Accepted Manuscript

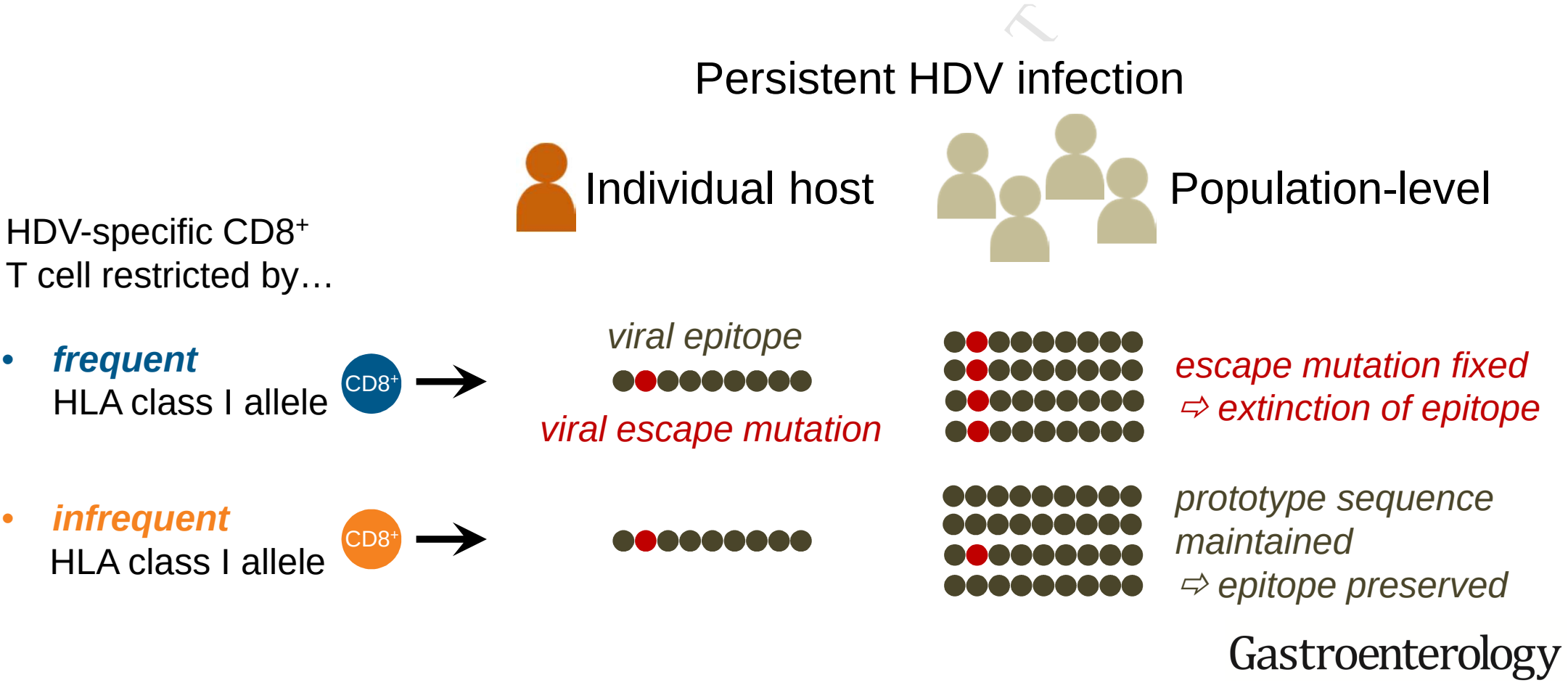
Mutations in Hepatitis D Virus Allow it to Escape Detection by CD8+ T Cells and Evolve at the Population Level

Hadi Karimzadeh, Muthamia M. Kiraithe, Valerie Oberhardt, Elahe Salimi Alizei, Jan Bockmann, Julian Schulze zur Wiesch, Bettina Budeus, Daniel Hoffmann, Heiner Wedemeyer, Markus Cornberg, Adalbert Krawczyk, Jassin Rashidi-Alavijeh, Francisco Rodríguez-Frías, Rosario Casillas, Maria Buti, Antonina Smedile, Seyed Moayed Alavian, Andreas Heinold, Florian Emmerich, Marcus Panning, Emma Gostick, David A. Price, Jörg Timm, Maike Hofmann, Bijan Raziorrouh, Robert Thimme, Ulrike Protzer, Michael Roggendorf, Christoph Neumann-Haefelin

