The aetiology of hepatitis E virus infection (HEV) is complex. Many individuals exposed to HEV infection remain asymptomatic while others go on to develop an acute hepatitis of varying severity that usually resolves within 2–3 months [reviewed in (1)]. Chronic infection lasting a year or more has been observed rarely, and then only in immunocompromised individuals. Another potential outcome of HEV infection is fulminant hepatitis (FH), an uncommon rapidly deteriorating state with a poor prognosis involving pathologies such as hepatic encephalopathy, necrosis of hepatic parenchyma, coagulopathy, renal failure or coma.

Several previous studies have suggested that the pathogenicity of HEV infection might be genotype or strain dependent. For example, a bias towards genotype 4 has been noted in Japanese patients with FH or severe disease (2–4), while a recent study reports an association between infection with genotype 4 virus in France and higher levels of ALT and the presence of jaundice (5). There are also reports that disease severity (including FH) is associated with a particular strain of HEV genotype 3 (6), and that FH is associated with particular strains of genotype 1 (7) or genotype 4 virus (8). Further, several recent publications have suggested that there may be a link between particular substitutions in the HEV genome and the development of FH. For example, progression to FH has been associated with 142 synonymous and 8 nonsynonymous substitutions of genotype 1 virus (9), with the presence of two or three synonymous substitutions of genotype 4 virus (10, 11), or with 12 unique amino acid substitutions in a genotype 4 virus from a FH patient (12).

However, an unacknowledged problem with some of these studies is the common geographical origin of the variants studied so the reported associations between particular virus substitutions and FH might simply occur because of epidemiological relationships amongst the viruses rather than because such substitutions were involved in the development of FH. This article re-examines published evidence for an association between FH and particular HEV genotypes, lineages or particular nucleotide substitutions.
Materials and methods

Complete HEV genome sequences downloaded from Genbank were as follows: Genotype 1: JF443726, JF443725, JF443724, JF443723, JF443722, FJ457024, X98292, M73218, JF443721, JF443720, JF443719, JF443718, JF443717, AF459438, AF076239, X99441, AF051830, DQ459342, M80581, D10330, AF185822, L08816, D11092, D11093, L25595, AY204877, AY230202, AB270034, AB270035, JQ655734, AF444002, AF444003, L25547, M94177, Genotype 3: AB291955, Genotype 4 human isolates: AB291967, AB291959, AB193176, AB220971, AB220972, AB220973, AB091935, JQ740781, AB291966, AB291965, AB291968, AB220974, AB291964, AB220975, AB220976, AB220977, AB220978, AB220979, AJ272108, AB08537, AB097812, FJ763142, KC492825, AB698654, JQ655735, JQ655733, HQ634346, HM439284, AB369690, AB369688, AB197674, AB197673, AB193178, AB099347, AB074917, AB074915, AB080575. Sequences were aligned and annotated using SSE v1.1 (13) and phylogenetic analysis was performed using Mega 6 (14). Nucleotide positions were numbered relative to AB220978. The significance of associations between substitutions at each genome position and FH status was measured using Fisher’s exact test using a significance level of $P < 0.01$ in a two-tailed test as implemented in R script available upon request from the authors. The same data sets were analysed using metaCATS (www.viprbrc.org) that uses both the chi-squared test of independence and Pearson’s chi-squared test (15). The genotype 1 data set differed from that of (9) in that three identical or near-identical sequences (L25547, M94177 and AF444002) were removed, and three recently reported genotype 1 non-FH sequences (AB720034, AB720035 and JQ655734) were added.

Results

Is FH genotype-specific?

There are reports of patients developing FH after infection with genotype 1 (7, 9, 16, 17), genotype 3 (6) and genotype 4 (8, 10, 12, 18–22). A large outbreak of HEV genotype 2 in Namibia was associated with FH in 3/600 (0.5%) of individuals (23); although nucleotide sequences were not reported from the fulminant cases it seems likely that genotype 2 virus was involved. Hence, all four of the currently identified HEV genotypes known to infect humans can result in FH.

Is FH strain-specific?

We next investigated the possibility that FH results from the infection with particular strains of HEV. Phylogenetic analysis of the complete genome sequences of HEV derived from FH patients (Fig. 1) reveals the presence of three lineages (groups of sequences supported by >70% of bootstrap replication) for both genotypes 1 and 4. Analysis of partial genome sequences of HEV from an additional 10 genotype 1 FH patients (EF05410, EF175962-4, EF206258&8 (7) and FJ230847–50), identified another genotype 1 lineage (Figure S1A, B), while analysis including the genotype 4 FH sequences (AB108659&60 (20), AB305793 (8), AB079762 (24) and AB114178) revealed a fourth genotype 4 lineage (Figure S1C–E). The ORF2 sequences of HEV genotype 3 FH isolates AB079763 (11), EF601404 and AB291955 (6) fell into three lineages (Figure S1F). Hence, these 31 FH-derived HEV variants comprise at least 11 lineages within genotypes 1, 3 and 4.

