site of the HAV particle (see below).

HAV-specific, HLA-restricted cytotoxic T cells have also been identified within the liver during acute HAV infection and may play an essential role in viral clearance and the induction of liver damage [3].

HAV replicates quietly, but efficiently, within the liver for several weeks during the incubation period of the disease, in which there is almost no liver damage, and this has been suggested to be due to the ability of the virus to avoid the type-I IFN innate response. In fact, a limited type I IFN response within the liver has been very recently shown in experimentally HAV-infected chimpanzees [39]. Such a response may be the result of a low synthesis of dsRNA viral intermediate molecules necessary for the initiation of the IFN-induction cascade and/or very efficient mechanisms to disrupt the cascade once initiated. Regarding this latter point, it has been shown that, during a HAV infection of cultured cells, two virus nonstructural protein intermediates, 3ABC and 3CD, cleave the MAV protein, which is involved in the RIG-I receptor pathway, and the TRIF, which is involved in the TLR3 pathway, respectively [24,40]. What is really striking is that compared with HCV, which also disrupts the signaling pathway of type I IFN synthesis, HAV is much more efficient in doing so, suggesting the occurrence of additional mechanisms in HAV to escape from this antiviral cellular response [39]. Thus, HAV infections should be considered a distinctly unique archetype in virus–host interactions.

Unique molecular features of HAV: the key for a quiescent replication & an extremely stable phenotype

Several unique molecular characteristics of the HAV genome make it entirely distinguishable from other picornaviruses (see Box 2), the first of which is the structure of the 5′ noncoding region, which contains the internal ribosome entry site (IRES). The HAV IRES is unique among picornaviruses and constitutes the type

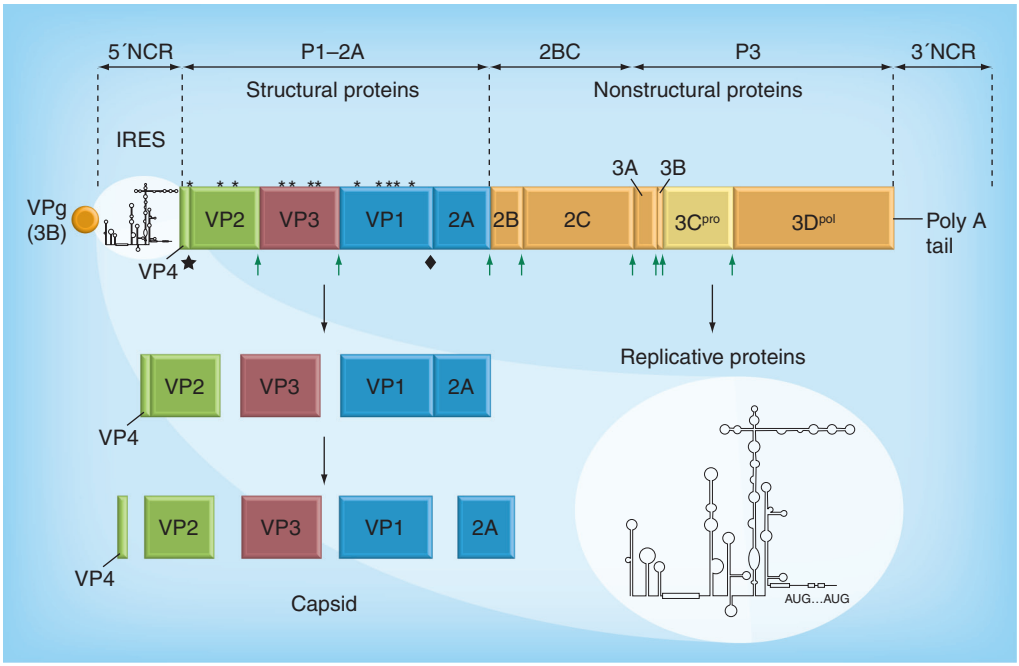

Figure 2. Hepatitis A virus genomic organization and expression. The hepatitis A virus genome is composed of an RNA molecule that can be directly translated (positive RNA genome) and which contains a single open reading frame encoding a polyprotein. Strategically located clusters of rare codons of the capsid-coding region (asterisks) have been proposed to play an essential role in capsid folding through the control of translation speed. The polyprotein is autoprocessed by the viral protease $3C^{pro}$ (yellow box) at all cleavage sites (green arrows), with the exception of two cleavages made by proteases other than the viral $3C^{pro}$: a yet-to-be-identified cellular protease (diamond), which releases the 2A fragment from the capsid, and an unknown proteolytic activity (star), which participates in the last capsid maturation process.
IRES: Internal ribosome entry site; NCR: Noncoding region.

III model [41,42], which shows a very low efficiency in directing translation [43] compared with other picornavirus IRESs.

Second, HAV possesses a complex internal stem-loop near the 5′ end of the polymerase-coding sequence that functions as a *cis*-acting replication element, which is distinguishable from other picornaviral cis-acting replication elements by its relatively large size and the length of its top loop [44].

Third, HAV encodes only protease, 3C, while other picornaviruses code for additional proteases such as the L protease in the genus *Aphtovirus*, or the 2A protease in the *Enterovirus* and *Rhinovirus* genera [45]. L and 2A proteases, when present, play a crucial role

Box 2. Unique molecular characteristics of hepatitis A virus compared with the rest of the picornaviruses.

- Hepatitis A virus (HAV) has a type III internal ribosome entry site, which is highly inefficient at directing translation
- HAV has a *cis*-acting replication element in the 5′ end of the polymerase-coding region, which is larger than other picornaviral *cis*-acting replication elements
- HAV encodes a single protease, 3C, which is not involved in the processing of the eIFG4 translation factor. In fact, HAV needs an intact eIFG4 factor for its own translation and consequently does not induce the protein cellular shut-off by this common method. Thus, HAV competes poorly with the cell for resources such as tRNAs
- HAV has significant codon usage and CpG biases
- The codon usage of HAV is not only highly biased but also highly deoptimized with respect to the cellular codon usage
- HAV has a great ability to inhibit the cellular IFN response
- HAV replicates in a quite quiescent mode
- The 2A protein of HAV is necessary during capsid morphogenesis, particularly for pentamer formation. The removal of 2A from the capsid is performed by a cellular protease(s)
- The capsid of HAV seems to be extremely smooth, and the absence of a pit or canyon region has been highlighted in cryo-electron microscopy images
- HAV's nucleotide diversity in the capsid-coding region is similar to that of other picornaviruses; however, the amino acid variability is much lower and HAV exists as a single serotype

in the primary cleavages of the viral polyprotein, while in those genera lacking these proteases, such as *Hepatovirus* and *Paraechovirus*, both primary and secondary cleavages are conducted by the 3C protease. But what is most important is that these additional proteases are involved in the cellular protein shut-off induction [45]. Since picornaviruses utilize a mechanism of translation that is cap-independent and IRES-dependent, the inhibition of nonessential, cap-dependent cellular translation could be advantageous to the virus. In doing so, the cellular translation machinery is utilized almost exclusively for the production of viral proteins [46]. An early event preceding the shut-off of host cell protein synthesis is the cleavage of the cellular translation initiation factor eIF4G, and evidence exists supporting the idea that the enzymes responsible for such a cleavage are 2A protease in enteroviruses and rhinoviruses, and L protease in aphtoviruses [46]. An immediate consequence of the lack of any of these proteolytic activities in HAV is its inability to induce cellular shut-off, which otherwise is directly related to its requirement for an intact uncleaved eIF4G factor for the formation of the initiation of the translation complex [47,48].

