Adeno-associated virus 2 infection in children with non-A–E hepatitis

An outbreak of acute hepatitis of unknown aetiology in children was reported in Scotland in April 2022 and has now been identified in 35 countries. Several recent studies have suggested an association with human adenovirus with this outbreak, a virus not commonly associated with hepatitis. Here we report a detailed case–control investigation and find an association between adeno-associated virus 2 (AAV2) infection and host genetics in disease susceptibility. Using next-generation sequencing, PCR with reverse transcription, serology and in situ hybridization, we detected recent infection with AAV2 in plasma and liver samples in 26 out of 32 (81%) cases of hepatitis compared with 5 out of 74 (7%) of samples from unaffected individuals. Furthermore, AAV2 was detected within ballooned hepatocytes alongside a prominent T cell infiltrate in liver biopsy samples. In keeping with a CD4+ T-cell-mediated immune pathology, the human leukocyte antigen (HLA) class II HLA-DRB1*04:01 allele was identified in 25 out of 27 cases (93%) compared with a background frequency of 10 out of 64 (16%; $P = 5.49 \times 10^{-12}$). In summary, we report an outbreak of acute paediatric hepatitis associated with AAV2 infection (most likely acquired as a co-infection with human adenovirus that is usually required as a ‘helper virus’ to support AAV2 replication) and disease susceptibility related to HLA class II status.

In April 2022, several hospitals in Scotland reported that children were presenting to medical practitioners with acute severe hepatitis of unknown aetiology (Fig. 1a). Elsewhere in the United Kingdom, 270 similar presentations were subsequently reported, for which 15 children required liver transplantation. The World Health Organization (WHO) has now registered 1,010 probable cases that fulfil their definition of this illness in 35 countries. Understanding the underlying cause of this new disease is a global public health imperative.

Detailed clinical investigations carried out as part of the public health response excluded common causes of acute hepatitis, including viral hepatitis, drug toxicity and autoimmune hepatitis. However, recent or active human adenovirus (HAdV) infection was identified in a high proportion of cases in Scotland, England and the United States. This finding was notable because HAdV is not a common cause of hepatitis. An increase in HAdV diagnoses in Scotland directly preceded the outbreak of unexplained hepatitis in children of a similar age (Fig. 1a,b).
SARS-CoV-2 had been circulating for 2 years and peaked several months before the increase in hepatitis cases3 (Fig. 1c). Human herpesvirus 6 (HHV6A and HHV6B) infections were not detected at higher levels during 2021 or 2022 (Fig. 1d).

**Research investigation**

To investigate the aetiology of these cases of acute hepatitis, we recruited 32 affected children who presented to a hospital between 14 March 2022 and 21 September 2022.
and 20 August 2022 and met the Public Health Scotland case definition criteria for inclusion in the International Severe Acute Respiratory and Emerging Infections Consortium (ISARIC) WHO Clinical Characterization Protocol United Kingdom (CCP-UK) (ISRCTN66726260). Samples from unaffected children (control samples) were obtained from the Diagnosis and Management of Febrile Illness using RNA Personalised Molecular Signature Diagnosis (DIAMONDS) study cohort and from the NHS Greater Glasgow & Clyde (Gg&C) Biorepository under appropriate ethics approval (Methods).

Clinical presentation

The median age of affected patients was 4.1 years (interquartile range (IQR) of 2.7–5.5 years) (Table 1). All patients were of white ethnicity, and 21 out of 32 (66%) were girls. Eighteen (56%) of the children reported a subacute history 2–12 weeks before acute hepatitis, which was characterized by an initial gastroenteritis-like illness followed by intermittent vomiting, abdominal pain and fatigue. The majority (23 out of 32) had no other medical conditions. One child had previously received a liver transplant, whereas none of the other patients were immunocompromised and none had received a COVID-19 vaccination. All routine blood tests for viral hepatitis, including hepatitis A, B, C and E, acute Epstein–Barr virus (EBV), cytomegalovirus (CMV), HHV6 and HHV7, and herpes simplex virus (HSV) were negative (Supplementary Table 1). Four patients had a low titre (1:80) of anti-nuclear antibodies and 3 patients had a low titre (1:40) of anti-smooth muscle antibodies, but other markers of autoimmunity were negative (Table 1 and Supplementary Table 2).

Following hospitalization, liver biopsy samples were obtained from five children. The samples showed evidence of lobular hepatitis with periportal and interface inflammation, intracellular inclusions, bile duct proliferation and ballooning of hepatocytes of varying severity (Fig. 1e–t). Mild-to-moderate fibrotic changes were noted, with no evidence of confluent fibrosis, and there was an inflammatory infiltrate that included cells expressing major histocompatibility complex class II (MHCI). Modified hepatic activity index scores (Ishak system)8,9 ranged from 6 to 11 (Extended Data Table 1), and the biopsy samples stained negative for complement.

Four patients required transfer to a specialist liver unit owing to significant synthetic liver dysfunction. Two of the patients were treated with steroid therapy and improved. One patient received supportive care only and spontaneously improved. The fourth patient had severe disease and required liver transplantation and was treated with cidofovir for HAdV viraemia and steroids after the liver transplant. The remaining 28 patients received supportive care only, with no antiviral or steroid treatment, and all showed gradual resolution of hepatitis over 2–3 months. There were no deaths. The median duration of hospital stay was 6 days (range of 1–68 days) (Table 1). In the patients with weakly positive autoantibodies, all had normal or normalizing transaminase levels at last follow up in the absence of treatment with an anti-inflammatory or immunosuppressant.

Pathogen detection by sequencing

As the epidemiology was in keeping with the emergence of an infectious pathogen, we undertook metagenomics and target enrichment (TE) next-generation sequencing (NGS) on all available clinical samples from the first nine recruited patients. The samples included plasma (n = 9), liver biopsy samples (n = 4), throat swabs (n = 6), faecal samples (n = 7) and a rectal swab (n = 1), and an average of 14 million sequence reads per sample were obtained (Fig. 2a–d). The samples were obtained between 7 and 80 days after initial symptom onset. Samples from the control group were restricted to children recruited in the United Kingdom between January 2020 and April 2022. Two comparison groups were used as controls: group 1 comprised serum or plasma samples from 13 age-matched healthy children (10 boys, 3 girls; age range of 3–5 years); and group 2 comprised serum or plasma samples from 12 children (8 boys, 4 girls; age range of 1–4 years) with HAdV infection confirmed by PCR and with normal transaminase levels. The children in group 2 had been diagnosed by nasopharyngeal aspirate (n = 10), by nose swab (n = 1) or by stool (n = 1) as part of the routine clinical investigation process and half of the patients required critical

### Table 1 | Demographic and clinical characteristics of the 32 patients with unexplained hepatitis

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)a</td>
<td>4.1 (2.7–5.5, 0.9–10.6)</td>
</tr>
<tr>
<td>Sex (girls)b</td>
<td>20 (63%)</td>
</tr>
<tr>
<td>Co-morbidityc</td>
<td>9 (28%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Biochemistry</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak bilirubinb (µmol l−1)</td>
<td>82 (36–160, 3–387)</td>
</tr>
<tr>
<td>Peak alanine transaminase (U l−1)</td>
<td>1,757 (708–2,763, 333–5,417)</td>
</tr>
<tr>
<td>Peak aspartate transaminase (U l−1)</td>
<td>2,048 (833–3,408, 424–6,908)</td>
</tr>
<tr>
<td>Peak γ-glutamyltransferase (U l−1)</td>
<td>124 (91–162, 18–720)</td>
</tr>
<tr>
<td>Peak international normalized ratio</td>
<td>1.2 (1.1–1.4, 1.0–2.9)</td>
</tr>
<tr>
<td>Peak C-reactive protein (mg l−1)</td>
<td>5 (3–11, 1–117)</td>
</tr>
<tr>
<td>Caeruloplasmin (n = 24) (g l−1)</td>
<td>0.36 (0.33–0.39, 0.22–0.52)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Key autoimmune parameters</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IgG (g l−1)</td>
<td>11.8 (9.9–14.3, 1.5–21.0)</td>
</tr>
<tr>
<td>Coeliac screen (TTG antibody) (n = 26)</td>
<td>Normal range</td>
</tr>
<tr>
<td>Anti-mitochondrial antibody</td>
<td>Negative</td>
</tr>
<tr>
<td>Anti-smooth muscle antibody</td>
<td>29 negative, 3 low positive (1:40)c</td>
</tr>
<tr>
<td>Anti-liver kidney microsomal antibody</td>
<td>Negative</td>
</tr>
<tr>
<td>Anti-nuclear antibody</td>
<td>28 negative, 4 weak positive</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clinical presentation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms at presentationc</td>
<td></td>
</tr>
<tr>
<td>• Vomiting</td>
<td>22 (69%)</td>
</tr>
<tr>
<td>• Jaundice</td>
<td>21 (66%)</td>
</tr>
<tr>
<td>• Poor appetite</td>
<td>12 (38%)</td>
</tr>
<tr>
<td>• Lethargy or fatigue</td>
<td>10 (31%)</td>
</tr>
<tr>
<td>• Abdominal pain</td>
<td>10 (31%)</td>
</tr>
<tr>
<td>• Diarrhoea</td>
<td>4 (13%)</td>
</tr>
<tr>
<td>Subacute symptoms for ≥14 days before presentation (n = 32)</td>
<td>18 (56%)</td>
</tr>
<tr>
<td>Subacute symptoms reported (n = 18)</td>
<td></td>
</tr>
<tr>
<td>• Intermittent vomiting</td>
<td>15 (83%)</td>
</tr>
<tr>
<td>• Initial gastroenteritis-like illness</td>
<td>12 (67%)</td>
</tr>
<tr>
<td>• Abdominal pain</td>
<td>9 (50%)</td>
</tr>
<tr>
<td>• Lethargy or fatigue</td>
<td>7 (39%)</td>
</tr>
<tr>
<td>• Poor appetite</td>
<td>6 (33%)</td>
</tr>
<tr>
<td>• Weight loss</td>
<td>6 (33%)</td>
</tr>
<tr>
<td>Approximate duration of subacute symptoms before presentationd</td>
<td>42 (27–52, 14–85) days</td>
</tr>
<tr>
<td>Length of hospital staye</td>
<td>6 (4–10, 1–68) days</td>
</tr>
<tr>
<td>Required transfer to tertiary liver unitf</td>
<td>4 (12.5%)</td>
</tr>
<tr>
<td>Required liver transplant</td>
<td>1 (3%)</td>
</tr>
</tbody>
</table>