PII: S0016-5085(19)30366-X
DOI: https://doi.org/10.1053/j.gastro.2019.02.003
Reference: YGAST 62453

To appear in: *Gastroenterology*
Accepted Date: 3 February 2019

Please cite this article as: Karimzadeh H, Kiraithe MM, Oberhardt V, Alizei ES, Bockmann J, zur Wiesch JS, Budeus B, Hoffmann D, Wedemeyer H, Cornberg M, Krawczyk A, Rashidi-Alavijeh J, Rodríguez-Frías F, Casillas R, Buti M, Smedile A, Alavian SM, Heinold A, Emmerich F, Panning M, Gostick E, Price DA, Timm J, Hofmann M, Raziorrouh B, Thimme R, Protzer U, Roggendorf M, Neumann-Haefelin C, Mutations in Hepatitis D Virus Allow it to Escape Detection by CD8+ T Cells and Evolve at the Population Level, *Gastroenterology* (2019), doi: https://doi.org/10.1053/j.gastro.2019.02.003.



Gastroenterology

Mutations in Hepatitis D Virus Allow it to Escape Detection by CD8+ T Cells and Evolve at the Population Level

Short title: Viral escape from HDV-specific $CD8^+$ T cells

Hadi Karimzadeh[1,2]*, Muthamia M. Kiraithe[3]*, Valerie Oberhardt[3,4]*, Elahe Salimi Alizei[3,5], Jan Bockmann[6,7], Julian Schulze zur Wiesch[6,7], Bettina Budeus[8], Daniel Hoffmann[8], Heiner Wedemeyer[7,9,10], Markus Cornberg[7,9], Adalbert Krawczyk[2,11], Jassin Rashidi-Alavijeh[10], Francisco Rodríguez-Frías[12], Rosario Casillas[12], Maria Buti[12], Antonina Smedile[13], Seyed Moayed Alavian[14], Andreas Heinold[15], Florian Emmerich[16], Marcus Panning[17], Emma Gostick[18], David A. Price[18], Jörg Timm[19], Maike Hofmann[3], Bijan Raziorrouh[20], Robert Thimme[3], Ulrike Protzer[1,7], Michael Roggendorf[1,2,7#], Christoph Neumann-Haefelin[3#]

*equally contributing primary authors

#equally contributing senior authors

1 Institute of Virology, Technical University of Munich/Helmholtz Zentrum München, Munich, Germany

2 Institute of Virology, University Hospital of Essen, University of Duisburg-Essen, Essen, Germany

3 Department of Medicine II, University Hospital Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany

4 Faculty of Biology, University of Freiburg, Freiburg, Germany

5 Faculty of Chemistry and Pharmacy, University of Freiburg, Freiburg, Germany

6 Department of Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[7] German Center for Infection Research (DZIF), Sites Hamburg-Lübeck-Borstel-Riems, Hannover-Braunschweig and Munich, Germany

[8] Department of Bioinformatics, University of Duisburg-Essen, Essen, Germany

[9] Department of Gastroenterology, Hepatology and Endocrinology, Hannover Medical School, Hannover, Germany

[10] Department of Gastroenterology and Hepatology, University Hospital of Essen, University of Duisburg-Essen, Essen, Germany

[11] Department of Infectious Diseases, University Hospital of Essen, University of Duisburg-Essen, Essen, Germany

[12] CIBERehd and Departments of Biochemistry/Microbiology and Hepatology, Vall d'Hebron Hospital, University Autònoma de Barcelona (UAB), Barcelona, Spain

[13] Department of Medical Sciences, University of Turin, Turin, Italy

[14] Baqiyatallah Research Center for Gastroenterology and Liver Diseases, Baqiyatallah University of Medical Sciences, Tehran, Iran

[15] Institute of Transfusion Medicine, University Hospital of Essen, University of Duisburg-Essen, Essen, Germany

[16] Institute for Transfusion Medicine and Gene Therapy, University Hospital Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany

[17] Institute of Virology, University Hospital Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany

[18] Division of Infection and Immunity, Cardiff University School of Medicine, Cardiff, United Kingdom

[19] Institute of Virology, Heinrich-Heine-University, University Hospital, Duesseldorf, Germany

[20] Department of Internal Medicine II, University Hospital Munich-Grosshadern, Munich, Germany

Grant Support: This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG) via grants to HK and BR (SFB-TR179/TP3), JSzW (SFB841, SFB1328), DH (TRR60/B1), JT (TI323/4-1), MH and RT (SFB-TRR-179/TP1), UP (SFB-TR179/TP14 and TRR36/B13), and CNH (SFB-TRR-179/TP2). Additional funding was provided by the Medical Faculty of the University of Duisburg-Essen / Technical University of Munich and by a registry grant from the European Association for the Study of the Liver (EASL). HK was supported by a grant from the German Academic Exchange Service (DAAD). DAP was supported by a Wellcome Trust Senior Investigator Award (100326Z/12/Z). JSzW, HW, MC, UP, and MR were supported by the German Center for Infection Research (DZIF).