What has been described up to now indicates that HAV must inefficiently compete for cellular translational machinery and consequently it has evolved a unique translation strategy. This highlights the third difference between HAV and other picornavirus members: codon usage. The preference of one codon over another synonymous codon to specify an amino acid is termed 'codon usage bias'. HAV presents a higher codon usage bias compared with other members of its family, which is characterized by the adaptation to use abundant and rare codons [49]. But what is more surprising is that the HAV codon usage has evolved to be complementary to that of human cells, never adopting as abundant codons those abundant for the host cell, and even in many instances using the latter as rare codons. A consequence of this special codon bias is an increase in the number of rare codons used. Overall, this increment is the result of the addition to the cellular rare codons, also used as rare by the virus, of those most abundant cellular codons. The rationale is that the cognate tRNAs of the codons abundantly used by the cell are unavailable for the virus and thus these codons are used by the virus at low frequency. In summary, HAV has a naturally deoptimized codon usage. The role of rare codons in the control of translation speed has been largely documented [50–52], in the sense that clusters of rare codons would induce a transient stop of the translational complex in order to seek a suitable tRNA present at a very low concentration among the pool of tRNAs. A function of these ribosome stallings has been suggested to be the assurance of the proper folding of the nascent protein [53–55]. Such a function has also been postulated for HAV, where highly conserved clusters of rare codons strategically located at the carboxy-ends of the structured elements of the capsid-coding region have been reported [49]. The critical role of HAV codon usage, and particularly of these clusters of rare codons of the capsid-coding region, has been shown in functional genomic studies during the process of adaptation of HAV to conditions of artificially induced cellular shut-off (replication in the presence of actinomicin D) [56]. An overall change in the codon usage of the capsid region was necessary to regain viral fitness, after an initial fitness loss, during the adaptation to actinomycin D, with a clear re-deoptimization with respect to the cellular codon usage, and particularly affecting the rare codons located at the aforementioned strategic positions of the capsid [56]. This mechanism of adaptation to the cellular shut-off proves translation kinetics – that is, the right combination of codons (common and rare) that allows a regulated ribosome traffic rate ensuring proper protein folding – as a driving selective force of the HAV codon usage at the capsid region (**Figure 3**). The accuracy of HAV capsid folding may contribute to the extremely resistant phenotype of HAV to high temperatures, acid pH and detergents [23,38], and to its high persistence in the environment [57,58] that consequently enables transmission by contaminated foods and drinking water [59–64]. In fact, those populations adapted to replicate in the presence of actinomycin D, which have re-deoptimized their codon usage, show an important decrease of resistance to all of these factors [Costafreda MI *et al.* Manuscript in Preparation]. Additionally, this special codon usage may answer the intriguing question of why HAV has evolved to have a highly inefficient IRES, since a very efficient translation machinery recruitment to the IRES combined with many ribosome stalls at the beginning of the coding region would not be a very convenient arrangement, but rather a very ineffective process hijacking many ribosomes on a few RNA molecules.

The highly deoptimized codon usage of HAV has also been interpreted as a subtle strategy to avoid, as much as possible, competition for

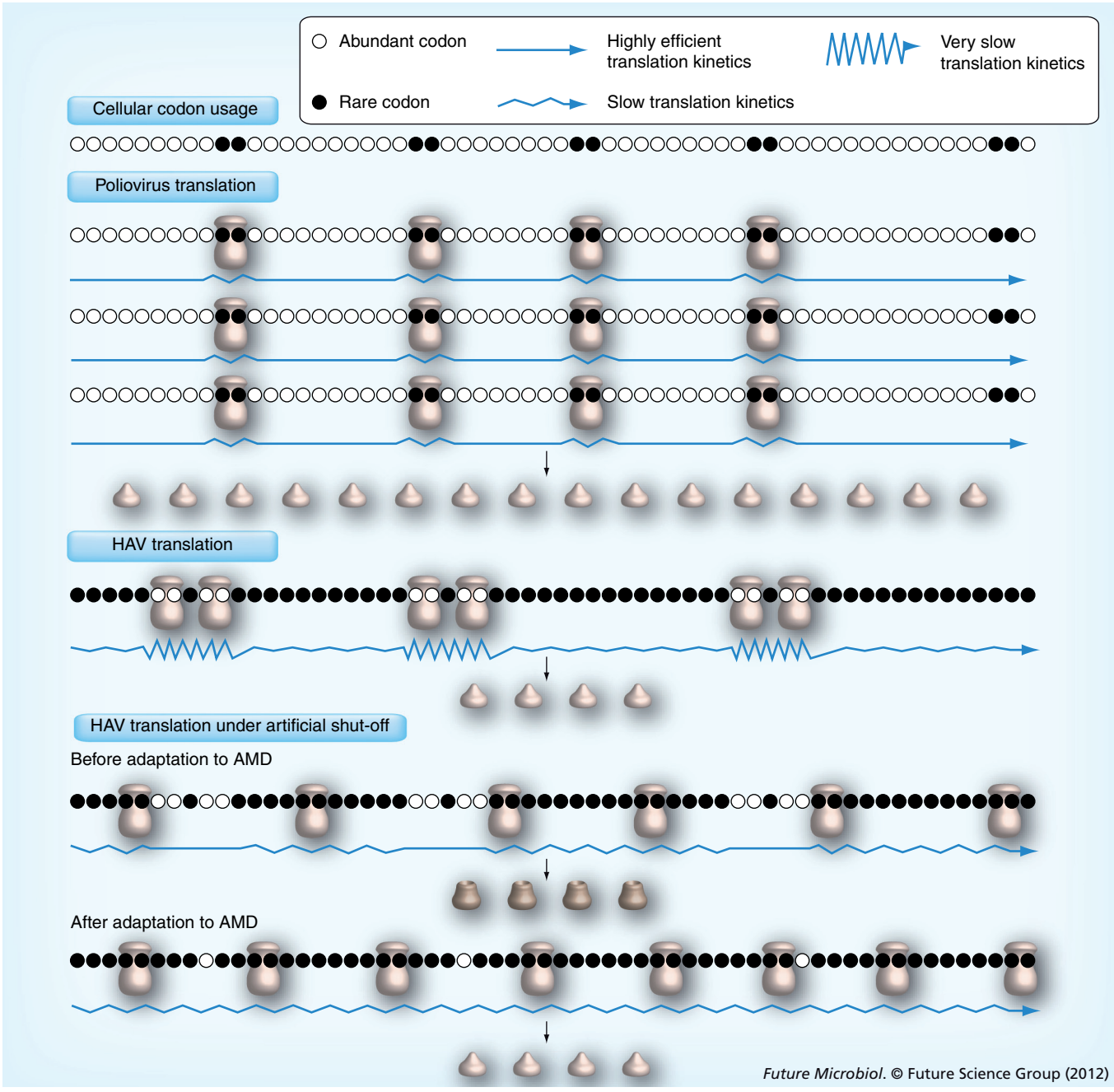

Figure 3. Fine-tuning translation kinetics selection in the capsid-coding region of hepatitis A virus. Codon usage bias is the result of a sum of pressures including mutation bias and diverse selection forces. Among these selection forces, at least two are translation-related processes, one of which is translational selection or the adaptation of the codon usage to the tRNA pools for a highly efficient and accurate rate of translation. In the virus world, this implies the optimization of the virus codon usage to that of the host cell, as happens in poliovirus. The other translation-related process is fine-tuning translation kinetics selection, or the right combination of common and rare codons to regulate the ribosome traffic rate and thus to ensure proper protein folding. Ribosome stallings at rare codon positions allow sequential folding, and consequently the occurrence of clusters of strategically located rare codons is essential. The codon usage of HAV is highly biased and highly deoptimized with respect to that of the cellular host. Since HAV does not possess mechanisms to induce cellular shut-off, it may compete poorly for tRNAs; thus, it has adapted to use codons that are not highly used by the cell as abundant codons, and codons that are used as rare, but also those that are abundantly used by the cells, as rare codons. Additionally, rare codons of the capsid-coding region, but not those of other genome regions, such as the polymerase-coding region, are strategically located at the carboxy-borders of the predicted α-helices and β-sheets. When cellular shut-off is artificially induced with AMD (with the associated changes in tRNA pools), HAV shows a fitness loss. HAV thereafter adapts to grow in the presence of AMD through an overall change of the codon usage of the capsid-coding region towards a re-deoptimization. Fitness variations correlate with changes of the capsid stability and antigenicity, which in turn may reflect modifications of capsid folding due to different kinetics of translation associated with changes in codon usage.
AMD: Actinomycin D; HAV: Hepatitis A virus.