aMedian (IQR, range). bNumber (%) denominator = 32 unless otherwise specified. cSee Supplementary Information for additional clinical details. dN = 16 patients with data available. eN = 30, one patient was a long-term in-patient for an unrelated condition, one patient was managed as an outpatient.

Nature | www.nature.com | 3
Fig. 2 | Detection of AAV2 in cases of paediatric hepatitis. a, Heatmap of AAV2 and HAdV read counts detected in cases of hepatitis by TE sequencing. Samples obtained for routine clinical investigation (plasma, liver, faecal, rectal swab and throat swab) were retrospectively sequenced following DNA or RNA extraction. AAV2 read counts are shown from 0 to >50 reads per million in green (top rows) and HAdV read counts are shown from 0 to >5 reads per million in red (bottom rows). b, Heatmap of viral reads of plasma samples from cases of hepatitis and of plasma or sera samples from controls. Plasma samples from cases of hepatitis (cases), and plasma or sera samples from children with HAdV infection (group 2 controls) and from age-matched healthy children (group 1 controls) were sequenced following DNA or RNA extraction. AAV2 read counts are shown from 0 to >50 reads per million in green and HAdV read counts are shown from 0 to >5 reads per million in red. The number of days between initial symptom onset and sample are indicated. c, AAV2 real-time RT–qPCR of serum or plasma samples from 32 cases of hepatitis (cases) and from 74 controls in four groups: 13 in group 1 (healthy controls); 12 in group 2 (HAdV-positive controls); 33 in group 3 (hepatitis controls); and 16 in group 4 (contemporaneous controls). The detection threshold of the assay (3,200 copies per ml) is shown as a dotted line. Values are shown as a scatter plot with a median line. d, AAV2 real-time RT–qPCR of liver biopsy samples from 5 cases of hepatitis and from 19 controls. e, IgM responses determined by ELISA in 22 cases of hepatitis and in 29 controls (13 in group 3, 16 in group 4). f, IgG responses determined by ELISA in 22 cases of hepatitis cases and in 29 controls (13 in group 3, 16 in group 4). For e–f, statistical analysis was performed using Mann Whitney test (two-tailed), and experiments were performed in triplicate.
Sequence and phylogenetic analyses

Near-full genomes of AAV2 were obtained from all nine patients with hepatitis (GenBank accession numbers OP019741–OP019749), and in all cases, two large open reading frames corresponding to the rep and cap genes, flanked by inverted terminal repeat sequences, were identified. Seven distinct sequences of AAV2 were noted (Extended Data Fig. 2), forming a single clade, alongside four AAV2 genomes previously detected in France between 2004 and 2015. Two out of three identical sequences were known to have come from individuals from the same household, therefore these two are epidemiologically linked. The third sequence was from a sample obtained around the same time but was not known to be linked to the other cases. Sequences from the liver samples matched those detected in plasma. Several mutations within the VP1–VP3 genes were noted to be over-represented in the sequences derived from patients with hepatitis when compared with reference sequences (Extended Data Fig. 2). Notably, nine of the mutations in the capsid gene that were over-represented in the cases of hepatitis (V151A, R447K, T450A, Q457M, S492A, E499D, F533Y, R585S and R588T) are associated with an AAV2 variant that has an altered phenotype. Characteristics of this variant include substantial evasion of neutralizing antibodies directed against wild-type AAV2, enhanced production yields, reduced heparin binding, increased virion stability and more localized spread in a mouse model.

A full genome of HAdV-F41 was obtained from a faecal sample (GenBank accession number OP019750) and was found to be closest phylogenetically to two genomes reported from Germany in 2019 and 2022 (Extended Data Fig. 2). Contigs matching to other human pathogens, including human coronavirus NL63, rhinovirus C, enterovirus B, human parainfluenza viruses 2 and 3, norovirus, and both betaherpesviruses and gammaherpesviruses were also detected across cases, albeit not consistently. These findings were confirmed by PCR (Supplementary Table 1).

Confirmatory PCR testing of cases of hepatitis

PCR testing for AAV2 was positive in all nine initial cases of hepatitis. Standards were used to estimate the viral loads of positive samples (Supplementary Fig. 2). All nine plasma samples tested negative by PCR for HHV6, HSV, CMV and EBV. Two out of the four liver biopsy specimens tested positive for HHV6 (cycle threshold (Ct) values of 33 and 36) (Supplementary Table 1). HAdV was detected in 3 out of 9 plasma samples, 3 out of 4 liver biopsy samples, 2 out of 6 throat swabs, 4 out of 7 faecal samples and 1 out of 1 rectal swab. The lower detection of HAdV and HHV6 by PCR compared with TE sequencing probably reflects a slightly lower sensitivity of the PCR assay. The low numbers of HAdV-positive samples detected using both assays may reflect the fact that plasma is a suboptimal sample type for HAdV detection (whole blood samples were unavailable).

Case–control study

To investigate the presence of AAV2 and the candidate helper viruses HAdV and HHV6B in plasma samples from cases of hepatitis, we undertook a case–control study in which samples from 32 cases of hepatitis were compared with samples from the group 1 and group 2 controls described above and with samples from two additional control groups (Fig. 2a–f). Group 3 controls comprised 33 children (18 boys and 15 girls aged 2–16 years) with increased transaminase levels that had tested negative by PCR for HAdV. This group was used to test the hypothesis that reactivation of AAV2 may occur in children with severe hepatitis and may be a correlate of liver dysfunction. The children comprising group 3 were older (median age of 10.2 years, IQR of 7–13.6 years, P < 0.001) than the patients from the case group (Extended Data Table 2b) and 15 out of 33 had required critical care for ventilatory or cardiovascular support. Group 4 controls comprised residual plasma or serum samples from 16 children in Scotland aged 10 years and were attending hospital contemporaneously with the children with hepatitis between March and April 2022. The group 4 controls were used to determine whether AAV2 was circulating widely in children in healthcare facilities across Scotland at the time the children with hepatitis were admitted to hospital. Clinical details, including liver function were not available for this group. To ensure that the quantification of AAV2 was accurately performed, we confirmed standard curve concentrations using droplet digital PCR (Methods).

Significance differences between groups for viral loads in plasma samples were calculated using a Mann–Whitney test (two-tailed). RT–qPCR of plasma samples showed that 26 out of 32 cases of hepatitis were positive for AAV2, with a median estimated copy number of 66,100 copies per ml (IQR of 13,461–300,277 copies per ml), a value higher than samples from all the control groups (P < 0.001 for all case–control comparisons). The median copy number in control groups 1–3 was below the detection limit. A median of 3,268 copies per ml (detection threshold of 3,200 copies per ml) was present in samples from control group 4, which suggested that AAV2 was circulating at low levels in children during March and April 2022 (Fig. 2c). Although five plasma samples from cases of hepatitis were positive for HAdV by PCR, and one tested positive by PCR for HHV6 DNA, these results were not significantly more common than in samples from the control group (Supplementary Fig. 3).

Next, five liver biopsy samples from cases of hepatitis were compared with 19 residual liver biopsy samples (controls) from children under 18 years old. The median AAV2 viral load was 3,721,497 copies per mm³ of liver (IQR of 3,308,243–6,717,616 copies per mm³) in cases of hepatitis compared with 64 copies per mm³ of liver (IQR of 20–83 copies per mm³) in samples from the control group (P < 0.001; Fig. 2d). Glyceraldehyde-3-phosphate dehydrogenase was used as a marker of extraction efficiency in all samples, and results were similar between the case and control groups. When outliers were removed, significance was retained (Supplementary Data 2, Supplementary Fig. 4).

Longitudinal sampling

To investigate AAV2 viraemia and liver function values over time, longitudinal PCR testing was performed in 14 cases of hepatitis from whom multiple retrospective plasma samples were available (Supplementary Fig. 5). Spearman’s rank correlation coefficients for the relationships
between the trajectories of viral load and alanine transaminase and bilirubin were positive for most cases. However, overall statistical significance could not be confirmed owing to the sample size.