Is FH transmissible?

Three published studies describe exposure from a common source (all HEV genotype 4) following which at least one individual developed FH. Of two men who consumed uncooked boar liver, one developed FH while the other individual had acute hepatitis (19). Similarly, of 13 individuals who had eaten grilled pig liver and intestines together, one developed FH and five seroconverted but were asymptomatic (21). One of these individuals transmitted HEV by blood transfusion to an individual who cleared the infection following interferon treatment. Finally, of 40 individuals who had eaten barbecued pig meat and entrails together, one developed FH, one had acute severe hepatitis, one had self-limited hepatitis, one seroconverted but had a subclinical infection, 10 were seronegative and the remainder were unavailable for follow up (22). These observations imply either that exposure to a particular strain of HEV is not sufficient to produce FH or that the penetrance of the trait is low. From the cases described above the penetrance would be 50%, 17% and 25%.

Is FH associated with specific virus substitutions?

We tested possibility that particular substitutions are responsible for the development of FH by tabulating sequence differences in two of the common source outbreaks for which complete genome sequences were available (Table 1). None of the 17 sites at which viruses derived from the same source differed, were also variable in the other transmission set, and none were overrepresented amongst other FH isolates. Similarly, although a genotype 4 virus isolated from a FH patient had 12 unique amino acid substitutions (12), none of these substitutions were present in other genotype 4 FH sequences.

Previous studies have used statistical methods to identify substitutions that are associated with isolates derived from FH patients infected with HEV genotype 4 (10, 11) or genotype 1 (9). However, interpretation of these results is complicated by the restricted geographical and temporal origin of the FH isolates; five of the seven genotype 1 FH isolates were from the single
Indian city of Pune and were sampled in consecutive years (9) while five of eight FH-derived genotype 4 complete genome sequences were obtained on the Japanese island of Hokkaido from 2002 to 2006 (10, 11). Our reanalysis of the genotype 1 data set using Fisher’s exact test identified 110 sites significantly ($P < 0.01$) associated with FH, but 136 sites if only the five FH isolates from Pune or Hokkaido, and four sites if only one of the Pune FH isolates was used. These four sites (positions 248, 946, 967 and 6312) were all synonymous, although position 6312 lies within a proposed secondary structure that is essential for virus replication (25). When we randomly assigned FH status to five of six non-FH isolates D11093, D11092, L08816, M80581, JQ655734 and L25595 that formed a well-defined lineage, >230 sites were detected. A three nucleotide deletion in ORF2/3 at position 5344 found in three FH isolates (9) was also present in two non-FH isolates, giving a non-significant association using Fisher’s exact test (3/7 compared to 2/23, $P = 0.06$). Together, these observations suggest that the previous reports of sites significantly associated with FH amongst genotype 1 viruses may be actually reflect the common geographical and temporal origin of 5/7 FH isolates.

Fig. 1. Phylogenetic analysis of complete HEV sequences from FH and non-FH patients. Distances between HEV genotype 1 (A) and genotype 4 (B) complete genome sequences are presented on a neighbor-joining tree (FH patients are indicated by symbols, ■ for FH isolates from Pune or Hokkaido, and ● for other FH isolates, □ for the genotype 3 outgroup). Branches supported by >70% of bootstrap replications (n = 500) are indicated.
Table 1. FH Associations in common-source genotype 4 infections