the cellular tRNAs in the absence of a precise mechanism of inducing shut-off of cellular protein synthesis [49] and, additionally, may also contribute to a slow genome replication since all the viral proteins are encoded in a single open reading frame and thus the expected low protein synthesis may also affect those proteins involved in RNA replication (see Figure 2). Thus, a low efficiency of viral replication will probably contribute to a low concentration of dsRNA intermediates. In fact, only 10% of HAV-infected liver cells were positive for dsRNA in experimentally infected chimpanzees [39], a rather low proportion compared with what is observed in HCV infection [65]. Additionally, as stated above, HAV has developed mechanisms to prevent or reduce cellular antiviral responses [40,66,67]. The delicate balance between the levels of dsRNA intermediates and some viral intermediate proteins (3ABC and 3CD; see above) might modulate the inhibition of the antiviral cell responses.

Other important differences exist between HAV and other picornaviruses at the morphogenetic/structural level. The role of both ends (amino-VP4 and carboxy-2A) of the capsid polyprotein in the virion assembly is still controversial [68], and while there is no agreement on the requirement of VP4 for the maturation of pentamers into capsids, a complete consensus exists on the necessity of 2A for pentamer formation [23]. The ulterior removal of 2A in the mature virion must be performed by a host cell protease [69,70], although the mature 2A protein has never been identified directly in infected cells. The x-ray crystallographic structure has not yet been solved, due to the low viral yields obtained by *in vitro* replication. However, 3D images of HAV produced by cryoelectron microscopy [Cheng H, Unpublished data] [23] have revealed important data, the most intriguing being the lack of a well-defined canyon around the fivefold axis of symmetry. The receptor-binding residues of many picornaviruses are located in the pit region [35,71]. However, whichever is the HAV receptor, the capsid region involved in receptor binding remains to be elucidated. By contrast, the capsid region interacting with the glycophorin A of human erythrocytes is indeed located around the putative pit area [32]. The capsid structure, however, is such that it will only tolerate this interaction in acidic conditions, being impaired at neutral biological conditions, enabling escape from erythrocyte attachment and thus constituting an advantage, as stated above.

Quasispecies dynamics of evolution & virus fitness

Viral genetic variability results from the universal mechanisms of mutation, recombination and genome segment reassortment, all of which are replication-dependent. Since virus populations replicate at exceptionally high rates, they may be extremely variable. All this is particularly critical in RNA viruses, since they rely on error-prone polymerases lacking proofreading activity, which leads to complex mutant genome populations or quasispecies. Viral quasispecies are dynamic distributions of nonidentical but closely related viral genomes subjected to a continuous process of genetic variation, competition and selection, and which act as a unit of selection (reviewed in [72]).

RNA viruses have the capacity to quickly explore large regions of sequence space thanks to their high mutation rates, which are in the range of 10^{-3}–10^{-5} substitutions per nucleotide copied [72]. However, their genome size and diverse selective constraints limit the diversity that is actually expressed [73].

Genetic & antigenic diversity of the virus: structural & biological constraints limit the antigenic variability

HAV, as an RNA virus, occurs as a swarm of mutants or quasispecies [74]. The nucleotide diversity is similar to that of other picornaviruses [49], and allows its differentiation into several genotypes and subgenotypes. The *VP1X2A* junction region is still the genomic region most in use worldwide to study the genetic diversity of HAV [75]. Six genotypes, whose genetic distance in the *VP1X2A* region – a highly variable genomic region – is >15% nucleotide variation, have been defined [76]. Three out of these six genotypes (I, II and III) are of human origin, while the others (IV, V and VI) are of simian origin. Genotypes I, II and III contain subgenotypes defined by a nucleotide divergence of 7–7.5%. Genetic diversity of HAV is evidenced by the emergence of new subgenotypes [77]. Genotypic characterization may be highly relevant to tracing the origin of outbreaks.

Despite this nucleotide variability, the diversity is limited at the amino acid level and only a few natural antigenic variants have been isolated [63,78], suggesting the occurrence of severe structural and biological constraints of the capsid that would prevent the emergence of new serotypes. Consequently, a single serotype of HAV exists [38], which represents another striking difference from other picornaviruses.

Table 1. *In vitro* and naturally isolated variants of the main antigenic sites of hepatitis A virus.

Capsid epitopes	Substituted residues in mAb-resistant mutants[†]	Substituted residues in natural antigenic mutants[‡]
Immunodominant site/multiple mAb binding site[§]	VP3: P65S, D70A, D70H, D70N, D70Y, S71P, Q74R VP1: S102L, N104D, K105R, V171E, A176D, Q232E	VP3: V72I VP1: V166G, V171A, Y181S, R189T, A280V, A280E
Glycophorin A binding site/H7C27 mAb binding site	VP1: G217D, K221E, K221M	None
4E7 mAb binding site	None	None

[†]Data taken from [30,76,77].
[‡]Data taken from [60,75].
[§]The immunodominant site is defined by most of the existing mAbs against hepatitis A virus with the exception of the mAbs H7C27 and 4E7.
mAb: Monoclonal antibody.

The antigenic structure of the HAV capsid is defined by three main epitopes (Table 1). The immunodominant site composed of closely clustered epitopes is defined by two major groups of escape mutants that include residues 70, 71 and 74 of VP3 and residues 102, 171 and 176 of VP1 [79,80]. A second epitope is the glycophorin A binding site, represented by mutants around residue 221 of VP1 [79,32]. Finally, there is a third and still undefined epitope, represented by escape mutants to the 4E7 monoclonal antibody (mAb).

Apart from the expected structural constraints due to those amino acid residues playing critical roles in capsid folding, a certain contribution of the codon usage to the low antigenic variability of the HAV capsid has also been suggested [81]. A total of 15% of the surface capsid residues are encoded by rare codons. These rare codons are highly conserved among the different HAV strains [49] and their substitution is negatively selected even under specific immune pressure [81]. Many of these capsid residues encoded by rare codons are surface exposed and located near or at the epitope regions, and this negative selection would prevent the emergence of antigenic variants. The need to maintain the clusters of rare codons responds to the requirement for proper capsid folding, which is controlled through the kinetics of translation (see above), and it is quite unlikely that a nucleotide substitution would give rise to a new codon of similar rarity and a compatible amino acid. Also, some biological constraints have been proposed to contribute to the low antigenic variability of HAV [82]. mAb-resistant mutants (MARs) representing the aforementioned epitopes show a completely different fitness pattern [81]. While MARs of the H7C27 mAb (glycophorin A binding site) show a similar fitness to that of wild-type viruses in *in vitro* assays, those of the K34C8 mAb (immunodominant site) show a significantly lower fitness than wild-type viruses. By contrast, among the few antigenic variants isolated from patients, only representatives of the immunodominant site are available (Table 1) [63,76,78,83]. This discrepancy may be explained by taking into consideration the biological constraints imposed by the enterohepatic cycle described above and the need to escape from erythrocyte attachment, since a conformational change in the glycophorin A binding site may result in an increased erythrocyte binding capacity, which is low in wild-type viruses at physiological conditions [32]. Avoiding blood clearance – that is, the removal of viruses from the fluid compartment of blood – may constitute an advantage for a viremic infectious agent whose target organ is the liver, contributing to the final fitness outcome *in vivo*.