Abbreviations: ALT, alanine aminotransferase; FCS, fetal calf serum; HCC, hepatocellular carcinoma; HDAg, hepatitis delta antigen; HDV, hepatitis D virus; HIV, human immunodeficiency virus; HLA, human leukocyte antigen; IFN, interferon; IL-2, interleukin 2; KIR, killer cell immunoglobulin-like receptor; KLRG1, killer cell lectin-like receptor G1; L-HDAg, large variant of the hepatitis delta antigen; PBMC, peripheral blood mononuclear cell; PCR, polymerase chain reaction; PD-1, programmed cell death protein 1; PE, phycoerythrin; PFA, paraformaldehyde; RPMI, Roswell Park Memorial Institute; S-HDAg, small variant of the hepatitis delta antigen; TCF1, T-cell factor 1.

Correspondence: Michael Roggendorf, MD, Institute of Virology, Technical University of Munich, Schneckenburger Strasse 8, 81675 Munich, Germany (Tel: +49/8941407360; Fax: +49/8941407444; E-mail: michael.roggendorf@tum.de) and Christoph Neumann-Haefelin, MD, Department of Medicine II, University Hospital Freiburg, Hugstetter Strasse 55, 79106 Freiburg,

Germany (Tel: +49/76127032800; Fax: +49/76127032800; E-mail: christoph.neumann-haefelin@uniklinik-freiburg.de).

Disclosures: The authors have no conflicts of interest.

Author Contributions: HK, MMK, MR, and CNH designed the study and interpreted the data; HK performed viral sequence and footprint analyses; MMK, VO, and ESA performed immunological experiments; JB, JSzW, HW, MC, AK, JRA, FRF, RC, MB, AS, SMA, MP, BR, and CNH enrolled patients and performed clinical evaluations; BB and DH assisted with statistical analyses; AH and FE performed HLA class I genotyping; EG and DAP generated HLA class I tetramers; JT, MH, RT, and UP helped design and interpret the experiments; HK, MMK, VO, DAP, MR, and CNH wrote the manuscript; all authors read and approved the final manuscript.

Abstract

Background & Aims: Hepatitis D virus (HDV) super-infection in patients with hepatitis B virus (HBV) is associated with rapid progression to liver cirrhosis and hepatocellular carcinoma. Treatment options are limited, and no vaccine is available. Although HDV-specific CD8+ T cells are thought to control the virus, little is known about which HDV epitopes are targeted by virus-specific CD8+ T cells or why these cells ultimately fail to control the infection. We aimed to define how HDV escapes the CD8+ T cell-mediated response.

Methods: We collected plasma and DNA samples from 104 patients with chronic HDV and HBV infection at medical centers in Europe and Asia, sequenced HDV, typed HLA class I alleles from patients, and searched for polymorphisms in HDV RNA associated with specific HLA class I alleles. We predicted epitopes in HDV that would be recognized by CD8+ T cells and corresponded with the identified virus polymorphisms in patients with resolved (n = 12) or chronic (n = 13) HDV infection.

Results: We identified 21 polymorphisms in HDV that were significantly associated with specific HLA class I alleles (P<.005). Five of these polymorphisms were found to correspond to epitopes in HDV that are recognized by CD8+ T cells; we confirmed that CD8+ T cells in culture targeted these HDV epitopes. HDV variant peptides were only partially cross-recognized by CD8+ T cells isolated from patients, indicating that the virus had escaped detection by these cells. These newly identified HDV epitopes were restricted by relatively infrequent HLA class I alleles, and bound most frequently to HLA-B. In contrast, frequent HLA class I alleles were not associated with HDV sequence polymorphisms.

Conclusions: We analyzed sequences of HDV RNA and HLA class I alleles that present epitope peptides to CD8+ T cells in patients with persistent HDV infection. We identified polymorphisms in the HDV proteome that associate with HLA class I alleles. Some variant peptides in epitopes from HDV were only partially recognized by CD8+ T cells isolated from patients—these could be mutations that allow HDV to escape the immune response, resulting in persistent infection. HDV escape from the immune response was associated with uncommon HLA class I alleles, indicating that HDV evolves, at the population level, to evade recognition by common HLA class I alleles.