Where samples were available, we screened for the presence of AAV2-specific IgM and IgG antibodies within samples from patients and samples from the group 1 healthy controls and group 4 contemporaneous controls (Fig. 2e,f and Supplementary Fig. 6). Anti-AAV2 IgM was detected in 15 out of 23 (65.2%) samples from cases of hepatitis, but only 1 out of 13 (7.7%) samples from group 1 healthy controls and 2 out of 16 (12.5%) samples from the group 4 contemporaneous controls from Scotland. For the samples from cases of hepatitis that tested negative for AAV2-specific IgM, samples from four patients were noted to be obtained fewer than 3 days after the onset of illness and samples from two patients were obtained more than 77 days after the onset of illness. IgG was detected in 21 out of 23 (91.3%) samples from cases of hepatitis, in 8 out of 13 (61.5%) samples from age-matched healthy controls (group 1) and in 9 out of 16 (56.3%) samples from healthy controls from Scotland (group 4). Of the two samples from patients who tested seronegative, both were obtained at early time points, probably sampled before expected seroconversion (less than 3 days after the onset of illness).

SARS-CoV-2 infection
Routine clinical investigation detected SARS-CoV-2 nucleic acid in nasopharyngeal samples from 3 out of 31 (9.6%) children at the time of illness, 2 of whom were also seropositive. The third became infected after the onset of hepatitis. SARS-CoV-2 was not detected by PCR or by sequencing in any of the samples from cases or controls available for analysis, including liver samples. Nevertheless, to investigate the possibility that unexplained hepatitis in children might relate to a previous infection with SARS-CoV-2 or other seasonal coronaviruses, we carried out serological analysis of 23 available residual samples from cases. IgG antibody titres were quantitatively measured against the spike protein, the amino-terminal domain (NTD) and receptor binding domain (RBD) of the spike protein and the nucleocapsid of SARS-CoV-2. IgG antibody titres were also measured for human seasonal coronaviruses 229E, OC43, NL63 and HKU1. Electrochemiluminescence assays (MesoECL) for coronavirus-specific IgG revealed previous exposure to seasonal coronaviruses, with strong responses detected against NL63 (17 out of 23) and OC43 (21 out of 23) (Extended Data Fig. 3a). By comparison, plasma samples from 12 out of 23 children displayed high reactivity against HKU1, whereas only 3 out of 23 samples reacted strongly against 229E. Plasma samples from 11 children reacted with 2 or more SARS-CoV-2 antigens (nucleocapsid, spike protein, NTD or RBD). One of the samples reacted solely with the nucleocapsid antigen, which indicated that in total, 12 out of 23 patients displayed serological evidence of previous exposure to SARS-CoV-2 (Extended Data Fig. 3b). In summary, 12 out of 23 (52%) of the children with hepatitis displayed evidence of previous exposure to SARS-CoV-2. This level is lower than SARS-CoV-2 seroprevalence in children aged 5–11 years in Scotland between 14 March and 27 June 2022 (when Public Health Scotland enhanced surveillance for COVID-19 was discontinued), which was reported as between 59.0% (95% confidence interval (CI) of 50.6–71.2) and 72.4% (95% CI of 53.9–78.8). This result indicates that there is no direct link between COVID-19 and the outbreak of acute hepatitis studied here.

Host genetics and HLA typing
We next investigated whether some children might be genetically more susceptible to non-A–E hepatitis. To that end, 27 samples from cases of hepatitis and 64 platelet apheresis samples from local donors in Scotland (controls) were genotyped using high-resolution typing for all HLA loci (HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DQA1, HLA-DQB1, HLA-DPA1 and HLA-DPB1). In total, 25 out of 27 (92.6%) samples from patients with hepatitis were positive for at least one copy of the HLA-DRB1*04:01 allele compared with 10 out of 64 (15.6%) of samples from controls. The allele frequency in patients was 0.54 compared with 0.08 in controls (odds ratio (OR) of 13.7 (95% CI of 5.5–35.1), \(P = 5.49 \times 10^{-12}\)). The frequency of the HLA-DRB1*04:01 allele (based on an imputation of HLA alleles) in a control set of unrelated participants from the UK Biobank (\(n = 29,379\)) was 0.11 (2.942 out of 29,379 allele carriers, OR of 112.3 (95% CI of 26.6–474.5), \(P = 3.27 \times 10^{-3}\). The frequency was also 0.11 in British/Irish North-West European individuals from the Anthony Nolan charity register13. To check for cryptic relatedness among patients and population stratification, we performed genome-wide microarray genotyping in 19 cases of hepatitis and excluded participants with a conservatively relatedness threshold (identity-by-state > 0.4). When compared with well-matched participants from the UK Biobank (Extended Data Fig. 4), similar signals for association with disease by allele frequency (\(P = 8.96 \times 10^{-5}\) and across the three possible biallelic genotypes at this locus (\(P = 1.2 \times 10^{-7}\) were obtained.

In addition to the association with the DRB1 allele, 23 out of 27 samples from patients with hepatitis were positive for HLA-DQA1*03:03 compared with 11 out of 64 samples from controls (allele frequency of 0.54 compared with 0.09, respectively, OR of 12.3 (5.1–30.7), \(P = 1.9 \times 10^{-11}\)). Moreover, 26 out of 27 samples from patients were positive for HLA-DRB4*01:03 compared with 21 out of 64 samples from controls (allele frequency of 0.67 compared with 0.17, respectively, OR of 9.4 (4.4–21.3), \(P = 1.8 \times 10^{-4}\)). Owing to strong linkage disequilibrium in this region of the genome, it is not possible to be certain which is the causal susceptibility allele.

In situ hybridization and immune typing
To investigate the presence of AAV2, HAdV and HHV6 in liver biopsy samples, we carried out in situ hybridization (ISH). Liver biopsy samples of all patients were characterized by the presence of AAV2 RNA within the nuclei and cytoplasm of ballooned hepatocytes and in arterial endothelial cells, which is indicative of the presence of replicating virus (Fig. 3a–h). AAV2-positive cells were quantified at a high level in all cases using QuPath in biopsy samples from five non-A-non-E hepatitis, ranging from 1.2 to 4.7. This level is similar to that seen in hepatitis associated with other viruses2,23. Consistent with low levels of HHV6B and HAdV sequence reads present in the biopsy samples from cases of hepatitis, negligible levels of viral RNA from these viruses were detected by ISH.

To investigate the possibility of an immune-mediated pathogenesis of disease in the liver, multiplex analysis of liver samples was carried out using co-detection by indexing (CODEX) for various immune cellular markers, including CD3, CD4, CD8, PD-L1, CD107a, CD20, CD31, CD44, CD68, MX1 and PanCK (Fig. 4a–d and Supplementary Figs. 7 and 8). In the explant liver sample of patient CVR35, prominent disordered proliferation of epithelial cells throughout the liver tissue was evident, with increased numbers of CD68+ macrophages, activated CD4+ and CD8+ T cells and CD20+ B cells. High expression of the interferon-induced GTP-binding protein MX1 was also noted, which indicated that the innate immune response was activated.

Conclusions and final statements
In this study, we reported the association of AAV2 infection and the class II HLA allele HLA-DRB1*04:01 with an outbreak of paediatric non-A–E hepatitis, with virus being detected independently by sequencing, real-time PCR and ISH. Liver biopsy tissue samples from all patients were characterized by the presence of AAV2 RNA (indicating replicating virus) within the nucleus and cytoplasm of ballooned hepatocytes and a dense infiltrate of CD4+ and CD8+ T cells in the liver with an
Fig. 3 | ISH of AAV2 in liver tissue. a–g, RNA ISH for the detection of AAV2 RNA in sections of formalin-fixed and paraffin-embedded liver tissues from children (one section per patient) with non A–E hepatitis. a, AAV2 RNA (red signal, indicated by an arrow) was detected in the endothelial cells of arteries in an explant liver section from patient CVR35. The vascular lumen is highlighted by an asterisk. b, A positive AAV2 signal was detected in the nuclei of hepatocytes with vacuolated morphology from patient CVR4 (indicated by arrows) and in a negative cell (indicated by the circle). c, d, A liver section from patient CVR1 showed AAV2 RNA both in the nucleus and in the cytoplasm (c), whereas for patient CVR9 (d), AAV2 RNA was found only in the nucleus (indicated by arrows). e, A high percentage of hepatocytes with a positive signal for AAV2 was present predominantly in the nucleus of hepatocytes in the samples from patient CVR1. f, AAV2 was not detectable in liver sections from samples from healthy individuals in either the endothelial cells or hepatocytes. g, Samples from patient CVR35 showed inclusion bodies in hepatocytes. Left, small, dark basophilic intranuclear inclusions next to the nucleus (indicated by arrows). Right, a large, pale basophilic, diffuse intranuclear inclusion body (suggestive of adenovirus infection; indicated by an arrow) next to a multinucleated giant cell in the liver (indicated by the asterisk). h, AAV2 positive cells were quantified using QuPath in biopsy samples from five patients with non A–E hepatitis (cases) and from controls. Patient CVR35 (who received a liver transplant) is highlighted in red. The entire section, cells were segmented to identify the nuclei and cytoplasm, and the algorithm was tuned to detect red signals. All samples were analysed using the same algorithm. Scale bars, 25 μm (insets of c, d), 50 μm (a–e, f, g) or 200 μm (e).