Potential emergence of variants escaping the protection of the available vaccines: the need to complete the vaccination schedules in the HIV-positive MSM group

The emergence of a new serotype requires extensive substitutions in the capsid that seem quite unlikely to occur in a virus with such severe genomic, structural and biological constraints. However, the emergence of new variants is plausible if virus populations are forced through bottleneck conditions such as immune selective pressures.

Recently, the isolation of several natural antigenic variants of the immunodominant site during an outbreak of hepatitis A in the MSM community of Barcelona (Spain) has been described [78]. Of particular interest is one of these variants with two amino acid substitutions at positions 166 (V to G) and 171 (V to A) of the capsid protein VP1, just in the core of the immunodominant site (Table 1). MAR mutants around

this location show a phenotype of complete resistance to the available vaccines and a lower fitness, compared with wild-type viruses, in normal conditions, but higher fitness in the presence of antibodies [74]. A total of 4% of the whole affected group of patients had been vaccinated, and among these, 62% were HIV-positive. Additionally, among those who were vaccinated, only 12% had followed the complete schedule of vaccination, while 88% had received only one dose. These are the optimal conditions for the selection of those variants that, despite their lower fitness, are able to escape the neutralization effect of antibodies.

The immunocompromised population, particularly HIV-positive individuals, have an impaired immunological response to HAV vaccines, which means that they may have lower concentrations of anti-HAV IgG in sera than healthy individuals after vaccination and as such they require additional vaccine doses [84,85]. The normal schedule for HAV vaccination usually requires two doses. Between both doses, the level of IgG in immunocompromised patients is very low and insufficient to completely neutralize the virus, particularly if the input virus is high. Conditions of high input virus occur in some risky sexual practices, bearing in mind that the HAV titer in feces can be as high as 10^{11} particles/g, particularly in HIV-positive individuals at 2 weeks after the onset of symptoms [6]. It cannot be forgotten that the peak of viruses in feces occurs even before the onset of symptoms, and thus the titers are even higher during the prodromic phase (Figure 1). Replication of the non-neutralized viruses in the presence of low concentrations of specific IgG may contribute to selecting, among the swarm of mutants generated by the quasispecies dynamics of replication, those variants that are resistant to the effect of the vaccines [86]. In other situations in which the level of input virus is low, such as in foodborne transmission, even low concentrations of IgG are able to neutralize virus infection [3].

Whether these newly emerged strains will further circulate worldwide or disappear is yet to be elucidated. However, what is imperative is to target the MSM community with more effective information on risky sexual practices and vaccination programs. Additionally, efforts should be made, particularly among HIV-positive MSM, to completely accomplish the HAV vaccination schedule to avoid the potential emergence of antigenic variants as much as possible.

What else can bioinformatics tell us about HAV genomic composition & evolution?

Bioinformatics and computational biology generate lots of information about the genomic composition of organisms and open the possibility to infer their evolution.

The degeneracy of the genetic code gives an organism the flexibility to encode a given protein sequence in its genome in an extraordinarily large number of ways [87]. In addition to codon bias, there are other genome biases such as codon-pair bias – that is, the preference or avoidance of certain codon pairs and dinucleotides biases, or in other words, the preference or avoidance of certain dinucleotides. The molecular and genetic mechanisms underlying these biases may be diverse, and it is likely that a combination of all of them is the actual evolutionary force acting within each genome. In the virus world, at least four mechanisms are envisaged, three of them in common with all organisms [88]; first, mutational bias and the specific nucleotide composition; second, translation selection, or the optimal codon adaptation to the tRNA pool in order to get a highly efficient and accurate translation; third, fine-tuning translation kinetics selection, or the right combination of codons to allow a regulated ribosome traffic rate that temporally separates protein folding events, ensuring 'beneficial' and avoiding 'unwanted' interactions within the growing peptide [89]; and fourth, selection for mechanisms to escape the antiviral cell responses.

HAV shows a significant codon usage bias, does not show codon-pair bias and has a very important dinucleotide bias [87,90]. All of these biases are obviously inherent to genomic compositional constraints [90,91], but probably not exclusively. The comparative analyses of the relationship between the effective number of codons and the G+C content at the third codon position of the HAV genome and other picornaviruses, RNA hepatitis viruses or other RNA viruses (Figure 4) allows us to conclude that the distance from the theoretical curve (compositional constraints as exclusive forces of codon usage bias) of HAV is the highest among all viruses tested, indicating that although compositional constraints play an important role in shaping its codon usage, other forces may also be involved. The percentage of deviation from the theoretical curve for HAV is 17.2%. Only hepatitis E virus shows a similar degree of deviation (16.7%) and is then followed by HIV (16%), Dengue virus (15%), yellow fever virus (12%), rabies and West Nile

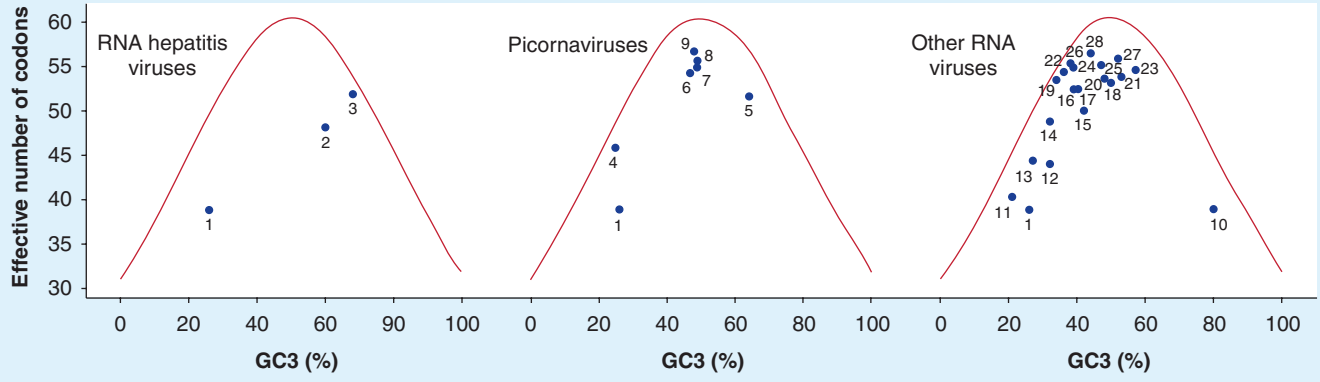

Figure 4. Influence of compositional constraints on codon usage bias of hepatitis A virus and other RNA viruses. Hepatitis A virus shows a significant codon usage bias, which is obviously inherent to genomic compositional constraints, but probably not exclusively. The comparative analyses of the relationship between the effective number of codons and the G+C content at the third codon position (GC3%) of RNA hepatitis viruses (1: hepatitis A virus; 2: hepatitis E virus and 3: HCV), picornaviruses (4: rhinovirus; 5: foot and mouth disease virus; 6: poliovirus; 7: encephalomyocarditis virus; 8: coxsackievirus A and 9: enterovirus 71) and other RNA viruses (10: rubella virus; 11: rotavirus; 12: HIV; 13: respiratory syncytial virus; 14: hantavirus; 15: Dengue virus; 16: vesicular stomatitis virus; 17: influenza A virus; 18: yellow fever virus; 19: Marburg virus; 20: rabies virus; 21: West Nile virus; 22: mumps virus; 23: western equine encephalitis virus; 24: astrovirus; 25: measles virus; 26: ebola virus; 27: Japanese encephalitis virus and 28: norovirus) allows us to conclude that the distance from the theoretical curve (compositional constraints as exclusive forces of codon usage bias) of HAV is the highest among all viruses tested, indicating that although compositional constraints play an important role in shaping its codon usage, other forces may also be involved.

virus (11%) and poliovirus (10%). The average deviation for the other viruses is in the range of 2–9%. How significant these deviations are is difficult to define. However, among the viruses with higher deviations, fine-tuning translation kinetics selection and translation selection have been proven by functional genomics to contribute to the delineation of codon usage in HAV [56] and HIV, respectively [92]. Consequently, those conclusions on the origin of codon usage bias based exclusively on computational analyses should be interpreted with caution.