KEY WORDS: cytotoxic T cell; MHC class I; TCR; antigen presentation

Introduction

Hepatitis D virus (HDV) has infected approximately 10% of hepatitis B virus (HBV)-seropositive individuals, affecting 15 to 20 million people worldwide.[1] Patients with simultaneous HDV/HBV infection often experience severe acute hepatitis, with an enhanced risk of fulminant disease, but frequently clear both viruses. In contrast, HBV-seropositive patients who become super-infected with HDV generally develop chronic HDV/HBV infection. These patients have a high risk of rapid progression to liver cirrhosis and hepatocellular carcinoma (HCC). At present, the sole treatment option for chronic HDV/HBV infection is pegylated interferon (IFN)-α, which is usually administered for 48 weeks. However, only around 30% of patients respond to treatment, and many subsequently develop viral relapse.[2]

Virus-specific $CD8^+$ T cells are thought to play a key role in the outcome of HDV/HBV infection. This contention has been supported by vaccination studies in mice and woodchucks.[3] However, little is known about naturally occurring HDV-specific $CD8^+$ T-cell responses, despite the urgent need for novel prophylactic and therapeutic interventions. The HDV genome encodes a single viral protein, hepatitis delta antigen (HDAg), translated as large and small variants (L-HDAg and S-HDAg, respectively). The only difference between these two proteins is a 19 amino acid extension at the C-terminus of L-HDAg. An initial report described two HDV-specific $CD8^+$ T-cell epitopes in human leukcoyte antigen (HLA)-A*02 transgenic mice and patients with resolved HDV infection, but these findings were not confirmed in patients with persistent HDV infection.[4] A subsequent report suggested that immune exhaustion may impair the efficacy of HDV-specific $CD8^+$ T cells *in vivo.*[5] In a recent study, we identified two HLA-B*27-restricted HDV-specific $CD8^+$ T-cell epitopes, both of which were subject to viral escape by mutation.[6] Of note, HLA-B*27 is associated with spontaneous clearance of hepatitis C virus (HCV) and elite control of human immunodeficiency virus (HIV), and HLA-B*27-restricted $CD8^+$ T-cell responses drive viral evolution in patients infected with HCV or HIV.[7] It nonetheless remains unclear if viral

escape is a generalizable phenomenon that underlies immune failure in patients infected with HDV.

In the present study, we analyzed HLA class I-associated viral sequence polymorphisms across the entire HDV genome as surrogate markers of $CD8^+$ T-cell-driven viral escape. These 'HLA footprints' were used to identify $CD8^+$ T-cell epitopes, confirm predicted escape mutations, and explore features of the host-virus interactome, providing mechanistic insights into adaptive immunity against HDV.

Materials and Methods

***Patients and samples*.** Patients with chronic HDV/HBV infection (n = 104) were recruited from eight medical centers located in Germany (Bonn, Düsseldorf, Essen, Hannover, and Munich), Spain, Italy, and Iran.[6] Viral sequences and HLA class I genotypes were determined according to standard protocols.[6] $CD8^+$ T-cell assays were performed using additional samples obtained from 12 patients with resolved HDV infection and 13 patients with chronic HDV/HBV infection, recruited from Freiburg, Hamburg, and Munich (Germany). All patients were infected with HDV genotype 1. Patient characteristics are summarized in supplementary Tables S1 and S2. Ethical approval was granted by the Ethik-Kommission der Albert-Ludwigs-Universität Freiburg (#369/15). Written informed consent was obtained in all cases according to federal guidelines and the Declaration of Helsinki. Venous blood samples (50 mL per draw) were collected in ethylene diamine tetraacetic acid (EDTA)-anticoagulated tubes. Peripheral blood mononuclear cells (PBMCs) were isolated using lymphocyte separation medium density gradients (PAA Laboratories, Austria) and resuspended in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10% fetal calf serum (FCS), 1% penicillin/streptomycin, and 1.5% HEPES buffer 1M (complete medium; all additives from Thermo Fisher Scientific, Germany).