AAV2 is a small non-enveloped virus with a single-stranded DNA genome of around 4,675 nucleotides in length and it belongs to the species adeno-associated dependoparvovirus A (genus Dependoparvovirus, family Parvoviridae). It was first described in 1965 and infects up to 80% of the adult population. Seroreversion occurs in early childhood following respiratory infection. In a prospective study in the United States, the earliest seroconversion to AAV2 infection occurred in a 9-month-old child, and its seroprevalence increased from 24.2% to 38.7% in 3-year-old and 5-year-old children, respectively. This age range coincides with that of the cases in this study, which suggests that illness may be related to primary infection with AAV2 rather than its reactivation. In line with this hypothesis, anti-AAV2 IgM reactivity was observed in the majority of affected children. AAV2 relies on co-infection with a helper virus for replication, most commonly HAdV or a herpesvirus. Most clinical samples taken at presentation with hepatitis were obtained more than 20 days after initial symptom onset, which could explain the absence of a helper virus in some samples and low viral loads in positive samples. In an exploratory study using NGS, we detected two candidate helper viruses at low level in the cases of hepatitis: HAdV and HHV6B (in 6 out of 9 cases and in 3 out of 9 cases, respectively). These viruses were not confirmed to be higher in cases than controls in plasma or liver samples in our larger case-control study. HHV6B was also present in two control groups that included children with severe HAdV infection and children with hepatitis of alternative aetiology. As HHV6 can establish latency and can integrate its genome into the human chromosome, it may reactivate following concomitant illness (or immunosuppression) and may represent either an opportunistic bystander or a pathogen.

We propose that AAV2 is directly implicated in the pathology of the 2022 outbreak of non-A–E hepatitis in children, which occurred following transmission as a co-infection with HAdV or less likely due to reactivation following HAdV or HHV6 infection. Our results also support an association between HLA class II haplotype and disease susceptibility. A CD4+ T-cell-mediated response may direct maladaptive immunopathology mediated by T cytotoxic cells or B cells. In support of this notion, a CD8+ cell-mediated response directed against the AAV2 viral capsid (VPI) in association with hepatitis was reported in early trials of AAV2 when used as a vector for gene therapy. Hepatitis AAV2 is a common phenomenon in recipients of gene therapy vectored by AAV, and this side effect is usually treated pre-emptively with steroids before and for several weeks after the gene therapy; in rare cases, AAV-mediated gene therapy has been associated with deaths from fulminant hepatic failure. As a result of this current study, further studies are needed to investigate the association between HLA status with severe illness in gene therapy recipients. Notably, we did not find features of autoimmune hepatitis (AIH), either by serology or histology, in affected children. In a study of children from Scotland with AIH, the majority had evidence of seropositive disease (100% of patients with type II AIH tested positive for anti-LKM1). Furthermore, patients with AIH were older in age (median age of 11.4 years compared with 4.1 years in our cohort) and had significantly lower median alanine transaminase levels at diagnosis (444 IU per litre compared with 1,756 IU per litre). None of the AIH patients improved without treatment.

An alternative explanation for our findings is that AAV2 is not directly involved in pathology and is instead a biomarker of infection with HAdV. More than half of the patients with hepatitis in our study had subacute symptoms, with a median onset of 42 days before the onset of jaundice. The opportunity to detect virus by sequencing was therefore reduced, as samples were collected after this stage of illness. Furthermore, whole blood samples might have increased the sensitivity of detection, but only serum or plasma samples were available. We consider this alternative hypothesis to be less likely because we did not detect AAV2 in a control group of children with HAdV infection who had normal liver function tests. However, HAdV41 is a common cause of diarrhoea in young children, and co-infection of AAV2 with HAdV41 may explain early gastrointestinal symptoms in affected children. By
show differences in cellular composition (present.
epithelial cells are stained green using cytokeratin (CK). Scattered macrophages
ducts in the liver biopsy from the control are highlighted by asterisks, and
macrophages (red), T cells (cyan) and activated T cells (yellow).

Fig. 4 | CODEX analysis of liver tissue. a–d, Images of liver tissue from patient
CVR35 (b, d) and a liver sample from an unaffected individual (control; a, c)
show differences in cellular composition (c, d). a, Regularly structured bile
ducts in the liver biopsy from the control are highlighted by asterisks, and
epithelial cells are stained green using cytokeratin (CK). Scattered macrophages
(CD68, red), T cells (CD3, cyan) and activated T cells (CD44, yellow) are also
present. b, By contrast, the explant liver from patient CVR35 shows prominent
proliferation of epithelial cells throughout the liver tissue (green), with increased
macrophages (red), T cells (cyan) and activated T cells (yellow). c, The control
liver shows scattered cytotoxic T cells (CD8, red), CD107a-positive cells (brown)
and CD4-positive cells (yellow) and low expression of the interferon-
induced GTP-binding protein MX1 (green). d, High numbers of all cell types
and high MX1 expression are observed in the explant liver from patient CVR35.
One section of liver was stained per individual, and the entire area was manually
outlined. Cells were segmented to identify the nuclei and cytoplasm, and the
algorithm was tuned to detect the colour signal in the cells. All samples were
analysed using the same algorithm for each stain. Scale bars, 50 μm.

c–d
contrast, although adenovirus-associated hepatitis has been previously
described, particularly among immunocompromised individuals24,
HAdV41 has not previously been associated with severe hepatitis.
In the recent outbreaks of unexplained hepatitis in children, it has been
inconsistently associated4–6,25–27.

We also investigated the possibility that the unexplained cases
of hepatitis were linked to a previous illness with COVID-19. Direct
SARS-CoV-2-induced liver injury is unlikely though, as few of our cases
of hepatitis (3 out of 31) were positive for SARS-CoV-2 by PCR on admission
to hospital, and we did not identify SARS-CoV-2 by PCR or sequenc-
ing in any of the clinical samples from cases, including liver biopsies.
Furthermore, the SARS-CoV-2 seroprevalence in cases of hepatitis was
lower than in the community at that time. This result is in keeping with a

case–control analysis by the UK Health Security Agency3, who found no
difference in SARS-CoV-2 PCR positivity between cases of hepatitis
and children presenting to emergency departments between January
and June 2022. Nevertheless, we cannot at this time fully exclude a
post-COVID-19 immune-mediated phenomenon, for example, a link
to HLA class II type, in susceptible children.

There are several limitations to this study. First, the presence of
AAV2 in cases of hepatitis but not controls in groups 1–3 may have
arisen because of seasonal variation in AAV2 transmission, as some
children in the control groups were sampled earlier in the year than
for cases. We included a contemporaneous control group (group 4)
to address this possibility. Low viral loads of AAV2 were detected in
a small number of samples from the group 4 controls, which is in
keeping with the presence of the circulating virus in children at the
time the cases of hepatitis occurred. Second, the presence of AAV2 in
the cases of hepatitis is an association and may not represent direct
aetiology, and AAV2 may be a useful biomarker of recent HAdV (or
less likely HHV6B) infection. We do not consider it probable that
AAV2 simply represents a marker of liver damage because it was not
present in cases of severe hepatitis of alternative aetiology and, sig-
nificantly, we detected AAV2 in ballooned hepatocytes by ISH. The
strong association of the HLA-DRB1*04:01 allele, known to be associ-
ated with autoimmune hepatitis type 129 and extra-articular manifes-
tations of rheumatoid arthritis29, with the cases of hepatitis provides
support for a large impact of host genetics on susceptibility. How-
ever, this analysis was affected by strong linkage disequilibrium, and
larger studies are required to confirm a definitive association with
this allele. The association between HLA status and the presence of
an activated T cell infiltrate together with AAV2-infected cells in the
liver is in keeping with a CD4 + cell-mediated immune pathology30.

We consider autoimmune disease to be less likely of a cause of the cases
of hepatitis studied here because of the absence of autoantibodies and
the absence of typical histology in liver specimens. It is also plausible
that simultaneous HAdV infection with a co-infecting or reactivated
AAV2 infection has resulted, for a proportion of children who are more
susceptible (owing to the HLA class II allele HLA-DRB1*04:01), in a more
severe outcome than might typically be expected for these commonly
circulating viruses. Peptide mapping experiments are recommended
in future studies to investigate the nature of the HLA class II-restricted
T cell response.

The 2022 outbreak of AAV2-associated paediatric hepatitis that we
described in this study may have arisen because of changes in exposure
patterns to AAV2, HAdV and HHV6B as an indirect consequence of the
COVID-19 pandemic. The circulation of common human viruses was
interrupted in 2020 by the implementation of non-pharmaceutical
interventions, including physical distancing and travel restrictions,
instituted to mitigate SARS-CoV-2 transmission. Once restrictions
were lifted, genetically susceptible children may have had a higher chance
of being exposed to both HAdV and AAV2 for the first time, creating a
synchronized wave of severe disease. Larger case–control studies are
needed to confirm the role of AAV2 and HLA status in the aetiology
of unexplained non-A–E paediatric hepatitis. Retrospective testing of samples from sporadic cases of unexplained hepatitis in children is also needed.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-05948-2.