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Nucleotide position</th>
<th>Syn/Nsyn</th>
<th>100</th>
<th>307</th>
<th>1963</th>
<th>2376</th>
<th>2968</th>
<th>3718</th>
<th>3921</th>
<th>4072</th>
<th>5203</th>
<th>5377</th>
<th>5541</th>
<th>5637</th>
<th>6053</th>
<th>6412</th>
<th>6648</th>
<th>7006</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Syn</td>
<td>Syn</td>
<td>Syn</td>
<td>NS</td>
<td>Syn</td>
<td>NS</td>
<td>Syn</td>
<td>NS</td>
<td>2/3</td>
<td>2/3</td>
<td>Syn</td>
<td>Syn</td>
<td>NS</td>
<td>NS</td>
<td>NS</td>
<td>Syn</td>
</tr>
<tr>
<td>AB291959</td>
<td>FH</td>
<td></td>
<td>T</td>
<td>t</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>AB291965</td>
<td>c</td>
<td>t</td>
<td>t</td>
<td>T</td>
<td>C</td>
<td>t</td>
<td>a</td>
<td>C</td>
<td>c</td>
<td>t</td>
<td>T</td>
<td>T</td>
<td>t</td>
<td>c</td>
<td>A</td>
<td>c</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>AB291966</td>
<td>c</td>
<td>t</td>
<td>t</td>
<td>T</td>
<td>C</td>
<td>t</td>
<td>a</td>
<td>C</td>
<td>c</td>
<td>t</td>
<td>C</td>
<td>t</td>
<td>t</td>
<td>c</td>
<td>A</td>
<td>c</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>AB291967</td>
<td>FH</td>
<td>c</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>c</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>t</td>
<td>c</td>
<td>A</td>
<td>c</td>
</tr>
<tr>
<td>AB291968</td>
<td>c</td>
<td>t</td>
<td>C</td>
<td>a</td>
<td>t</td>
<td>C</td>
<td>G</td>
<td>t</td>
<td>c</td>
<td>C</td>
<td>T</td>
<td>c</td>
<td>t</td>
<td>c</td>
<td>g</td>
<td>c</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>AB193176</td>
<td>FH</td>
<td>c</td>
<td>t</td>
<td>C</td>
<td>T</td>
<td>t</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>c</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>c</td>
<td>A</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>AB220971</td>
<td>FH</td>
<td>c</td>
<td>t</td>
<td>C</td>
<td>T</td>
<td>t</td>
<td>C</td>
<td>G</td>
<td>c</td>
<td>c</td>
<td>T</td>
<td>c</td>
<td>T</td>
<td>C</td>
<td>c</td>
<td>A</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>AB220972</td>
<td>FH</td>
<td>c</td>
<td>t</td>
<td>C</td>
<td>T</td>
<td>c</td>
<td>C</td>
<td>a</td>
<td>c</td>
<td>t</td>
<td>T</td>
<td>t</td>
<td>c</td>
<td>A</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>AB220973</td>
<td>FH</td>
<td>c</td>
<td>t</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>a</td>
<td>c</td>
<td>c</td>
<td>t</td>
<td>T</td>
<td>T</td>
<td>t</td>
<td>c</td>
<td>A</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td>AB091395</td>
<td>FH</td>
<td>c</td>
<td>t</td>
<td>C</td>
<td>T</td>
<td>t</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>c</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>c</td>
<td>A</td>
<td>T</td>
<td>C</td>
</tr>
<tr>
<td>JQ740781</td>
<td>FH</td>
<td>c</td>
<td>a</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>t</td>
<td>G</td>
<td>C</td>
<td>c</td>
<td>C</td>
<td>T</td>
<td>c</td>
<td>t</td>
<td>c</td>
<td>A</td>
<td>c</td>
<td>t</td>
</tr>
<tr>
<td>Other human</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>genotype 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>T1</td>
<td>C3</td>
<td>C12</td>
<td>T23</td>
<td>C12</td>
<td>G15</td>
<td>C25</td>
<td>T9</td>
<td>C18</td>
<td>T27</td>
<td>T18</td>
<td>C10</td>
<td>c27</td>
<td>A27</td>
<td>T16</td>
<td>C20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>t21</td>
<td>t13</td>
<td>c4</td>
<td>a6</td>
<td>a8</td>
<td>a11</td>
<td>t2</td>
<td>c18</td>
<td>t9</td>
<td>c9</td>
<td>t11</td>
<td>c11</td>
<td>t7</td>
<td>a2</td>
<td>g2</td>
<td>g3</td>
</tr>
</tbody>
</table>

Nucleotide substitutions present in a FH patient but not in a non-FH patient infected from the same source are capitalised. Nucleotides are numbered relative to AB220978; Syn, synonymous; NS, non-synonymous; 2/3 – overlap between ORF2 and ORF3 where substitutions are synonymous in one frame and non-synonymous in the other. Sequences derived from a common source of infection are enclosed between bold lines. The number of isolates with each nucleotide present amongst other non-FH genotype 4 sequences (n = 27) is indicated.
Similar re-analysis of genotype 4 complete genome sequences from FH (n = 8) and non-FH (n = 30) patients revealed only two sites significantly associated with FH (positions 1963 and 4795, both synonymous). These FH-associated substitutions were also present in >41% of the non-FH sequences and differed from the two sites reported from analysis of 22 isolates (10), or the eight sites reported from analysis of 28 isolates (11). Position 1963 differed between one FH and two non-FH isolates derived from a common source of infection (Table 1). Removing two or three of the most closely related Hokkaido sequences left no sites significantly associated with FH. The frequency of the double (synonymous) substitution U3148 and C5907 in FH isolates reached significance at the 5% level in Fisher’s exact test (5/8 vs. 6/24, $P = 0.031$); a previous study reported a much lower $P$ value of 0.0042 (11).