***Peptides, antibodies, and tetramers*.** Peptides were synthesized with a free amine NH_2 terminus and a free acid COOH terminus using standard Fmoc chemistry (Genaxxon Bioscience, Germany). Anti-CD8 (RPA-T8, 1:100; SK1, 1:300), anti-CD38 (Hb7, 1:200), anti-CD45RA (HI100, 1:50), anti-IFN-γ (4S.B3, 1:50), and anti-PD-1 (EH12.1, 1:33) were purchased from BD Biosciences, Germany. Anti-Bcl-2 (Bcl-2/100, 1:200), anti-CCR7 (G043H7, 1:33), anti-CD8 (RPA-T8, 1:400), and anti-CD127 (A019D5, 1:33) were purchased from BioLegend, UK. Anti-CD14 (61D3, 1:100), anti-CD19 (HIB19, 1:100), anti-Eomes (WD1928, 1:50), anti-KLRG1 (13F12F2, 1:33), and anti-T-bet (4B10, 1:200) were purchased from Thermo Fisher Scientific,

Germany. Anti-TCF1 (C63D9, 1:100) was purchased from Cell Signaling Technology, Germany. Tetrameric complexes of HLA-B*15:01/L-HDAg$_{170–179}$ SMQGVPESPF were generated as described previously.[8]

***Analysis of HDV sequences and HLA class I-associated viral polymorphisms*.** Total RNA was extracted from serum samples and reverse transcribed into cDNA using Moloney Murine Leukemia Virus Reverse Transcriptase (Promega, Madison, WI) and the HDV-specific primer 771R (5'-CGGTCCCCTCGGAATGTTG-3'). The L-HDAg-encoding region was amplified from cDNA using Pfu DNA Polymerase (Promega, Madison, WI) in a two-step nested polymerase chain reaction (PCR) incorporating the HDV-specific primers 891-F (5'-AGGTCGGACCGCGAGGAGGT-3'), 339-R (5'-GCTGAAGGGGTCCTCTGGAGGTG-3'), 912-F (5'-GAGATGCCATGCCGACCCGAAGAG-3'), and 1674-R (5'-AGAAAAGAGTAAGAGYACTGAGG-3'). The following thermal profile was used for all PCRs: (i) 94 ℃ for 10 minutes; (ii) 94 ℃ for 30 seconds, 54 ℃ for 45 seconds, and 72 ℃ for 90 seconds (35 cycles); and (iii) 72 ℃ for 7 minutes. PCR products were purified using a QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) and sequenced on an ABI 3730xl DNA Analyzer using the internal primers 912-F and 1674-R. All sequences were submitted to GenBank (accession numbers MF175257–MF175360). HDV sequences were aligned using ClustalX2 software.[9] HLA class I associations with amino acid residues at each alignment position were tested using the R package SeqFeatR (https://seqfeatr.zmb.uni-due.de).[10] Fisher's exact test was used to determine significant associations in a 2 x 2 contingency table format with counts for observed combinations of amino acids and HLA class I allotypes at each alignment position. Further analyses were based on a cut-off P value of < 0.005.

***Prediction of HDV-specific CD8$^+$ T-cell epitopes*.** Viral amino acid sequences 15 residues N-terminal and 15 residues C-terminal of the identified HLA class I-associated viral sequence

polymorphisms were analyzed for the corresponding binding motifs using four online prediction tools: ANN 3.4 and netMHCpan 2.8 on the Immune Epitope Database webpage,[11, 12] SYFPEITHI,[13] and BIMAS.[14] 8mer, 9mer, and 10mer peptides were tested if available via the respective prediction tool. An IC_{50} of ≤ 1,000 nM, a SYFPEITHI score of ≥ 20, and a BIMAS score of ≥ 20 were used as cut-offs. Candidate epitopes were ranked against epitopes predicted across the entire sequence of L-HDAg.

***Peptide-specific CD8$^+$ T-cell lines*.** PBMCs were activated with peptides as described previously.[15] Briefly, 4 x 10^6 PBMCs were stimulated once with 10 µg/mL peptide and 0.5 µg/mL anti-CD28 (BD Biosciences, Germany) and fed every 3 days with complete medium containing 20 U/mL recombinant IL-2 (Miltenyi Biotec, Germany). Peptide-specific CD8$^+$ T-cell lines were used for experimental purposes after 14 days.