27. van Gerven, N. M. et al. HLA-DRB1*03:01 and HLA-DRB1*04:01 modify the presentation of DR4 by Felty's syndrome. Hum. Immunol. 975803 (2022).

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s). author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023
Methods

ISARIC CCP-UK recruitment, Biorepository and DIAMONDS studies

Ethics approval for the ISARIC CCP-UK study was given by the South Central-Oxford C Research Ethics Committee in England (13/SC/0149), the Scotland A Research Ethics Committee (20/SS/0028) and the WHO Ethics Review Committee (RPC571 and RPC572). Thirty-two children aged <16 years were prospectively recruited by written informed consent (parent or guardian) from the ISARIC WHO CCP-UK cohort admitted to hospital with increased transaminase levels (defined as alanine transaminase levels of >400 IU per litre and/or aspartate aminotransferase levels of >400 IU per litre) that was not due to viral hepatitis A–E, AIH or poisoning. Nine patients had available clinical samples for further investigation. Three additional patients had HLA typing performed, but samples were not available for further analysis. Samples for the control groups were obtained from children (aged <16 years) recruited to the DIAMONDS study, an ongoing multi-country study that aims to develop a molecular diagnostic test for the rapid diagnosis of severe infection and inflammatory diseases using personalized gene signatures (ISRCTN12394803). Ethics approval was given by the London-Dulwich Research Ethics Committee (20/HRA/1714). Controls included healthy individuals (n = 13; group 1), children with PCR-confirmed adenoviral infection with normal transaminase levels (n = 12; group 2), and children with increased transaminase levels without adenoviral infection (n = 33; group 3), recruited between 19 May 2020 to 8 January 2022. Surplus plasma samples from individuals in Scotland (aged <10 years; March to April 2022; group 4) and liver biopsy control samples (from individuals aged <18 years; January 2021 to July 2022) from the Diagnostic Pathology/Blood Sciences archive were obtained with NHS GG&C Biorepository approval (application no. 717; REC 22/WS/0020). Samples from adults that had tested negative by PCR for SARS-CoV-2 were used as an additional group for serological analysis of coronaviruses as a negative control group, also with NHS GG&C Biorepository approval. These adult samples were used without consent on the basis of Human Tissue Act legislation on consent exemption.

Viral PCR

RNA extraction was carried out using the protocol from Biomerieux Easymag. In total, 300 μl of plasma or sera was extracted and eluted into 80 μl of water.

AAV2 RT-qPCR was performed to detect a 62 bp amplicon of the AAV2 inverted terminal repeat region (ITR) as previously described using the forward ITR primer (5’-GGGCGCTAGTGATGAGTTGATGTT-3’) and the reverse ITR primer (5’-GGGCGCTAGTGATGAGTTGATGTT-3’). The AAV2 ITR hydrolysis probe was labelled with fluorescein (6FAM) and quenched with Black Hole quencher (BHQ) 5’-[6FAM]-CA CTCCCTCTCTGCGCGCTCG-[BHQ1]3’. AAV2 primers and probe were synthesized by Merck Life Sciences. RT–qPCR analysis was performed using an ABI7500 Fast Real-Time PCR system (Applied Biosystems). A LUNA Universal One-Step RT PCR kit (New England Biolabs) was used for the amplification and detection of the AAV2 ITR target. RT–qPCR assays were performed in a 20 μl volume reaction (Luna Universal One-Step reaction mix, Luna WarmStart RT enzyme mix, 400 nM forward and reverse primers, 200 nM AAV2 ITR probe and 1–2.5 μl of template DNA) as per the manufacturer’s instructions. To quantify the number of copies, serial dilutions of plasmid containing the 62 bp ITR product were used to generate a standard curve, which was then used to interpolate the copy number of AAV2 copies in the samples. Wells with no template were used as negative controls. RT–qPCR reactions were performed in triplicate. The RT–qPCR program consisted of an initial reverse transcription step at 55 °C for 10 min, an initial denaturation step at 95 °C for 1 min followed by 45 cycles of 95 °C denaturation for 10 s and extension at 58 °C for 1 min. A qPCR detection limit between 31 and 32 cycles was calculated as the threshold Ct value at the last dilution of DNA standards that were within the linear range. A PCR result was considered positive if all three reactions tested positive at ≤31 cycles.

Digital droplet PCR was performed according to the manufacturer’s instructions using the digital droplet PCR supermix for probes (no dUTP) (Bio-Rad, 1863023) and analysed using a QX200 Droplet Digital PCR system (Bio-Rad, 1864001).

The West of Scotland Specialist Virology Centre, NHS Greater Glasgow and Clyde, conducted diagnostic real-time RT–PCR to detect HAdV, SARS-CoV-2-positive samples and other viral pathogens associated with hepatitis (for example, hepatitis A–E) following nucleic acid extraction utilizing the NucliSENS easyMag and Roche MG96 platforms. HIV6 (ref. 32) and HAdV41 (ref. 33) were tested by qPCR as previously described using Invitrogen platinum qPCR mix (i1730-025) and Quanta Biosciences qPCR mix mastermix (733-1273), respectively, on an ABI7500 system and amplified for 40 cycles. A 6 μl extract was amplified in a total reaction volume of 15 μl.

Measurement of antibody response to coronaviruses by electrochemiluminescence

IgG antibody titres were quantitatively measured against the spike protein, the NTD, the RBD or nucleocapsid of SARS-CoV-2, and against the spike glycoproteins of human seasonal coronaviruses 229E, OC43, NL63 and HKU1 using MSD V-PLEX COVID-19 Coronavirus Panel 2 (K15369) and Respiratory Panel 1 (K15365) kits. Multiplex meso scale discovery electrochemiluminescence (MSD-ECL) assays were performed according to manufacturer’s instructions. Samples were diluted 1:5,000 in diluent and added to the plates along with serially diluted reference standard (calibrator) and serology controls L1, L2 and L3. Plates were read using a MESSO Sector S 600 plate reader. Data were generated using Methodological Mind software and analysed using MSD Discovery Workbench (v.4.0). Results are expressed as MSD arbitrary units per ml (AU ml−1). Adult negative and positive pools gave the following values: negative pool: spike, 56.6 AU ml−1; NTD, 119.4 AU ml−1; RBD, 110.5 AU ml−1; and nucleocapsid, 20.7 AU ml−1; SARS-CoV-2-positive pool: spike, 1,331.1 AU ml−1; NTD, 1,545.2 AU ml−1; RBD, 1,156.4 AU ml−1; and nucleocapsid, 1,549.0 AU ml−1 in the same assay. NIBSC 20/130 reference serum was used and the following values obtained: spike, 547.7 AU ml−1; NTD, 538.8 AU ml−1; RBD, 536.9 AU ml−1; and nucleocapsid 1,840.2 AU ml−1.

Metagenomics sequencing

Full protocols on the detection of RNA and DNA viruses using metagenomics NGS and TE sequencing methods can be found in refs. 32, 34.

In summary, residual nucleic acid from 27 samples from cases with hepatitis (from 9 patients with a combination of plasma, liver, faeces, rectal, and throat and nose samples), 12 samples from HAdV-positive individuals and 13 samples from healthy individuals (control samples were either plasma or sera) underwent metagenomics NGS sequencing at the MRC-University of Glasgow Centre for Virus Research Genomics facility. In brief, each nucleic acid sample was split into two library preparations to improve the chances of detecting RNA and DNA viruses. The protocol used to improve detection of RNA viruses included treatment with DNaseI (Ambion DNase I, ThermoFisher), ribosomal depletion (Ribo-Zero Plus rRNA Depletion Kit, Illumina), except for plasma samples, reverse transcription (SuperScript III, Invitrogen) and double-strand DNA synthesis (NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module, NEB). The protocol used to detect DNA viruses included partial removal of host DNA (NEBNext Microbiome DNA Enrichment Kit, NEB). Following this, both sets of samples were used to prepare libraries using a KAPA LTP kit (Roche) with unique dual indices (NEBNext Multiplex oligos for Illumina, NEB). The resulting libraries were pooled in equimolar amounts and sequenced using a NextSeq500 (Illumina) to obtain paired-end reads using 150 × 150 cycles.
**TE sequencing**
Following the library preparation step described above, DNA-derived and RNA-derived libraries were pooled separately and were incubated with VirCapSeq- VERT Capture Panel probes (Roche) following the manufacturer’s guidelines. The Roche VirCapSeq- VERT Capture Panel covers the genomes of 207 taxa of viruses known to infect vertebrates (including humans). Enriched DNA-derived and RNA-derived libraries were further amplified using 14 PCR cycles, then pooled and sequenced using a NextSeq500 (Illumina) to obtain paired-end reads using 150 × 150 cycles.

**Bioinformatics analysis**
Reads for each sample were first quality checked. Illumina adapters were trimmed using Trim Galore (https://github.com/FelixKrueger/TrimGalore) and then mapped to the human genome using BWA-MEM (https://github.com/lh3/bwa). Only reads that did not map to the human genome were used for metagenomics analyses. Reads per million were calculated as the number of viral reads per million reads sequenced to normalize for variation in sample sequencing depth. Non-human reads were then de novo assembled using MetaSPAdes (https://github.com/ablab/spades) to generate contigs for each sample. Contigs were compared against a protein database of all NCBI RefSeq organisms (including virus, bacteria and eukaryotes) with BLASTX using DIAMOND (https://github.com/bbuchfink/diamond). In addition, non-human reads for each sample were aligned to a small panel of HAdV NCBI RefSeq genomes (HAdV-A, HAdV-B1, HAdV-B2, HAdV-C, HAdV-D, HAdV-E, HAdV-F, HAdV-I, HAdV-2, HAdV-5, HAdV-7, HAdV-35, HAdV-54 and HAdV-F41).

