Chapter 2
Etiology

2.1 Structure of the Virus and Genome Organization

In1983, the hepatitis E virus was first identified by immune electron microscopy in the feces of patients with enterically transmitted non-A-C hepatitis. Subsequent establishment of the disease in cynomolgus macaques led to cloning and sequencing of HEV in 1990 (Balayan et al. 1983; Kane et al. 1984).

The hepatitis E virus (HEV) is a small (27–34 nm), icosahedral, non-enveloped single-stranded positive-sense RNA virus.

The HEV genome is approximately 7.2 kb in length, and presents a 7-methylguanosine cap followed by three overlapping open reading frames (ORFs) and a second non-coding region of about 65–74 nucleotides with a 3′ poly A tail. The genome length slightly varies between animal strains; shorter genomes have been detected in rat HEV in Vietnam (6927 nt) (Li et al. 2013), in avian strains (6654 nt) (Huang et al. 2004) and in recently detected bat viruses (6767 nt) (Drexler et al. 2012), although the genome organization seemed to be conserved in all these cases. ORF1 (5073–5124 nt) codes for a non-structural poly-protein of about 1,690 amino acids, which is involved in viral genome replication and viral protein processing. The poly-protein functional domains include a methyltransferase (MeT), flanked by the Y domain, a papain-like cysteine protease (PCP), flanked by the macro domain (X domain), an RNA helicase, and an RNA-dependent RNA polymerase. The N-terminal portion of ORF1 serves as a viral methyltransferase (MeT) that catalyzes the capping of both genomic and subgenomic viral RNAs (Rozanov et al. 1992). Capping of the viral RNA determines its translation and, as recently shown, it is decisive for virus defense from the innate host response, inhibiting interferon cascade activation (Pichlmair et al. 2006). The MeT is followed by the so-called Y domain, which shares significant homology with non-structural proteins of other positive-stranded RNA viruses. This is followed by the papain-like cysteine proteases (PCP) and by the formerly designated X domain, more recently renamed as the "macro domain", which may be involved in the binding of ADP-ribose and its polymeric form (Neuvonen and Ahola 2009). As for other described PCPs, also in

F. M. Ruggeri et al., *Hepatitis E Virus*, SpringerBriefs in Food, Health, and Nutrition,
DOI: 10.1007/978-1-4614-7522-4_2, © Franco Maria Ruggeri, Ilaria Di Bartolo,
Marcello Trevisani, Fabio Ostanello 2013

the HEV genome, the presence of a macro domain following the PCP sequence strengthens its functional homology with proteases of diverse origin, although the specific role of HEV PCP in polyprotein (pORF1) processing still remains undefined (Ahmad et al. 2011). pORF1 is characterized by the presence of a region rich in proline residues and without a predicted secondary structure, which might act as a flexible hinge within the protein. The predicted helicase domain of HEV contains a full complement of conserved helicase motifs (Karpe and Lole 2010), including the seven conserved motifs proposed to contain both the NTPase activity and an RNA binding domain. The HEV helicase possesses an RNA 5′-triphospatase activity involved in the first step of RNA capping (Karpe and Lole 2010). The C-terminal domain of pORF1 has RNA-dependent RNA polymerase (RdRp) activity (Agrawal et al. 2001); this is an essential enzyme for RNA virus replication through the synthesis of an anti-genomic RNA intermediate. The endoplasmic reticulum was identified as the site of replicase localization, and the intracellular membranes are the possible sites where RNA replication occurs (Rehman et al. 2008). ORF3 (366–369 nt) follows ORF1 and overlaps the N-terminal portion of ORF2, in a different reading frame, and encodes for a small phosphoprotein (pORF3), which is expressed at the intracellular level. The protein contains two hydrophobic and two proline rich domains; these regions contain amino acid motifs involved in signal transduction (Korkaya et al. 2001), and a PSAP motif is present and conserved in all HEV isolates, including avian HEV. pORF3 does not show homology with any other known protein; its role still remains unclear. Recent studies of the biology of HEV replication have shown that pORF3 may be involved in virus release from infected cells (Okamoto 2011), since it is associated with the cytoskeleton and is present on the virion surface (Yamada et al. 2009) (Fig. 2.1). Moreover, pORF3 downregulates innate host responses through the reduction of the expression of acute phase proteins and promotion of the secretion of α1-microglobulins (Panda et al. 2007).