HAV shows a very low occurrence of CpG dinucleotide [87,90]. This low CpG occurrence cannot be explained by the overall low G+C content in the HAV genome (37%) since the dinucleotide GpC is eight-times more frequent [87]. The cytosine in CpG is the primary target of cellular DNA methylation. Methylated cytosine, in turn, is prone to deamination, creating a thymine in the process, which results in a transition mutation in the next round of DNA replication and systematic loss of cytosine over evolutionary time frames. In somatic cells, the majority of the remaining genomic CpGs outside active gene promoter regions are in fact methylated [93]. As a result of the scarcity of freely accessible, non-methylated CpG in the cell (that is not bound by cellular CpG-binding proteins or components of the transcriptional machinery), DNA sequences rich in unmethylated CpG, if encountered by the cellular machinery (e.g., as part of a pathogen's genome), are recognized as foreign, and innate defense pathways are activated. While such CpG-mediated innate immune recognition is well established for DNA pathogens [94], evidence for a similar mechanism of CpG in RNA genomes is very sparse [95]. However, it can be speculated that the elimination, as much as possible, of CpG RNA motifs in the HAV genome is an evolutionary mechanism to avoid antiviral responses. This would again fit with the quiescent HAV replicative cycle to prevent triggering the cell antiviral responses.

Mutational pressure and different selective forces shape the HAV genome composition and contribute to its codon and dinucleotide biases, which in turn may play an important role in the virus replicative cycle. Thus, the endless question of whether mutation or selection were first remains to be answered.

Conclusion

Although HAV as an RNA virus exists as a swarm of mutants or quasiespecies, due to important structural and biological constraints, its phenotypic variability is rather low and a single serotype exists. Nevertheless, this situation should not be misunderstood; indeed, HAV, similar to any other RNA virus, continuously mutates, and although a long history of evolution has selected a constrained phenotype, new pressures may induce the emergence of viruses adapting to novel situations. An example of such

Executive summary

Hepatitis A infection

- Hepatitis A is an acute infection of the liver by the hepatitis A virus (HAV), which is transmitted through the fecal–oral route.
- Infection is mostly asymptomatic among children under 5 years of age, while in older children and adults, it proceeds with symptoms.
- During the prodromic phase, virus titers in feces are very high, while the peak of viremia extends from the prodromic phase to 2 weeks after the onset of symptoms.
- Hepatitis A infection never develops to a chronic hepatitis and induces a long-lasting immunity.
- Highly effective inactivated vaccines are available today although, due to the difficulties in growing the virus, they are quite expensive.

Hepatitis A virus

- HAV belongs to the *Picornaviridae* family. Its replicative cycle has been studied *in vitro* with cell-adapted strains and, although it shares most of its replicative steps with the rest of picornaviruses, it is characterized by highly prolonged replication times.
- HAV enters the host by the oral route and, to reach the target organ, the liver, an enterohepatic cycle has been proposed. The input virus could replicate in the intestine, cross it into the bloodstream and reach the hepatocytes. After liver replication, progeny viruses would be released into the intestine through the biliary canaliculus and excretion in the feces and/or reinitiation of the cycle would proceed.
- HAV can use different mechanisms for its entry process: direct virus binding to the hepatitis A cellular receptor 1; indirect virus interaction with the asialoglycoprotein receptor, which binds and internalizes IgA molecules and consequently IgA-coated HAV; or direct and indirect binding to the hepatitis A cellular receptor 1 through the dual interaction of the virus particle and the IgAλ chain of an IgA–HAV complex. Thus, both virus-free particles and IgA-coated particles are able to infect susceptible cells.
- HAV posseses a highly inefficient internal ribosome entry site to initiate translation, which constitutes a unique type III internal ribosome entry site, cannot induce the cellular shut-off and requires an intact cellular translation initiation factor eIF4G. This means that the tRNAs, among other resources, will be limiting for the virus.
- HAV has a highly biased codon usage that is also highly de-optimized with respect the cellular codon usage. This means that HAV uses codons that are not highly used by the cell as abundant codons and uses as rare codons those that are rare for the cell but also those that are highly used by the cell and whose pairing tRNAs will not be easily available for the virus. In the capsid-coding region, there are clusters of residues encoded by rare codons strategically located in the carboxy-borders of the highly structured elements and folded near or at the surface epitopes. One underlying molecular mechanism explaining why the clusters of rare codons are kept is the need to control the speed of translation. A correct combination of rapidly translated codons (those pairing with abundant tRNAs) and slowly translated codons (those pairing with scarce tRNAs) allows a regulated ribosome traffic pathway that is compatible with very precise protein folding, giving highly cohesive and stable capsids.
- HAV exists as a single serotype, another striking difference from the rest of picornaviruses.

Quasispecies dynamics of evolution & virus fitness

- HAV as an RNA virus replicates as a complex mutant genome population, or quasispecies, which is subjected to a continuous process of genetic variation, competition and selection.
- Genome size and diverse selective constraints limit the diversity that is actually expressed.
- Several structural constraints, both at the amino acid and codon usage levels, may contribute to the low variability of the HAV capsid. The clusters of rare codons needed to control the speed of translation are highly conserved among different HAV strains and their replacement is negatively selected even under immune pressure. Thus, the need to keep the clusters of rare codons, with the aim to preserve a proper capsid folding, contributes to the low antigenic variability of HAV and does not enable the emergence of a new serotype.
- The occurrence of an enterohepatic cycle calls for the need for a highly stable capsid to acid pH and biliary salts, which is able to escape blood clearance mechanisms. HAV has evolved to have the ineffective binding to the glycophorin A decoy factor present in the erythrocyte membrane, and the region involved in such low interaction must be preserved.
- With so many constraints, the fitness landscape of HAV is predicted to be narrow and antigenic diversity is expected to be low. However, the ability of RNA viruses to explore the sequence space is enormous and, given the existence of selective pressures, such as the immune pressure, the emergence of new antigenic variants should not be neglected.
- To avoid as much as possible the emergence of antigenic variants escaping the effect of the available vaccines, the replication of the virus in the presence of low concentrations of neutralizing antibodies should be avoided. Thus, the complete vaccination schedule is a must, particularly among the HIV-positive MSM group.

What else can bioinformatics tell us about HAV genomic composition & evolution?

- HAV presents a highly biased codon usage but also a highly biased dinucleotide composition. HAV has an extremely low proportion of the CpG dinucleotide.
- Bioinformatics tells us that HAV genomic composition plays an important role in such biases. However most bioinformatic tools are not sensitive enough to provide information on the impact of different selective pressures in delineating such biases. Among them, fine-tuning control of the translation kinetics and avoiding the antiviral cell responses may play a significant role as mechanisms underlying the codon usage and dinucleotide bias, respectively.

a situation is to give the virus the opportunity to replicate in the presence of antibodies, which will probably end up in the selection of resistant mutants. The available vaccines are highly effective; however, to avoid the emergence of resistant viruses, incomplete vaccination schedules among the immunocompromised population, and particularly in the HIV-positive MSM group, must be totally avoided.