The nine AAV2 near-complete genome contigs from the plasma sample were assembled and compared with sequences in GenBank using BLASTN (nucleotide database). Each of these AAV2 genomes had numerous close hits (exhibiting >95% similarity across 95% of the genome) with various existing AAV2 sequences; those most closely related were reported in a previous publication. All linear complete AAV2 genomes returned from BLAST against the GenBank nucleotide database with a query coverage of >75% were selected and combined with the AAV sequences de novo assembled here and aligned using MAFFT. The terminal ends of this alignment were trimmed off, and IQ-TREE2 was used to infer a phylogenetic tree. For the single terminal ends of this alignment were trimmed off, and IQ-TREE2 was used to infer a phylogenetic tree.

**Anti-AAV2 ELISA**
AAV2 pAAV-CAG-tddTomato viral preparation (codon diversified) was a gift from E. Boyden (Addgene viral preparation number 59462-AAV2; http://n2t.net/addgene:59462; RRID:Addgene_59462). AAV2 particles, obtained from Addgene (59462-AAV2) were diluted in PBS and used to coat a Immulon 2HB 96-well flat bottom plate (Immunochemistry Technologies) at a concentration of 1 × 10⁸ particles per well. The plates were incubated on an orbital shaker overnight at 4 °C. Plates were then blocked with PBS-T (PBS with 0.1% Tween-20) containing components of the CODEX staining kit (Akoya Biosciences, 44600, ThermoFisher). Antibodies were custom conjugated to a unique oligonucleotide barcode according to the manufacturer’s instructions using an antibody conjugation kit (7000009, Akoya Biosciences) and stored at 4 °C for at least 48 h before use. Conjugated antibodies were stored at 4 °C.

**Immunohistochemistry, ISH and special staining**
Formalin-fixed and paraffin-waxed embedded liver samples were cut at around 3 µm thickness and mounted on glass slides. A retinulin (1936) and Masson trichrome (1929) special staining method (Gordon and Sweets method (1936)) was performed. Antibodies used for immunohistochemistry are listed in Supplementary Table 6.

**Liver histopathology grading**
Liver scoring was performed as previously described.

**Quantification of immune cells**
After scanning of the whole slide, liver tissue was outlined and the number of positively stained cells (DAB signal for immunohistochemistry or Fast Red signal for ISH) was assessed using software-assisted image analysis (QuPath, v.0.3.2). For each marker, the cell detection algorithm was tuned, and data were plotted using GraphPad Prism (v.9.4.1).

**Spatial analysis (CODEX Phenocycler)**
Formalin fixed, paraffin-waxed-embedded liver samples (patient 228742A and patient 145808) were sectioned at 2–4 µm thickness on 22 × 22 mm glass coverslips (Akoya Biosciences, 7000005) coated in 0.1% poly-l-lysine (Sigma-Aldrich, P8920). Antigen retrieval was performed by pressure cooking with citrate buffer at pH 6. Carrier-free, pre-conjugated antibodies were purchased directly from Akoya Biosciences or purchased from other suppliers in preparation for custom conjugation. If carrier-free antibodies were not available, alternatives were purchased and purified using a Pierce antibody cleanup kit (44600, ThermoFisher). Antibodies were custom conjugated to a unique oligonucleotide barcode according to the manufacturer’s instructions using an antibody conjugation kit (7000009, Akoya Biosciences) and stored at 4 °C for at least 48 h before use. Conjugated antibodies were stored at 4 °C.

**Image acquisition**
Image acquisition was achieved using a Keyence BZ-X710 microscope equipped with 4 fluorescent channels (1 nuclear stain, 3 for antibody visualization). In a 96-well plate (Akoya Biosciences, 7000006), a maximum of three oligonucleotide reporters were used per well (cycle) (5 µl each) and added to between 235 µl and 245 µl reporter
stock solution that was made according to the manufacturer’s instructions. Plates were sealed with aluminium film (Akoya Biosciences, 70000007) and stored at 4 °C until use. Pictures were captured using QuPath (v.0.3.2)36.

Host genetics and HLA typing
High-resolution typing for all HLA loci (HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-DQA1, HLA-DQB1, HLA-DPA1 and HLA-DPB1) was performed using an AllType FASTplex NGS assay (One Lambda) run on an Illumina Mi-Seq platform. HLA typing was undertaken on 27 ISARIC participants who provided consent. One patient was omitted from analysis as they were a sibling of another case. HLA typing from 64 Scottish National Blood Transfusion Service apheresis platelet donors, self-identified as white British (n = 15) or white Scottish (n = 49) were used as control samples for comparison with patient HLA allele frequencies. Genotyping was performed using Illumina Global Screening Array v.3.0 + multi-disease beadchips (GSAMD-24v3-0-EA) and Infinium chemistry. This consists of three steps: (1) whole genome amplification; (2) fragmentation followed by hybridization; and (3) single-base extension and staining. Arrays were imaged on an Illumina iScan platform, and genotypes were automatically called using GenomeStudio Analysis software (v.2.0.3), GSAMD-24v3-0-EA_20034606_A1.bpm manifest and a cluster file provided by the manufacturer.

Given the small sample size, it was not possible to implement quality control processes using GenomeStudio and the manufacturer’s published recommendations. As genotyping was conducted using the same genotyping array used for the GenOMICC study, variants that passed quality control for the GenOMICC study were retained as previously described37. After further excluding variants with call rates of <0.95, a total of 478,692 variants were used for downstream analysis.

Kinship and population structure
To identify close relatives up to third degree, King 2.1 was used, which confirmed the presence of a pair of siblings with no further close relatives identified. Genotypes of 19 patients were combined with imputed genotypes of a subset of unrelated participants from the UK Biobank, which was obtained by removing one individual in each pair with estimated kinship larger than 0.0442. The resulting genotypes were filtered to exclude variants with a mean allele frequency of <5%, a genotype missingness rate of <1.5% and Hardy–Weinberg equilibrium of P < 10−8. Principal component analysis was conducted with gcta 1.955 in the set of unrelated individuals with pruned single nucleotide polymorphisms using a window of 1,000 markers, a step size of 50 markers and a r2 threshold of 0.01. Analyses were performed once including all UK Biobank participants and once including only UK Biobank participants who were born in Scotland (UK Biobank data-field 1647) and of Caucasian genetic ancestry (UK Biobank data-field 22006).

Statistics
Differences between cases and control groups were tested using Fisher’s exact test for categorical variables and Mann–Whitney (two tailed) for continuous variables using R studio (v.1.2.5033), R (v.4.1.2) and GraphPad (v.9.0.0). For coronavirus serology experiments, comparisons were carried out using one-way analysis of variance (ANOVA) and Tukey’s multiple comparison test, carried out in GraphPad (v.8.4.3). HLA analysis used the Bridging ImmunoGenomic Data-Analysis Workflow Gaps (BIGDAWG) R package to derive OR and corrected P values for individual HLA alleles.36 The Bonferroni-corrected P value significance threshold, adjusted for multiple comparisons (168 HLA alleles), was 0.3 × 10−4.

Figures
Figures were prepared using Microsoft Office Excel 2010, Microsoft Office PowerPoint 2010 and Adobe Illustrator 2022.

Code availability
Freely available bioinformatic and statistical software were used, see links in the Methods section.

Data availability
Datasets generated in the current study are appended as supplementary tables. Data, protocols and all documentation regarding this analysis may be made available to academic researchers after authorization from the independent data access and sharing committee. Clinical data and analysis scripts are available on request to the Independent Data Management and Access Committee at https://isaric-4c.net/sample-access. Restrictions apply to the availability of identifiable clinical data. Owing to the relatively small number of cases, de-aggregation of data is potentially disclosure, as is the patient-level line list data. Therefore, a formal data-sharing agreement is required for data access. The Independent Data and Material Access Committee considers requests as they arrive; most responses are made within 28 days. Use of clinical samples are also restricted under ethical approvals obtained for their use. Genome sequences are available at GenBank with accession numbers OP019741–OP019749 for AAV2 and OP019750 for HAdV-F41. Source data are provided with this paper.