An additional three ORFs have been described in rat and bat HEV genomes, but their function remains yet unknown. No suggestive similarity of the putative gene products of the internal reading frame to any described protein domain could be detected by BLAST comparison (Johne et al. 2010a; Drexler et al. 2012).

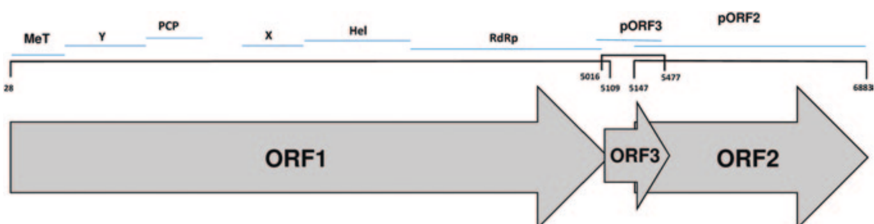

Fig. 2.1 Genomic organization of HEV, including the three ORFs. Nucleotide positions are referred to a prototype strain (Acc. No. M73218). On the *top*, functional domains are indicated: *MeT* methyltransferase; *Y* domain; *PCP* protease; *X* domain; *Hel* helicase; *RdRp* RNA dependent RNA polymerase; *pORF2* capsid protein; pORF3

The viral capsid protein encoded by ORF2 works for particle assembly, binding to host cells, and elicitation of neutralizing antibodies. pORF2 is glycosylated, and three asparagine (Asn) residues for N-linked glycosylation sites have been identified. The virus capsid is made up of subunits containing 30 homodimers of pORF2 (Yamada et al. 2009). The crystal structure of a truncated recombinant pORF2 protein has been obtained, but the real size of the protein in mature virions remains unknown (Yamashita et al. 2009). Among four major mammalian HEV genotypes, sequence identity among the amino acid residues of the capsid protein was over 85 %, and many amino acid divergences were found in the N-terminal 111 residues. The N-terminal region of the HEV capsid protein is most likely to represent the shell domain, whereas the C-terminal region of pORF2 is more variable and is considered to be the protruding domain of the HEV capsid protein (Li et al. 2009). The initial contact with host cells in order to initiate viral infection is believed to occur through these protrusions (Pichlmair et al. 2006). Expression of a truncated capsid protein lacking the first 111 amino acids and/or the C-terminal 59 amino acids in insect cells by the baculovirus expression system resulted in self-assembly of the capsid protein and in the production of two types of HEV-like particle (HEV-VLP) with different diameters (Li et al. 1997; Caprioli et al. 2005; Xing et al. 2011), corresponding to different proteolytic cleavages. As demonstrated by protein expression in the baculovirus insect cell expression system, the minimum requirement for assembly was inclusion of amino acid residues 126–601 (Li et al. 2005a). The N-terminal domain followed by the signal sequence (residues 28–101) is an arginine-rich domain resembling the RNA-binding domain of the coat proteins of tombusviruses. The capsid protein binding to Huh-7 liver cells has been studied, and it appeared to be mediated by heparin sulfate proteoglycans (HSPGs), specifically syndecans, as demonstrated using the baculovirus expressed pORF2, assembled in VLPs (Kalia et al. 2009).

Recent studies on HEV particles have provided useful information about the HEV life cycle as well as for the possible development of monovalent or polyvalent vaccines. Indeed, the recombinant HEV capsid protein is currently undergoing clinical trials as a vaccine candidate (Zhu et al. 2010; Zhao et al. 2012), and genotype 2 HEV VLPs have been proposed as a useful carrier for foreign DNA (Takamura et al. 2004) or epitopes into mucosal epithelial cells (Niikura et al. 2002). However, several knowledge gaps still remain concerning the structure of HEV capsid, such as what role pORF3 may possibly have in the virion architecture and function. Further studies will be needed to answer these questions.

2.2 Taxonomy and Nomenclature

Because of limitations in allowing it to grow reproducibly and efficiently in vitro, HEV classification has been mainly based on the analysis of the viral RNA by sequencing and phylogenetic techniques (Korkaya et al. 2001).