Future perspective

The available inactivated vaccines against HAV are highly effective. This is indeed very good news for public health, but it has resulted in a lack of interest in the biology of HAV by the scientific community.

HAV has traditionally been considered an invariable virus. Nothing could be further from the truth. As an RNA virus, it is continuously evolving and exploring the sequence space. However, how wide the mutant spectrum could be depends on the fitness landscape, and indeed HAV suffers many structural and biological constraints, which will narrow this landscape. Additionally, although the mutation frequency of HAV is similar to that of other RNA viruses, the mutant repertoire is narrower, since the population size is small due to a slow replication ratio. In fact, the quasispecies theory predicts that in populations of RNA viruses with rapid replication, the narrower the fitness landscape, the higher the number of individuals in the progeny that will be lost due to their much lower fitness than the parent. Thus, given the predicted narrow landscape of the capsid-coding region of HAV, it would not be a good investment to have a rapid replication phenotype. Overall, in such a context, it could be expected that HAV will forever exist as a single serotype. However, since the sequence space is almost infinite, the emergence of new viruses adapting to new conditions remains an open possibility. From the public health point of view, molecular epidemiology surveys to control the potential circulation of newly arising strains escaping the effect of the available vaccines should be performed, particularly among the HIV-positive MSM group.

From the biotechnology point of view, the new insights that genomic information has produced must be applied in the postgenomic era to provide new methods to obtain more affordable HAV reagents through the selection of a virus bearing a more efficient IRES and having an optimized genomic composition and codon usage, which could allow a higher antigen production.

Financial & competing interests disclosure

Research in the Enteric Virus Laboratory of the University of Barcelona is supported by Spanish Ministry of Science and Innovation (projects BIO2008-01312, BIO2011-23461 and CSD2007-00016), Generalitat de Catalunya (project 2005SGR00966 and Biotechnology Reference Network) and European Commission (project Food-4B-2005-36306).

The authors would like to disclose that the Enteric Virus Laboratory of the University of Barcelona owns intellectual property in the field of hepatitis A virus genetics and hepatitis A virus antigen production. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as:
- ▪ of interest
- ▪▪ of considerable interest

1. Cockayne EA. Catarrhal jaundice, sporadic and epidemic, and its relation to acute yellow atrophy of the liver. *QJM* 6, 1–29 (1912).
2. Previsani N, Lavanchy D, Siegl G. Hepatitis A. In: *Viral Hepatitis Molecular Biology, Diagnosis, Epidemiology and Control.* Mushahwar IK (Ed.). Elsevier, The Netherlands, 1–30 (2004).
3. Cuthbert J. Hepatitis A: old and new. *Clin. Microbiol. Rev.* 14, 38–58 (2001).
4. Sjogren MH, Tanno H, Fay O *et al.* Hepatitis A virus in stool during clinical relapse. *Ann. Intern. Med.* 106, 221–226 (1987).
5. Glikson M, Galun E, Oren R, Tur-Kaspa R, Shouval D. Relapsing hepatitis A. Review of 14 cases and literature survey. *Medicine (Baltimore)* 71, 14–23 (1992).
6. Costafreda MI, Bosch A, Pinto RM. Development, evaluation, and standardization of a real-time TaqMan reverse transcription-PCR assay for quantification of hepatitis A virus in clinical and shellfish samples. *Appl. Environ. Microbiol.* 72, 3846–3855 (2006).
7. Bower WA, Nainan OV, Han X, Margolis HS *et al.* Duration of viremia in hepatitis A virus infection. *J. Infect. Dis.* 182, 12–17 (2000).
8. Akriviadis EA, Redeker AG. Fulminant hepatitis A in intravenous drug users with chronic liver disease. *Ann. Intern. Med.* 110, 838–839 (1989).
9. Germinario C, Luigi Lopalco P, Chicanna M, Da Villa G. From hepatitis B to hepatitis A and B prevention: the Puglia (Italy) experience. *Vaccine* 18(Suppl. 1), S83–S85 (2000).
10. Van Damme P, Van Herck K. Effect of hepatitis A vaccination programs. *JAMA* 294, 246–248 (2005).
11. Domínguez A, Oviedo M, Carmona G *et al.* Impact and effectiveness of a mass hepatitis A vaccination programme of preadolescents seven years after introduction. *Vaccine* 26, 1737–1741 (2008).
12. Pinto RM, Alegre D, Dominguez A *et al.* Hepatitis A virus in urban sewage from two Mediterranean countries. *Epidemiol. Infect.* 135, 270–273 (2007).

13. Cianciara J. Hepatitis A shifting epidemiology in Poland and eastern Europe. *Vaccine* 18(Suppl. 1), S68–S70 (2000).
14. Tallo T, Norder H, Tefanova V *et al.* Sequential changes in hepatitis A virus genotype distribution in Estonia during 1994 to 2001. *J. Med. Virol.* 70, 187–193 (2003).
15. Barzaga NG. Hepatitis A shifting epidemiology in south-east Asia and China. *Vaccine* 18(Suppl. 1), S61–S64 (2000).
16. Tanaka J. Hepatitis A shifting epidemiology in Latin America. *Vaccine* 18(Suppl. 1), S57–S60 (2000).
17. Lee D, Cho YA, Park Y *et al.* Hepatitis A in Korea: epidemiological shift and call for vaccine strategy. *Intervirology* 51, 70–74 (2008).
18. Van Herck K, Jacquet JM, Van Damme P. Antibody persistence and immune memory in healthy adults following vaccination with a two-dose inactivated hepatitis A vaccine: long-term follow-up at 15 years. *J. Med. Virol.* 83, 1885–1891 (2011).
19. Samandari T, Bell BP, Armstrong GL. Quantifying the impact of hepatitis A immunization in the United States, 1995–2001. *Vaccine* 22, 4342–4350 (2004).
20. Wasley A, Samandari T, Bell BP. Incidence of hepatitis A in the United States in the era of vaccination. *JAMA* 294, 194–201 (2005).
21. Dagan R, Leventhal A, Anis E, Slater P, Ashur Y, Shouval D. Incidence of hepatitis A in Israel following universal immunization of toddlers. *JAMA* 294, 202–210 (2005).
22. Vacchino MN. Incidence of hepatitis A in Argentina after vaccination. *J. Viral Hepat.* 15(Suppl. 2), 47–50 (2008).
23. Martin A, Lemon SM. Hepatitis A virus: from discovery to vaccines. *Hepatology* 43, S164–S172 (2006).
24. Yang Y, Liang Y, Qu L *et al.* Disruption of innate immunity due to mitochondrial targeting of a picornaviral protease precursor. *Proc. Natl Acad. Sci. USA* 104, 7253–7258 (2007).