Acknowledgements
We wish to acknowledge the contribution of the participating children and their parents who agreed to participate in the ISARIC CCP-UK and DIAMONDS studies, and to the research teams who recruited the patients; S. Bennett-Slater from NHS Greater Glasgow and Clyde for assisting with sample location and testing, the histopathology team, Veterinary Diagnostic, University of Glasgow, for excellent technical assistance; P. Murcia for providing resources and advice; P. Olmo for administrative assistance, and E. J. Kremer from the Institut de Génétique Moléculaire de Montpellier, Université de Montpellier and A. Baker, University of Edinburgh, for advice. The work was funded by Public Health Scotland, the National Institute for Health Research (NIHR; award CO-CIN-01) and the Medical Research Council (MRC; grants MR/X010252/1, MC_UU_120412, MC_UU_120182, MC_PC_18008, MC_PC_18025 and MC_PC_20040). DIAMONDS is funded by the European Union Horizon 2020 programme; grant 848196. M.P. acknowledges funding support from the Wellcome Trust (206369/Z/17/Z). J.K.B. acknowledges funding support from The Pandemic Institute, Liverpool and the NIHR grant 848196). M.G.S. acknowledges funding support from The Pandemic Institute, Liverpool and the NIHR (MRC; grants MR/UO13732/1, BB/P013759/1, and the Intensive Care Society of the United Kingdom. We acknowledge the support of Ballie Giffard and the Ballie Gifford Science Pandemic Hub at the University of Edinburgh. Parts of this research has been conducted using the UK Biobank Resource under project 788 and we would like to acknowledge the assistance.

References
of A. Tenesa in making this possible. Additional replication was also conducted using the UK Biobank Resource (Project 26041). This research was also funded by the National Institute for Health and Care Research (CO-CIN-01) and jointly by NIHR and UK Research and Innovation (CV220-169, MC_PC_19059). The views expressed in this article are those of the author(s) and not necessarily those of UKRI, the NIHR, or the Department of Health and Social Care. We also acknowledge the support of NHS Research Scotland (NRS) Greater Glasgow and Clyde Biorepository team. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author contributions

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-023-05948-2.

Correspondence and requests for materials should be addressed to Emma C. Thomson.

Peer review information Nature thanks Frank Tacke, Leif Sander and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at http://www.nature.com/reprints.
Extended Data Fig. 1 | AAV2, HAdV and human herpesvirus detection by target enrichment sequencing in cases and controls. Read counts per million are plotted for a) HAdV; b) AAV2; c) HHV6B; d) HSV1; e) HSV2; f) VZV; g) HHV6A; h) HHV7; i) HHV8; and j) CMV in cases, Group 1 healthy controls and Group 2 controls (HAdV positive children with normal liver function). Statistical significance was estimated using a Mann-Whitney test (two-sided).
Extended Data Fig. 2 | Phylogenetic and sequence analysis of AAV2 genomes. a) Maximum likelihood phylogeny of AAV2 from hepatitis cases CVR1-9. The nine AAV2 genome sequences generated from the plasma samples via target enrichment (highlighted in green) were aligned with a range of the closest AAV GenBank sequences 39. AAV2 reference sequences are denoted by accession number, country and year of sampling. b), Phylogeny of HAdV41 genome from case 5. The HAdV41 genome sequence from the faecal sample of patient 5 (red) was combined with complete genomes of HAdV41 from GenBank. Bootstrap values >70 are indicated. HAdV41 reference sequences are denoted by accession number, country and year of sampling. c), Key mutations and hierarchical clustering of AAV2 genomes. Mutations in published AAV2 sequences are highlighted in (blue) and case sequences (green). d) Mutations over-represented in hepatitis cases versus controls. Mutations in VP1-3, Rep78 and 52 and AAP are highlighted by % representation in case sequences (green) and published sequences (blue).
Extended Data Fig. 3 | Reactivity of sera from paediatric hepatitis cases against human seasonal coronaviruses and SARS-CoV-2. Sera from the paediatric hepatitis cases were screened for reactivity against spike proteins from a) seasonal coronaviruses 229E, OC43, NL63 and HKU1, and b) SARS-CoV-2 nucleocapsid (N), spike (S), and N-terminal domain (NTD) and receptor binding domain (RBD) of S by electrochemiluminescence (MSD-ECL). Reactivity of the 23 samples (Hepatitis) was compared with 16 sera from contemporaneous control samples from children (Group 4 Controls), and three groups of sera from adults of known SARS-CoV-2 status; Negatives (never tested positive for SARS-CoV-2; n = 30), Vaccinated two doses (n = 28) and Infected (n = 39).
**Extended Data Fig. 4 | Principal component analysis (PCA) plots.** PCA plots showing the first four genome-wide principal components to confirm genetic ancestry matching. 

a) Genomic PCA using full United Kingdom Biobank cohort as background population (grey), showing the subgroup of unrelated United Kingdom Biobank participants who were born in Scotland and of Caucasian ancestry (blue) and the hepatitis cases reported here (red).

b) plots showing only the subgroup born in Scotland and of Caucasian ancestry.
### Extended Data Table 1 | Modified hepatic activity index scores

<table>
<thead>
<tr>
<th></th>
<th>CVR1</th>
<th>CVR3</th>
<th>CVR4</th>
<th>CVR9</th>
<th>CVR35</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modified Ishak score</td>
<td>6/18</td>
<td>10/18</td>
<td>10/18</td>
<td>8/18</td>
<td>11/18</td>
</tr>
<tr>
<td>Portal inflammation</td>
<td>2/4</td>
<td>4/4</td>
<td>4/4</td>
<td>2/4</td>
<td>2/4</td>
</tr>
<tr>
<td>Confluent necrosis</td>
<td>0/6</td>
<td>0/6</td>
<td>0/6</td>
<td>0/6</td>
<td>1/6</td>
</tr>
<tr>
<td>Lobular inflammation</td>
<td>2/4</td>
<td>2/4</td>
<td>2/4</td>
<td>2/4</td>
<td>4/4</td>
</tr>
</tbody>
</table>
### Extended Data Table 2 | Characteristics of cases and controls

#### a) used in metagenomic and target enrichment analysis

<table>
<thead>
<tr>
<th></th>
<th>Cases - metagenomic &amp; TE sequencing analysis (n=9)</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group 1 DIAMONDS Healthy (n=13)</td>
<td>P value*</td>
</tr>
<tr>
<td>Sex - male</td>
<td>4 (44.4)</td>
<td>10 (76.9)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>3.9 (3.4-5.1)</td>
<td>4.1 (3.6-4.8)</td>
</tr>
<tr>
<td>Recruitment period</td>
<td>14 March – 20 April 2022</td>
<td>6 Nov 2020 - 6 Jul 2021</td>
</tr>
</tbody>
</table>

#### b) used in PCR analysis

<table>
<thead>
<tr>
<th></th>
<th>Cases – PCR (n=32)</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group 1 DIAMONDS Healthy (n=13)</td>
<td>P value*</td>
</tr>
<tr>
<td>Sex - male</td>
<td>11 (34.4)</td>
<td>10 (76.9)</td>
</tr>
<tr>
<td>Age (years) †</td>
<td>4.1 (2.7-5.5)</td>
<td>4.1 (3.6-4.8)</td>
</tr>
<tr>
<td>Recruitment period</td>
<td>14 March – 20 August 2022</td>
<td>6 Nov 2020 - 6 Jul 2021</td>
</tr>
</tbody>
</table>

*Fisher’s Exact or chi-squared test for categorical and Mann-Whitney (two-sided) test for continuous variables.

†Age and sex of Group 4 controls unavailable.
Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

<table>
<thead>
<tr>
<th>Item</th>
<th>Confirmation</th>
</tr>
</thead>
<tbody>
<tr>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>☑ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement</td>
<td>Yes</td>
</tr>
<tr>
<td>☑ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly</td>
<td>Yes</td>
</tr>
<tr>
<td>☑ The statistical test(s) used AND whether they are one- or two-sided</td>
<td>Yes</td>
</tr>
<tr>
<td>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</td>
<td></td>
</tr>
<tr>
<td>☑ A description of all covariates tested</td>
<td>Yes</td>
</tr>
<tr>
<td>☑ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons</td>
<td>Yes</td>
</tr>
<tr>
<td>☑ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)</td>
<td>Yes</td>
</tr>
<tr>
<td>☑ For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted Give P values as exact values whenever suitable.</td>
<td>Yes</td>
</tr>
<tr>
<td>☑ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings</td>
<td>Yes</td>
</tr>
<tr>
<td>☑ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes</td>
<td>Yes</td>
</tr>
<tr>
<td>☑ Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection No commercial code was used for data collection in this study

Data analysis

**HLA ANALYSIS**

The Bridging ImmunoGenomic Data-Analysis Workflow Gaps (BIGDAWG) R package to derive OR and corrected p values for individual HLA alleles. 30 Bonferroni corrected p value significance threshold, adjusted for multiple comparisons (168 HLA alleles), was p < 3.0x10^-4.

**BIOMINFORMATICS ANALYSIS**

Reads for each sample were first quality checked, Illumina adapters were trimmed using Trim Galore version 0.6.6 (https://github.com/FelixKrueger/TrimGalore), and reads were then mapped to the human genome using BWA-MEM version 0.7.17 (https://github.com/lh3/bwa). Only reads that did not map to the human genome were used for metagenomic analyses. Non-human reads were then de novo assembled using MetaSPAdes version 3.15.5 (https://github.com/aiblab/spades) to generate contigs for each sample. Contigs were then compared against a protein database of all NCBI RefSeq organisms [including virus, bacteria, eukaryotes] with BLASTX using DIAMOND version 2.0.15 (https://github.com/bbuchfink/diamond). In addition, non-human reads for each sample were aligned to a small panel of HAdV NCBI RefSeq genomes (HAdV-A, B1, B2, C, D, E, F, 1, 2, 5, 7, 35, 54 as well as HAdV-F41).