Table 2.1 Genotypes and host range of the hepatitis E viruses. Adapted from Meng et al. (2011)

HEV strains	Natural host
Mammalian HEV	
Genotype 1	Humans
Genotype 2	Humans
Genotype 3	Humans, domestic pigs, wild boar, deer, mongooses, rabbits
Genotype 4	Humans, domestic pigs, wild boar
Novel unclassified genotype, Rat HEV	Rats
Novel unclassified genotype, Boar HEV	Wild boar in Japan
Novel unclassified genotype, Bat HEV	Bats
Novel unclassified genotype, Ferret HEV	Ferrets
Avian HEV	
Genotype 1	Chickens
Genotype 2	Chickens
Genotype 3	Chickens
Genotype ?	Chickens (Hungary)
Trout HEV	
Genotype ?	Cutthroat trout (USA)

HEV was initially classified within the *Caliciviridae* family, but the increasing numbers of sequences collected afterwards have clearly unmasked significant differences with other caliciviruses, and since 2004 HEV has been classified as a new genus called Hepevirus in the family of *Hepeviridae* (Emerson and Purcell 2003). HEV strains detected in humans and other mammalian species represent the major genus of the *Hepeviridae* (Table 2.1). Although avian HEV strains share only 50–60 % nucleotide identity with mammalian HEV strains (Meng 2010a), specific antibodies are able to cross-react with the capsid protein of both groups of viruses, demonstrating the presence of common epitopes (Haqshenas et al. 2001). Nonetheless, avian HEV strains have never been associated with cases of infection in human beings (Kamar et al. 2012), causing hepatitis-splenomegaly syndrome (HS) only in chickens (Haqshenas et al. 2001). Consequently, it has been proposed to assign avian HEV to a separate genus, consisting of at least three different genotypes (Bilic et al. 2009; Marek et al. 2010).

Hepeviridae includes four genotypes of mammalian HEV, which primarily infect humans, domestic pigs, wild boar, deer, and rabbits (Meng et al. 2012). However, genetically distant HEV strains have more recently been identified in the rat (Johne et al. 2010b), ferrets (Raj et al. 2012), wild boar (Takahashi et al. 2011), bats (Drexler et al. 2012), and cutthroat trout (*Oncorhynchus clarkii*) (Batts et al. 2011), suggesting that the *Hepeviridae* family classification should be reviewed. A proposed revision includes introduction of separate clades: one genus would comprise human HEV genotypes and closely related animal viruses, while the others would include viruses from chiropteran (bat), rodent (rat), and avian (chicken) hosts (Drexler et al. 2012). The “cutthroat” hepevirus is genetically the most