***Intracellular IFN-γ staining*.** Procedures were carried out as described previously.[16] Briefly, expanded CD8$^+$ T cells or peptide-specific CD8$^+$ T-cell lines (0.2 x 10^6 cells per well in a 96-well plate) were stimulated with peptides (10 µg/mL) in the presence of 50 U/mL recombinant IL-2 and 1 µL/mL brefeldin A (BD Biosciences, Germany). After 5 hours, cells were stained with 7-aminoactinomycin D and anti-CD8, fixed/permeabilized with Cytofix/Cytoperm, and stained with anti-IFN-γ (all reagents from BD Biosciences, Germany). Stained cells were fixed in phosphate buffered saline (PBS) containing 2% paraformaldehyde (PFA). Data were acquired using an BD FACSCanto II flow cytometer (BD Biosciences, Germany) and analyzed with FlowJo software version 10 (FlowJo LLC, Ashland, OR).

***HLA class I tetramer-based analysis and cell enrichment*.** Tetramer staining procedures were carried out as described previously.[17] Briefly, 1 x 10^6 PBMCs per well were incubated in a 96-well plate with the relevant HLA class I tetramer for 15 minutes at 37 ℃. Cells were then washed three times with PBS containing 1% FCS and stained with the indicated surface and/or

intracellular antibodies. Dead cells were excluded from the analysis using the fixable viability dye eFluor780 (1:5,000, eBioscience, Germany). Cytoplasmic and nuclear molecules were revealed using a FoxP3/Transcription Factor Staining Buffer Set (eBioscience, Germany). Stained cells were fixed in PBS containing 2% PFA. Tetramer-based enrichment was performed as described by Alanio *et al.*[18] Briefly, 10–15 x 10^6 PBMCs were labeled for 30 minutes at room temperature with the HLA-B*15:01/L-$HDAg_{170–179}$ SMQGVPESPF tetramer coupled to phycoerythrin (PE) and enriched using anti-PE beads with MACS technology (Miltenyi Biotec, Germany). Frequencies of virus-specific $CD8^+$ T cells were calculated as described previously.[18] Data were acquired using an LSRFortessa flow cytometer and analyzed with FlowJo software version 10 (FlowJo LLC, Ashland, OR).

Results

Identification of HLA class I-associated viral sequence polymorphisms in L-HDAg

To assess the impact of HDV-specific $CD8^+$ T-cell-mediated selection pressure *in vivo*, we analyzed a cohort of 104 patients with chronic HDV/HBV infection for HLA class I-associated viral sequence polymorphisms (HLA footprints) in L-HDAg. Using a highly significant cut-off value ($P < 0.005$) based on studies of other persistent viral infections,[19, 20] we identified 21 viral sequence polymorphisms that were associated with specific HLA class I alleles (Table 1).

The strongest association linked residue 170 in L-HDAg with HLA-B*15. All eight HLA-$B*15^+$ patients harbored viruses with a serine (S) to asparagine (N) substitution at this position (S170N) compared with only 16/96 HLA-$B*15^-$ patients ($P = 2.9 \times 10^{-8}$) (Figure 1A and C).

Of note, only 5/21 significant polymorphisms (24%) were associated with HLA-A alleles, whereas 16/21 significant polymorphisms (76%) were associated with HLA-B alleles (Table 1). Moreover, these HLA footprints were distributed quite evenly across the central region of HDAg, sparing the N-terminal 30 amino acids of S/L-HDAg, the C-terminal 25 amino acids of S-HDAg, and the entire C-terminal extension unique to L-HDAg (Figure 2A).

Validation of $CD8^+$ T-cell epitopes and viral escape mutations in L-HDAg

To confirm the relevance of these associations, we used *in silico* prediction tools to analyze viral amino acid sequences for HLA class I binding motifs, guided by the corresponding footprint data (see Materials and Methods). In addition to 8mer, 9mer, and 10mer peptides, an extended window incorporating 15 residues either side of the HLA footprint was tested to allow the identification of $CD8^+$ T-cell epitopes associated with 'flanking' mutations, which can disrupt antigen processing and facilitate viral escape.[21-25] Using this approach with defined cut-offs (an $IC_{50} \leq 1{,}000$ nM, a SYFPEITHI score ≥ 20, or a BIMAS score ≥ 20), we identified 18 candidate

$CD8^+$ T-cell epitopes spanning 17 of the 21 HLA footprints (Figures 1B, 3B, and 4B; Table 1 and supplementary Table S3). As expected, these candidates included two HLA-B*27-restricted epitopes identified previously using the same sequence dataset, validating the overall strategy.[6] Of note, 12/18 HLA footprints were located within the respective candidate epitopes, whereas 6/18 HLA footprints were located in the regions flanking the respective candidate epitopes. These latter footprints did not impact antigen processing *in silico*, as predicted using an online algorithm (www.iedb.org),[26] but negative results were also obtained in similar analyses of flanking mutations known to impair the generation of epitopes derived from HCV and HIV (supplementary Table S4).