▪ **A mechanism of inhibition of IFN synthesis.**

25. Asher LVS, Binn LN, Mensing TL, Marchwicki RH, Vassell RA, Young GD. Pathogenesis of hepatitis A in orally inoculated owl monkeys (*Aotus trivirgatus*). *J. Med. Virol.* 47, 260–268 (1995).
26. Schulman AN, Dienstag JL, Jackson DR *et al.* Hepatitis A Antigen particles in liver, bile, and stool of chimpanzees. *J. Infect. Dis.* 134, 80–84 (1976).
27. Ticehurst J, Rhodes LL Jr, Krawcznski K *et al.* Infection of owl monkeys (*Aotus trivirgatus*) and cynomolgus monkeys (*Macaca fascicularis*) with hepatitis E virus from Mexico. *J. Infect. Dis.* 165, 835–845 (1992).
28. Blank CA, Anderson DA, Beard M, Lemon SM. Infection of polarized cultures of human intestinal epithelial cells with hepatitis A virus: vectorial release of progeny virions through apical cellular membranes. *J. Virol.* 74, 6476–6484 (2000).
29. Ouzilou L, Caliot E, Pelletier I, Prévost MC, Pringault E, Colbère-Garapin F *et al.* Poliovirus transcytosis through M-like cells. *J. Gen.Virol.* 83, 2177–2182 (2002).
30. Dotzauer A, Brenner M, Gebhardt U, Vallbracht A. IgA-coated particles of hepatitis A virus are translocalized antivectorially from the apical to the basolateral site of polarized epithelial cells via the polymeric immunoglobulin receptor. *J. Gen. Virol.* 86, 2747–2751 (2005).
31. Dotzauer A, Gebhardt U, Bieback K *et al.* Hepatitis A virus-specific immunoglobulin A mediates infection of hepatocytes with hepatitis A virus via the asialoglycoprotein receptor. *J. Virol.* 74, 10950–10957 (2000).

▪ **A mechanism of entry of hepatitis A virus (HAV) into hepatocytes.**

32. Sanchez G, Aragones L, Costafreda MI, Ribes E, Bosch A, Pinto RM. Capsid region involved in hepatitis a virus binding to glycophorin A of the erythrocyte membrane. *J. Virol.* 78, 9807–9813 (2004).
33. Kaplan G, Totsuka A, Thompson P, Akatsuka T, Moritsugu Y, Feinstone SM. Identification of a surface glycoprotein on African green monkey kidney cells as a receptor for hepatitis A virus. *EMBO J.* 15, 4282–4296 (1996).
34. Tami C, Silberstein E, Manangeeswaran M *et al.* Immunoglobulin A (IgA) is a natural ligand of hepatitis A virus cellular receptor 1 (HAVCR1), and the association of IgA with HAVCR1 enhances virus-receptor interactions. *J. Virol.* 81, 3437–3446 (2007).

▪ **A mechanism of entry of HAV into cells.**

35. Racaniello VR. *Picornaviridae*, the viruses and their replication. In: *Fields Virology.* Knipe DM, Howley PM (Eds). Lippincott Williams & Wilkins, PA, USA, 795–838 (2007).
36. Bishop NE, Anderson DA. Uncoating kinetics of hepatitis A virus virions and provirions. *J. Virol.* 74, 3423–3426 (2000).
37. Lemon SM, Binn LN. Serum neutralizing antibody response to hepatitis A virus. *J. Infect. Dis.* 148, 1033–1039 (1983).
38. Hollinger FB, Emerson SU. Hepatitis A virus. In: *Fields' Virology.* Knipe DM, Howley PM (Eds). Lippincott Williams & Wilkins, PA, USA, 911–947 (2007).
39. Lanford RE, Feng Z, Chavez D *et al.* Acute hepatitis A virus infection is associated with a limited type I interferon response and persistence of intrahepatic viral RNA. *Proc. Natl Acad. Sci. USA* 108, 11223–11228 (2011).
40. Qu L, Feng Z, Yamane D *et al.* Disruption of TLR3 signaling due to cleavage of TRIF by the hepatitis A virus protease-polymerase processing intermediate, 3CD. *PLoS Pathog.* 7, E1002169 (2011).

▪ **A mechanism of inhibition of IFN synthesis.**

41. Brown EA, Zajac AJ, Lemon SM. *In vitro* characterization of an internal ribosomal entry site (IRES) present within the 5′ nontranslated region of hepatitis A virus RNA: comparison with the IRES of encephalomyocarditis virus. *J. Virol.* 68, 1066–1074 (1994).
42. Ehrenfeld E, Teterina NL. Initiation of translation of picornavirus RNAs: structure and function of the internal ribosome entry site. In: *Molecular Biology of Picornaviruses.* Semler BL, Wimmer E (Eds). ASM Press, Washington, DC, USA, 159–169 (2002).
43. Whetter LE, Day SP, Elroystein O, Brown EA, Lemon SM. Low efficiency of the 5′ nontranslated region of hepatitis A virus RNA in directing cap-independent translation in permissive monkey kidney cells. *J. Virol.* 68, 5253–5263 (1994).

▪▪ **Evidence of the low efficiency of the HAV internal ribosome entry site.**

44. Yang Y, Yi M, Evans DJ, Simmonds P, Lemon SM. Identification of a conserved RNA replication element (CRE) within the 3Dpol-coding sequence of hepatoviruses. *J. Virol.* 82, 10118–10128 (2008).
45. Leong LEC, Cornell CT, Semler BL. Processing determinants and functions of cleavage products of picornavirus. In: *Molecular Biology of Picornaviruses.* Semler BL, Wimmer E (Eds). ASM Press, Washington, DC, USA, 187–198 (2002).
46. Kuechler E, Seipelt J, Liebig H-D, Sommergruber W. Picornavirus proteinase-mediated shutoff of host cell translation: direct cleavage of a cellular initiation factor. In: *Molecular Biology of Picornaviruses.* Semler BL, Wimmer E (Eds). ASM Press, Washington, DC, USA, 301–312 (2002).
47. Borman AM, Kean KM. Intact eukaryotic initiation factor 4G is required for hepatitis A virus internal initiation of translation. *Virology* 237, 129–136 (1997).

▪▪ **Need for an intact eIF4G factor for HAV translation.**

48. Ali IK, McKendrick L, Morley SJ, Jackson RJ. Activity of the hepatitis A virus IRES requires association between the cap-binding translation initiation factor (eIF4E) and eIF4G. *J. Virol.* 75, 7854–7863 (2001).
49. Sanchez G, Bosch A, Pinto RM. Genome variability and capsid structural constraints of hepatitis A virus. *J. Virol.* 77, 452–459 (2003).
50. Chou T, Lakatos G. Clustered bottlenecks in mRNA translation and protein synthesis. *Phys. Rev. Lett.* 93(19), 198101 (2004).
51. Robinson M, Lilley R, Little S *et al.* Codon usage can affect efficiency of translation of genes in *Escherichia coli*. *Nucleic Acids Res.* 12, 6663–6671 (1984).
52. Sorensen MA, Kurland CG, Pedersen S. Codon usage determines translation rate in *Escherichia coli*. *J. Mol. Biol.* 207, 365–377 (1989).
53. Adzhubei AA, Adzhubei IA, Krasheninnikov IA, Neidle S. Non-random usage of 'degenerate' codons is related to protein three-dimensional structure. *FEBS Lett.* 399, 78–82 (1996).
54. Gavrilin GV, Cherkasova EA, Lipskaya GY, Kew OM, Agol VI. Evolution of circulating wild poliovirus and vaccine-derived poliovirus in an immunodeficient patient: a unifying model. *J. Virol.* 74, 7381–7390 (2000).
55. Evans MS, Clarke TF, Clark PL. Conformations of co-translational folding intermediates. *Protein Pept. Lett.* 12, 189–195 (2005).
56. Aragonès L, Guix S, Ribes E, Bosch A, Pintó RM. Fine-tuning translation kinetics selection as the driving force of codon usage bias in the hepatitis A virus capsid. *PLoS Pathog.* 6, E1000797 (2010).