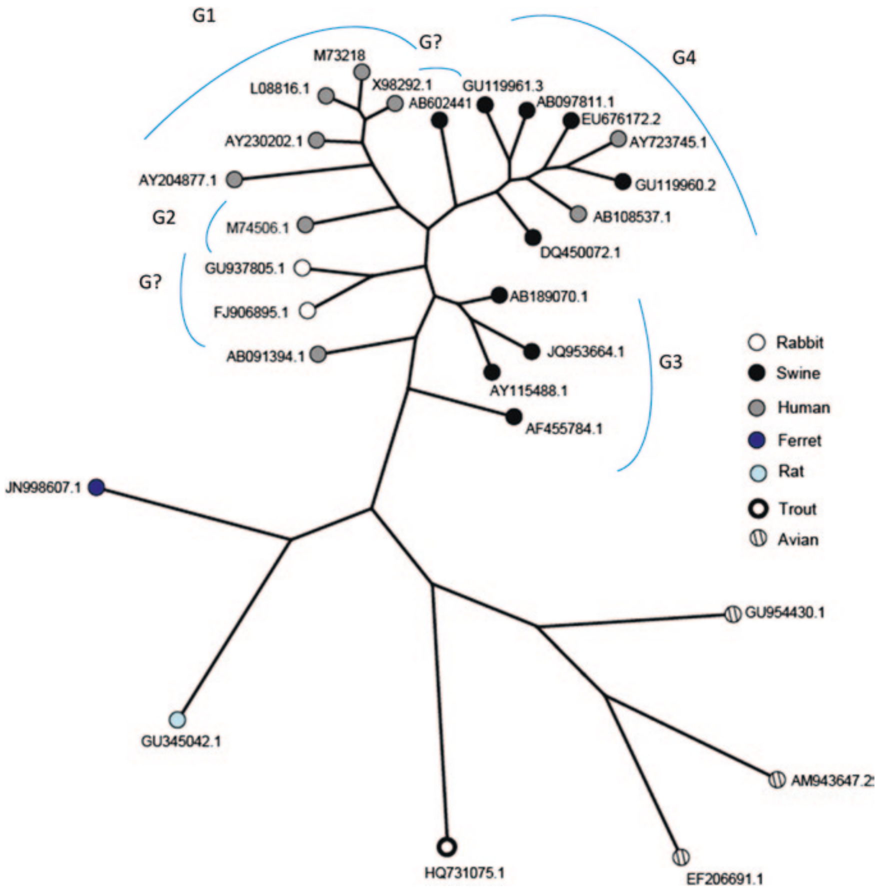

Fig. 2.2 Phylogenetic tree illustrating different genotypes of hepatitis E. The tree is based on full-length sequences of HEV strains of either human or animal origin. The GenBank accession number of each strain used in the tree is indicated

distant animal strain, and it might correspond to a separate taxonomic unit of a higher rank, e.g., a subfamily.

Based on the sequence comparisons of the HEV genome currently available, strains are classified in genotypes and sub-types. The four major genotypes identified in mammalians to date include (Fig. 2.2 and Table 2.1): genotype 1 (Burmese-like Asian strains), genotype 2 (a Mexican strain and some African strains), genotype 3 (strains from humans and animals, worldwide), and genotype 4 (strains from human cases and animal strains, in Asia and Europe) (Meng 2011). Avian HEV, genetically distant from mammalian strains, has been classified separately and comprises at least three different genotypes (Marek et al. 2010), since a putative novel avian HEV genotype has been identified but not yet classified

(Banyai et al. 2012). However, with the recent identification of genetically distinct HEV strains from rats (Johne et al. 2010a), ferrets (Raj et al. 2012), and wild boar (Takahashi et al. 2011), new previously unrecognized genotypes have been proposed.

Despite the knowledge that different HEV genotypes occur, the virus seems otherwise to exist as a single serotype (Aggarwal and Naik 2009).

Most human infections that occur in Asia and Africa are caused by genotype 1, whereas genotype 2 is commonly found in Mexico and West Africa (Nigeria and Chad). In industrialized countries, where until a few years ago hepatitis E was considered non-endemic, autochthonous cases appear to be related to HEV strains belonging to genotypes 3 and 4 (Emerson and Purcell 2003; Okamoto 2007; Kamar et al. 2012; Scobie and Dalton 2013), that are still considered the only zoonotic genotypes. Genotype 3 HEV was first identified in human cases locally acquired in the USA, when the two human strains, called US-1 and US-2, showed only 74–75 % nucleotide identity with genotypes 1 and 2, being classified separately (Meng et al. 1997). Since then, genotype 3 has been detected throughout the entire world, associated with sporadic cases and small outbreaks in North America, Europe, Japan and New Zealand (Tsang et al. 2000; Mansuy et al. 2004; Dalton et al. 2007a, b). This genotype is also commonly detected in animals, and a strict genetic correlation has been observed between human and animal strains circulating in the same geographical area. The first animal strain of HEV was identified in swine in the USA in 1997. The virus was shown to belong to genotype 3, and presented a high identity with some human strains (Meng et al. 1997). In particular, the virus shared a 92 % nucleotide identity in ORF2 with the two HEV genotype 3 autochthonous human strains (US-1 and US-2) detected in the same period in the USA. Given the strict genetic correlation, the two viruses were classified in the same genotype 3, and since then pigs have been considered a reservoir of HEV (Meng et al. 1997).

Recently, a study conducted in Japan identified a new potential reservoir of genotype 3 HEV virus in the mongoose. As revealed by phylogenetic analysis, the HEV RNAs detected belonged to genotype 3 and were classified into two groups, one of which contains sequences very similar to mongoose HEV previously detected in the same area as well as to HEV identified in a pig (Nidaira et al. 2012).