We then evaluated the immunogenicity of candidate epitopes in functional assays using peptide-stimulated $CD8^+$ T-cell lines generated from HLA class I-matched patients with resolved HDV infection (n = 12) or chronic HDV/HBV infection (n = 13) (supplementary Table S2). Five novel HDV-specific $CD8^+$ T-cell epitopes were identified using this approach (Table 1), including one restricted by HLA-B*15 (Figure 1), one restricted by HLA-B*18 (Figure 3), two restricted by HLA-B*37 (Figure 4), and one restricted by HLA-B*41 (data not shown). In further experiments, we generated peptide-specific $CD8^+$ T-cell lines targeting either the prototype or variant epitopes, and tested the impact of viral sequence variation using IFN-γ production assays. As shown in Figures 1D, 3D, and 4D, and more clearly in serial dilution assays (Figures 1E, 3E, and 4E), the variant peptides were recognized suboptimally by some, but not all, prototype-specific $CD8^+$ T cells. Conversely, variant-specific $CD8^+$ T cells responded poorly to the corresponding prototype and variant peptides without exception, consistent with population-level viral escape (Figures 1D, 1E, 3D, 3E, 4D, and 4E).

An association of special interest was detected at residue 47 (Figure 3A and C). Here, the consensus aspartate (D) was conserved in 13/14 HLA-$B*18^+$ patients, whereas the variant glutamate (E) was found in 42/90 HLA-$B*18^-$ patients (P = 0.0027). This observation is

compatible with the notion that HLA-B*18 drives population-level evolution from glutamate, which likely represents the 'ancestral' residue, to aspartate and further suggests that the E47D substitution does not markedly impair viral fitness, given the relative paucity of sequence revertants in the absence of HLA-B*18. In line with these interpretations, the variant peptide (47E) induced a small but reproducible IFN-γ response, whereas the consensus peptide (47D) was not recognized in parallel assays (Figure 3D and E).

Viral evolution during super-infection with HDV

To probe the biological implications of these findings, we characterized the emergence of viral escape mutations in longitudinal samples from an HLA-B*15^{+} patient with acute HDV/HBV super-infection (P18 in supplementary Table S2). This patient presented with alanine aminotransferase (ALT) levels approaching 2,000 U/L, which declined to < 500 U/L over a period of one week, and serum HDV RNA levels approaching 10^7 IU/mL, which declined by approximately two orders of magnitude in the first week and subsequently relapsed to plateau at > 10^7 IU/mL (Figure 5A). Over a period of > 200 weeks prior to initiation of antiviral therapy with pegylated IFN-α, viral sequence analysis revealed a total of four amino acid substitutions in L-HDAg. Two of these mutations were reversions toward the consensus sequence, and two mutations encoded *de novo* substitutions away from the consensus sequence (Figure 5B). Of note, one of the reversion mutations occurred at an HLA-A*02 anchor residue, suggesting primary acquisition of an escape virus from an HLA-A*02^{+} individual (patient P18 lacked HLA-A*02). More importantly, one of the *de novo* mutations (S170N) was a *bone fide* escape variant located in L-HDAg$_{170–179}$ SMQGVPESPF, matching the HLA-B*15 footprint identified in our cohort of patients with chronic HDV/HBV infection (Figure 1 and Table 1). These observations clearly demonstrate the longitudinal accumulation of viral escape mutations in L-HDAg and

tentatively link this phenomenon with $CD8^+$ T-cell failure in the context of super-infection with HDV.

HDV sequence polymorphisms are associated with infrequent HLA class I alleles

As noted above, most of the HLA class I footprints identified in our cohort were associated with HLA-B alleles (Table 1). In further analyses, we realized that common HLA class I alleles, such as HLA-A*01, -A*02, -A*03, -B*07, -B*08, -B*35, and -B*44, all of which occurred at frequencies > 10% in the German reference population, were not associated with viral sequence polymorphisms (Figure 2B). Conversely, significant footprints in the HDV genome were almost exclusively associated with relatively uncommon HLA class I alleles, namely HLA-A*29, -A*30, -B*13, -B*14, -B*15, -B*18, -B*27, -B*37, -B*38, -B*41, -B*49, and -B*51 (Figure 2B and Table 1).