**STATISTICAL ANALYSIS**

Differences between cases and control groups were tested using Fisher’s Exact Test for categorical variables and Mann-Whitney [two-tailed] for continuous variables. Spearman’s rank correlation coefficients were calculated for the relationships between the trajectories of viral load and ALT and bilirubin. We used R studio version 1.2.5033, R version 4.1.2 and GraphPad version 9.0.0 for most statistical analyses. For coronavirus serology experiments, comparisons were carried out with one way ANOVA and Tukey’s Multiple Comparison test, carried out in
Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Datasets generated in the current study are appended as Source Data, Extended Data Tables and Supplementary Tables. Data, protocols, and all documentation around this analysis may be made available to academic researchers after authorisation from the independent data access and sharing committee. Clinical data and analysis scripts are available on request to the Independent Data Management and Access Committee at https://Isaric4c.net/samples_access. Restrictions apply to the availability of identifiable clinical data. Due to the relatively small number of cases, de-aggregation of data is potentially disclosive, as is the patient level list data. Therefore, a formal data sharing agreement is required for data access. The Independent Data and Material Access Committee considers requests as they arrive; most responses are made within 28 days. Use of clinical samples are also restricted under ethical approvals obtained for their use. Genome sequences are available in GenBank with accession numbers for AAV2: OP019741-OP019749 and for HAdV-F41: OP019750.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender

For the case control study, recruited patients were female (n=20) and male (n=12).
Group 1 healthy control subjects were restricted to 13 children recruited in the UK between January 2020 and April 2022 and were age-matched but not sex-matched due to availability of samples (10 male, 3 female; age range 3-5 years).
Group 2 subjects were children (8 male, 4 female; age range 1-4 years) with PCR-confirmed HAdV infection and normal transaminases.
Group 3, 33 children (18 male, 15 female; age range 2-16 years) with raised transaminases who were HAdV PCR negative.
Group 4 included 16 residual samples from children from the NHS GG&C biorepository aged <18 years. Further information was not available under ethical protocols for the use of residual biorepository samples.
For the HLA analysis a further 3 cases were recruited to the ISARIC CCP-UK cohort and had HLA typing carried out but further clinical samples and additional clinical data were not available.

Population characteristics

The median age of affected patients was 4.1 years [IQR: 2.7 to 5.5 years] (Table 1). Twenty of the 32 (62%) children were female, and all were of white ethnicity. Eighteen (56%) of the children reported a subacute history 2-12 weeks prior to acute hepatitis, characterised by an initial gastroenteritis-like illness followed by intermittent vomiting, abdominal pain and fatigue.
Most of the affected children (23/32) had no other medical conditions: one child previously received a liver transplant, none of the other cases were immunocompromised and none had received COVID-19 vaccination. All routine blood tests for viral hepatitis, including hepatitis A, B, C, E, acute Epstein-Barr virus (EBV), cytomegalovirus (CMV), human herpes virus (HHV) 6/7 and herpes simplex virus (HSV) were negative.

Recruitment

To investigate the aetiology of the acute hepatitis cases, we recruited 32 of the earliest affected children who presented to hospital between 14 March and 4 April 2022 into the International Severe Acute Respiratory and Emerging Infections Consortium (ISARIC) WHO Clinical Characterisation Protocol UK (CCP-UK) [ISRCTN 66726260]7. All cases who fulfilled the case definition and were willing to participate were recruited. For the HLA analysis a further 3 cases were recruited to the ISARIC CCP-UK cohort and had HLA typing carried out but further clinical samples and additional clinical data were not available.
Control samples (Groups 1, 2 and 3) were obtained from the Diagnosis and Management of Febrile Illness using RNA Personalised Molecular Signature Diagnosis study cohort (DIAMONDS; https://www.diamonds2020.eu). This study recruited children presenting with suspected infection or inflammation. Patients were recruited with the informed written consent of parents or guardians.

Ethics oversight

32 affected children, who presented to hospital between 14 March and 20 August 2022 and who met the PHS case definition into the International Severe Acute Respiratory and Emerging Infections Consortium (ISARIC) WHO Clinical Characterisation Protocol UK (CCP-UK) [ISRCTN 66726260]. 7 Ethical approval was given by the South Central–Oxford C Research Ethics Committee in England (13/SC/0149), the Scotland A Research Ethics Committee (20/SS/0028), and the WHO Ethics Review Committee (RPC571 and RPC572).
Control samples (Groups 1-3) were obtained from the Diagnosis and Management of Febrile Illness using RNA Personalised Molecular Signature Diagnosis study cohort (DIAMONDS; https://www.diamonds2020.eu). Patients were recruited with the written informed consent of parents or guardians.
Contemporary Scottish surplus plasma and liver biopsy control samples (Control Group 4) from the Diagnostic Pathology/Blood Sciences archive were obtained with NHS GG&C Biorepository approval (application #717; REC 22/NW/0020). These samples were used without consent following HRA legislation on consent exemption.
Genetic (HLA) control data was obtained using the UK Biobank Resource (project 788; 21/NW/0157). Participants in the UK Biobank have been recruited with written informed consent.

Note that full information on the approval of the study protocol must also be provided in the manuscript.
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☑ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-list.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

All available cases were selected. All available healthy control samples that could be age-matched to cases were obtained from the DIAMONDS cohort (we planned for up to a 1:3:1 ratio of controls:cases). We selected all available control subjects in group 2 (HAdV positive with normal LFTs) from the DIAMONDS cohort. These were all children but were not age matched. We also selected all available control subject samples in group 3 (hepatitis of alternative aetiology) from the DIAMONDS cohort. We used all available residual samples from children from the same time period as cases in group 4.

Data exclusions

We excluded any cases that did not meet the PHS definition for non-A-E paediatric hepatitis on the basis of age (over 10 years of age or with an alternative diagnosis or from whom clinical data was not available).

We excluded 5 plasma samples in the case control study from the NGS analysis of herpesviruses because during nucleic extraction in the relevant clinical laboratory, murine cytomegalovirus (CMV) had been used as an extraction control. This was not used for other sample extractions in the case control study. Clinical specimens taken from cases (throat, rectal swab, faeces and liver samples) also had murine CMV added to the samples and were also excluded from the NGS herpes read count analysis. However, all samples were tested for human herpesviruses by PCR.

Replication

PCR experiments were carried out in triplicate, other than GAPDH PCR which was carried out in duplicate or triplicate. Results were highly concordant. There were four AAV2 Ct values that were borderline (transversing the limit of detection). These were considered negative as weakly positive results were not reproducible and read counts for all samples were negative by metagenomic and target enrichment NGS. Next generation sequencing experiments were repeated on separate runs using different methods (metagenomic sequencing and then semi-agnostic target enrichment sequencing). Results were also confirmed by PCR for AAV2, HAdV and HHV6.

Randomization

As described above, all available cases were selected. All available healthy control samples that could be age-matched to cases were obtained from the DIAMONDS cohort (we planned for up to a 1:3:1 ratio of controls:cases). We selected all available control subjects in group 2 (HAdV positive with normal LFTs) from the DIAMONDS cohort. These were all children but were not age matched. We also selected all available control subject samples in group 3 (hepatitis of alternative aetiology) from the DIAMONDS cohort.

Blinding

The first sequencing run was of samples from 5 cases carried out urgently at the request of Public Health Scotland and investigators were not blinded to these as ethical permissions were not in place for the use of control samples. Subsequent runs were carried out when control samples were available and included 4 further cases. These were analysed with blinding of case/control status and then samples were compared to look for viruses present in cases and controls.

PCR, serology and histology experiments were carried out with blinding in place.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

<table>
<thead>
<tr>
<th>Materials &amp; experimental systems</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>n/a</td>
<td>Involved in the study</td>
</tr>
<tr>
<td>☑ Antibodies</td>
<td></td>
</tr>
<tr>
<td>☑ Eukaryotic cell lines</td>
<td></td>
</tr>
<tr>
<td>☑ Palaeontology and archaeology</td>
<td></td>
</tr>
<tr>
<td>☑ Animals and other organisms</td>
<td></td>
</tr>
<tr>
<td>☑ Clinical data</td>
<td></td>
</tr>
<tr>
<td>☑ Dual use research of concern</td>
<td></td>
</tr>
</tbody>
</table>

Antibodies

Antibodies used

For AAV2 ELISA, bound human antibody was detected with either anti-human IgM or anti-human IgG (Merck, UK cat no. A9794 and A1543, respectively). Antibodies for ILIC are listed below.

AAntigen Dilution Clone Product code, company Antigen retrieval Detection system
Validation

For IHC, C4d was validated on kidney with acute rejection versus normal kidney tissue. Validation policies and procedures were carried out in accordance with ISO accreditation ISO 15189. CD3, CD4, CD8 and CD20 were validated with normal tonsill control tissue and positive case tissue (lymphoma) from a minimum of 3 cases. Procedures were carried out in accordance with ISO 15189. Test subject tissue for these cases was compared with normal controls and negative controls.

For the CODEX analysis, we carried out validation of antibodies by 1) replacing the primary antibody with isotype serum and 2) checking in each section that stained cells had the morphology of the cell to be stained, e.g., MHCII macrophage-like cells close to the sinus using concentrations, as recommended by the manufacturer.