Genotype 4, the other genotype transmitted by the zoonotic route, is indigenous to Asia, where it has been recovered from both pigs and humans (Wang et al. 2012). Since its first report in China in 1999, genotype 4 has been increasingly described as being endemic in pigs and as the cause of sporadic cases of hepatitis E in humans and infection in the swine in China and Japan, and more recently in Europe (Hakze-van der Honing et al. 2011; Colson et al. 2012; Garbuglia et al. 2013). The question for the recent and increasingly frequent detection of genotype 4 in Europe is of concern for public health, and raises the question whether genotype 4 was somehow introduced into domestic pigs and may be expected to spread further on farms or whether the problem is to be correlated to importation of pig meat of Asian origin into Europe (Colson et al. 2012).

The recent availability of increasing numbers of HEV sequences emphasizes the genome diversity among HEV isolates. According to the nucleotide identity, the four genotypes were further subdivided into subtypes. HEV strains belonging to genotype 1 are more conserved and can be further classified into five subtypes (1a–1e). Genotype 2 sequences are classified into two subtypes (2a and 2b). Genotypes 3 and 4 are extremely diverse and are divided into ten (3a–3j) and seven subtypes (4a–4g), respectively. Thus, in total, at least 24 subtypes of HEV exist in nature (Lu et al. 2006). The diversity of genotypes 3 and 4 appears related to their zoonotic origin from a variety of animals in different parts of the world, whereas the relative conservation of genotypes 1 and 2 is consistent with their primary circulation in humans and a less frequent isolation from animals (Pavio et al. 2010).

Nevertheless, several studies have supported the existence of intra-host quasispecies in both humans (Grandadam et al. 2004) and swine (Bouquet et al. 2012b). A recent study conducted using next generation sequencing (NGS) confirmed the existence of intra-host quasispecies in experimentally infected swine, underlining that the range of quasispecies diversity is lower than for other human viruses, but is in line with other zoonotic viruses (Bouquet et al. 2012b). In human patients undergoing solid-organ transplantation with chronic HEV infection, the quasispecies diversification seems to be related to rapid development and progression of liver fibrosis, since patients had lower quasispecies diversification during the first year than had patients without liver fibrosis progression (Lhomme et al. 2012).

2.3 Viral Replication Cycle

Because cell culture and small animal models to investigate HEV infection was developed only recently, the viral replication and regulation processes involved are poorly understood, being mainly based on HEV genome analysis and homology with other positive-stranded RNA viruses. Nevertheless, the recent identification of various cell lines permissive to HEV replication (hepatic cell lines HuH7, PCL/PRF/5, HepG2; lung carcinoma cell lines A549; human hepatoma-derived cell line, HepaRG; porcine embryonic stem cell-derived cell line, PICM-19) (Okamoto 2011; Rogee et al. 2012) has significantly contributed to clarify some essential steps of the replicative cycle. Although no cellular receptors have been identified up to now, a role of heparin sulfate proteoglycans (HSPGs) in mediating virus attachment has been demonstrated (Kalia et al. 2009). The mechanisms by which the virus is released into the cytoplasm are unknown. Genomic RNA appears to be translated directly into the ORF1 polyprotein, even though it is unclear whether the polyprotein is or is not processed into individual functional units. The RdRp mediates the replication of the positive-sense genomic RNA into negative-sense RNA transcripts, that serve as a template for the synthesis of the full genome, and into the single 2.2 kb subgenomic RNA including the overlapping ORF2 and ORF3 that are translated in the capsid protein pORF2 and in pORF3. pORF2 is assembled in a capsid particle where the RNA genome

is packaged to construct the newly formed virion. Although the role of pORF3 remains largely unclear, this protein mediates virus budding likely based on the interaction of lipid-associated virions with plasma membranes or endomembranes (Ahmad et al. 2011). Binding of the cellular TSG101 (tumor susceptibility gene 101) to pORF3 through amino acid PSAP motif (i.e., amino acids proline, serine, alanine, and proline) has been demonstrated (Surjit et al. 2006). TSG101 has been identified as the critical protein for budding of enveloped viruses, such as the human immunodeficiency virus type-1 (HIV) from the plasma membrane (Martin-Serrano et al. 2001). It is likely that pORF3 mediates virus budding by recruiting the TSG101 (Okamoto 2011).

http://www.springer.com/978-1-4614-7521-7

Hepatitis E Virus
An Emerging Zoonotic and Foodborne Pathogen
Ruggeri, F.M.; Di Bartolo, I.; Ostanello, F.; Trevisani, M.
2013, VI, 88 p. 8 illus. in color., Softcover
ISBN: 978-1-4614-7521-7