Repeating all these analyses using a different test for association, the Metadata-comparison analysis tool (meta-CATS, www.viprbrc.org) (15) that uses the chi-squared test of independence and Pearson’s chi-squared test revealed fewer sites significantly associated with FH for both the full genotype 1 data set (45 significant sites compared to 110), and the full genotype 4 data set (0 significant sites compared to 2).

Discussion

In contrast to several previous studies (6, 7, 9–12), we have been unable to identify specific HEV strains or genomic substitutions that are associated with FH, although the presence of synonymous substitutions at positions 3148 and 5907 (11) was significantly associated with FH at a reduced level of significance (5%). Progression to FH is not a genotype-specific property of HEV as all four of the genotypes currently known to infect humans have been associated with FH. In addition, FH is not a strain-specific property of particular lineages within each HEV genotype (Fig. 1). Not all individuals infected from a common source develop FH and neither do particular substitutions appear to be associated with FH as no common mutations were observed in individuals with and without FH infected from a common source. Our re-analysis of the association between FH and substitutions at particular positions in the genome of genotype 1 and genotype 4 viruses suggests that previously reported significant associations may have been influenced by the restricted geographical and temporal sampling of FH isolates.

The identification of virus lineages or substitutions associated with the development of FH would be more difficult if this trait had incomplete penetrance. However, in this case, many more isolates from FH patients would be required. For example, a recent survey of virus determinants of FH following infection with hepatitis B virus that included 50 cases of FH with age- and sex-matched controls was able to demonstrate an association with the G1896A pre-core mutation (26), a substitution found in several other common source outbreaks in which there was a high frequency of FH (27). A different study identified two different substitutions (T1961 not T and C1962 not C) as significantly associated with FH (28). On the other hand, a study of 10 FH cases following acute infection with hepatitis A virus did not detect any difference in virus genotype between patient groups (29).

Although FH can develop following infection with any of the HEV genotypes known to infect humans, evidence that FH might be associated with particular genotypes comes from a survey of HEV in Japan over the last decade (30). Of 199 HEV patients, seven developed FH, these comprising 8.1% of those infected with genotype 4 but only 0.8% of those infected with genotype 3. However, the northerly island of Hokkaido which has <5% of Japan’s population contributed 70% of the genotype 4 infections, and almost 70% of HEV infections there could be attributed to the consumption of uncooked pig liver, while in other regions most infections had no known source. Infection with HEV genotype 4 has also been associated with more severe disease (compared to genotype 3) amongst patients from France (5). In both cases, it is possible that the reported association between FH and HEV genotype 4 could reflect underlying epidemiological factors rather than a difference in virus pathogenicity. Experiments in animal models are currently limited to genotypes 1, 2 and 3 (31).

Alternatively, the development of FH consequent to HEV infection could be because of patient-specific factors. A precedent for this comes from previous studies showing a correlation between the severity of HEV infection (including FH) and pregnancy (32, 33), pre-existing liver disease (2, 34–37) and either a low (38) or high viral load (39). More detailed study of HEV infected FH patients from around the world should provide a definitive answer to the role of virus variation in this aspect of pathogenesis.

Acknowledgements

We are grateful to Professor Jihong Meng for providing additional information in relation to AY230202, to Dr Philippe Colson for confirming that EF061404 was sampled from a patient with fulminant HEV, and to Dr Al Ivens for induction into the mysteries of R. This work was supported by a grant from The Welcome Trust to the Centre for Immunity, Infection and Evolution at Edinburgh University.

Conflict of interest: The authors do not have any disclosures to report.

References

2. Mizuo H, Yazaki Y, Sugawara K, et al. Possible risk factors for the transmission of hepatitis E virus and for the severe


**Supporting information**

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** Phylogenetic analysis of subgenomic HEV regions from FH and non-FH patients. Maximum composite likelihood distances between sequences are presented as neighbour joining trees for (A) genotype 1 (positions 4295–4601), (B) genotype 1 (4460–4786), (C) genotype 4 (123–448), (D) genotype 4 (4001–4821), (E) genotype 4 (5979–6390) and (F) genotype 3 (6034–6335). Sequences isolated from FH patients are indicated by ● and the genotype 2 outgroup (M74506) by □. The percentage of bootstrap replicates supporting individual branches is indicated.