▪▪ **Role of capsid rare codons in HAV fitness.**

57. Abad FX, Pinto RM, Bosch A. Survival of enteric viruses on environmental fomites. *Appl. Environ. Microbiol.* 60, 3704–3710 (1994).
58. Bosch A. Human enteric viruses in the water environment: a minireview. *Int. Microbiol.* 1, 191–196 (1998).
59. Reid TMS, Robinson HG. Frozen raspberries and hepatitis A. *Epidemiol. Infect.* 98, 109–112 (1987).
60. Rosemblum LS, Mirkin IR, Allen DT, Safford S, Hadler SC. A multistate outbreak of hepatitis A traced to commercially distributed lettuce. *Am. J. Public Health* 80, 1075–1080 (1990).
61. Bosch A, Lucena F, Diez JM, Gajardo R, Blasi M, Jofre J. Waterborne viruses associated with hepatitis outbreak. *J. Am. Water Works Assoc.* 83, 80–83 (1991).
62. Dentinger CM, Bower WA, Nainan OV *et al.* An outbreak of hepatitis A associated with green onions. *J. Infect. Dis.* 183, 1273–1276 (2001).
63. Sanchez G, Pinto RM, Vanaclocha H, Bosch A. Molecular characterization of hepatitis a virus isolates from a transcontinental shellfish-borne outbreak. *J. Clin. Microbiol.* 40, 4148–4155 (2002).
64. Pinto RM, Costafreda MI, Bosch A. Risk assessment in shellfish-borne outbreaks of hepatitis A. *Appl. Environ. Microbiol.* 75, 7350–7355 (2009).
65. Liang Y, Shilagard T, Xiao S *et al.* Visualizing hepatitis C virus infections in human liver by two-photon microscopy. *Gastroenterology* 137, 1448–1458 (2009).
66. Brack K, Berk I, Magulski T, Lederer J, Dotzauer A, Vallbracht A. Hepatitis A virus inhibits cellular antiviral defense mechanisms induced by double-stranded RNA. *J. Virol.* 76, 11920–11930 (2002).
67. Fensterl V, Grotheer D, Berk I, Schlemminger S, Vallbracht A, Dotzauer A. Hepatitis A virus suppresses RIG-I-mediated IRF-3 activation to block induction of beta interferon. *J. Virol.* 79, 10968–10977 (2005).
68. Probst C, Jecht M, Gauss-Muller V. Intrinsic signals for the assembly of hepatitis A virus particles. Role of structural proteins VP4 and 2A. *J. Biol. Chem.* 274, 4527–4531 (1999).
69. Graff J, Richards OC, Swiderek KM *et al.* Hepatitis A virus capsid protein VP1 has a heterogeneous C terminus. *J. Virol.* 73, 6015–6023 (1999).
70. Martin A, Bénichou D, Chao SF, Cohen L, Lemon SM. Maturation of the hepatitis A virus capsid protein VP1 is not dependent on processing by the $3C^{pro}$ proteinase. *J. Virol.* 73, 6220–6227 (1999).
71. Fry EE, Stuart DA. Virion structure. In: *The Picornaviruses.* Ehrenfeld E, Domingo E, Roos R (Eds). ASM Press, Washington, DC, USA, 59–71 (2010).
72. Domingo E, Martin V, Perales C, Grande-Perez A, Garcia-Arriaza J, Arias A. Viruses as quasispecies: biological implications. *Curr. Top. Microbiol. Immunol.* 299, 51–82 (2006).
73. Overbaugh J, Bangham CRM. Selection forces and constraints on retroviral sequence variation. *Science* 292, 1106–1109 (2001).
74. Sanchez G, Bosch A, Gomez-Mariano G, Domingo E, Pinto RM. Evidence for quasispecies distributions in the human hepatitis A virus genome. *Virology* 315, 34–42 (2003).
75. Robertson BH, Jansen RW, Khanna B *et al.* Genetic relatedness of hepatitis A virus strains recovered from different geographical regions. *J Gen. Virol.* 73, 1365–1377 (1992).
76. Costa-Mattioli M, Cristina J, Romero H *et al.* Molecular evolution of hepatitis a virus: a new classification based on the complete VP1 protein. *J. Virol.* 76, 9516–9525 (2002).
77. Perez-Sautu U, Costafreda MI, Lite J *et al.* Molecular epidemiology of hepatitis A virus infections in Catalonia, Spain, 2005–2009: circulation of newly emerging strains. *J. Clin. Virol.* 52, 98–102 (2011).
78. Perez-Sautu U, Costafreda MI, Cayla J *et al.* Hepatitis A virus vaccine escape variants and potential new serotype emergence. *Emerg. Infect. Dis.* 17, 734–738 (2011).

▪▪ **Emergence of HAV antigenic variants in the HIV-positive MSM risk group.**

79. Ping LH, Lemon SM. Antigenic structure of human hepatitis A virus defined by analysis of escape mutants selected against murine monoclonal antibodies. *J. Virol.* 66, 2208–2216 (1992).
80. Nainan OV, Brinton MA, Margolis HS. Identification of amino-acids located in the antibody-binding sites of human hepatitis-A virus. *Virology* 191, 984–987 (1992).
81. Aragonès L, Bosch A, Pintó RM. Hepatitis A virus mutant spectra under the selective pressure of monoclonal antibodies: codon usage constraints limit capsid variability. *J. Virol.* 82, 1688–1700 (2008).

▪ **Role of rare capsid codons in the low antigenic variability of HAV.**

82. Pinto RM, Aragones L, Costafreda MI, Ribes E, Bosch A. Codon usage and replicative strategies of hepatitis A virus. *Virus Res.* 127, 158–163 (2007).
83. Gabrieli R, Sanchez G, Macaluso A *et al.* Hepatitis in Albanian children: molecular analysis of hepatitis A virus isolates. *J. Med. Virol.* 72, 533–537 (2004).
84. Neilsen GA, Bodsworth NJ, Watts N. Response to hepatitis A vaccination in human immunodeficiency virus-infected and -uninfected homosexual men. *J. Infect. Dis.* 176, 1064–1067 (1997).
85. Weissman S, Feucht C, Moore BA. Response to hepatitis A vaccine in HIV-positive patients. *J. Viral Hepat.* 13, 81–86 (2006).
86. Borrego B, Novella IS, Giralt E, Andreu D, Domingo E. Distinct repertoire of antigenic variants of foot-and-mouth disease virus in the presence or absence of immune selection. *J. Virol.* 67, 6071–6079 (1993).
87. Bosch A, Mueller S, Pintó RM. Codon biases and viral fitness. In: *The Picornaviruses.* Ehrenfeld E, Domingo E, Roos R (Eds). ASM Press, Washington, DC, USA, 271–283 (2010).
88. Hershberg R, Petrov DA. Selection on codon bias. *Annu. Rev. Genet.* 42, 287–299 (2008).

89. Komar AA. A pause for thought along the co-translational folding pathway. *Trends Biochem. Sci.* 34, 16–24 (2009).

90. D'Andrea L, Pintó RM, Bosch A, Musto H, Cristina J. A detailed comparative analysis on the overall codon usage patterns in hepatitis A virus. *Virus Res.* 157, 19–24 (2011).

91. Zhang Y, Liu Y, Liu W *et al.* Analysis of synonymous codon usage in hepatitis A virus. *Virol. J.* 8, 174 (2011).

92. van Weringh A, Ragonnet-Cronin M, Pranckeviciene E, Pavon-Eternod M, Kleiman L, Xia X. HIV-1 modulates the tRNA pool to improve translation efficiency. *Mol. Biol. Evol.* 28, 1827–1834 (2011).

93. Tucker KL. Methylated cytosine and the brain: a new base for neuroscience. *Neuron* 30, 649–652 (2001).

94. Takeshita F, Gursel I, Ishii KJ, Suzuki K, Gursel M, Klinman DM. Signal transduction pathways mediated by the interaction of CpG DNA with Toll-like receptor 9. *Semin. Immunol.* 16, 17–22 (2004).

95. Sugiyama T, Gursel M, Takeshita F *et al.* CpG RNA. Identification of novel single-stranded RNA that stimulates human $CD14^+CD11c^+$ monocytes. *J. Immunol.* 174, 2273–2279 (2005).