To extend these observations, we compared HLA class I allele frequencies in our cohort of patients with chronic HDV/HBV infection with the corresponding HLA class I allele frequencies in Germany, Italy, and Iran, where the primary medical centers were located for recruitment purposes (Figure 2B). Additional patients were recruited from a medical center in Barcelona, but HLA class I allele frequency data were not available for the general population in Spain. The overall distribution of HLA class I allele frequencies in our cohort of patients with chronic HDV/HBV infection was similar to the overall distribution of HLA class I allele frequencies in Germany, Italy, and Iran (Figure 2B). However, the footprint-linked allele HLA-B*51 was relatively common in our study cohort and in the general populations of Italy and Iran (10–11%), and we were unable to identify an HLA-B*51-restricted candidate epitope in HDV (Table 1 and supplementary Table S3). Other immune selection pressures, such as interactions between the Bw4 motif and killer cell immunoglobulin-like receptors (KIRs), may therefore drive viral mutation in the context of HLA-B*51. Similarly, the footprint-linked allele HLA-B*18 was relatively common

in our study cohort (6.7%) and in the general population of Italy (9.7%). This observation is compatible with the negative association between HLA-B*18 and sequence variation at residue 47 in L-HDAg.

These data can be explained by the concept of viral evolution at the population level. Accordingly, relatively infrequent HLA class I allotypes continue to drive the accumulation of escape variants *in vivo*, whereas more common HLA class I allotypes shape the circulating quasispecies and therefore no longer select for *de novo* mutations in HDV.

HDV-specific CD8$^+$ T cells are maintained at very low frequencies after viral escape

In line with previous studies,[4, 6] we were unable to detect HDV-specific CD8$^+$ T cells directly *ex vivo*, even with the aid of fluorochrome-labeled tetramers corresponding to the HLA-B*15-restricted epitope L-HDAg$_{170–179}$ SMQGVPESPF (Figure 6A, upper panels). We therefore characterized these cells using a tetramer-based enrichment strategy.[27, 28] Distinct populations of tetramer$^+$ CD8$^+$ T cells were detected after enrichment in all seven HLA-B*15$^+$ patients with chronic HDV/HBV infection (P13, P14, P16, P17, P18, P21, and P23) (Figure 6A, lower panels, and Figure 6B). The frequencies of these HLA-B*15-restricted HDV-specific CD8$^+$ T cells were low, however, subordinate even to rarely detectable virus-specific CD8$^+$ T cells in patients with chronic HCV infection (Figure 6B). In patients P13, P14, P17, P18, P21, and P23, SMQGVPESPF-specific CD8$^+$ T cells displayed a predominant effector-memory phenotype, whereas in patient P16, SMQGVPESPF-specific CD8$^+$ T cells displayed a predominant central-memory phenotype (Figure 6C). These HDV-specific CD8$^+$ T cells expressed intermediate to high levels of PD-1 (Figure 6D), intermediate levels of the inhibitory receptor KLRG1 (Figure 6D), and intermediate levels of the transcription factors T-bet and Eomes (Figure 6E). In line with these characteristics, which contrast with the typical phenotype of terminally exhausted cells (PD-1^{hi}KLRG1$^+$T-bet^{dim}Eomes^{hi}), SMQGVPESPF-specific CD8$^+$ T cells also expressed low levels

of the activation marker CD38, akin to HCV-specific $CD8^+$ T cells targeting escape variants (Figure 6F). Moreover, SMQGVPESPF-specific $CD8^+$ T cells preferentially displayed a PD-1^+CD127^+ phenotype (Figure 6G) and expressed relatively high levels of TCF1 (Figure 6H), a transcription factor that defines memory-like cells with proliferative capacity.[29] Similar levels of TCF1 were expressed by SMQGVPESPF-specific $CD8^+$ T cells in a patient (P5) with resolved HDV infection (Figure 6H). Expression of the pro-survival factor BCL2 was more variable, however, with the highest levels expressed by SMQGVPESPF-specific $CD8^+$ T cells in patient P5 (Figure 6H). These phenotypic characteristics broadly mirror those reported previously for HCV-specific $CD8^+$ T cells in the absence of antigenic stimulation, reflecting viral escape or viral clearance.[17, 29] Although further studies are required to establish such parallels across different viral infections, it is notable that SMQGVPESPF-specific $CD8^+$ T cells in patient P14, who retained a subpopulation of prototype viruses, expressed relatively low levels of TCF1 and BCL2 in conjunction with an atypical $CD38^+$PD-1^+CD127^- phenotype (Figure 6F and G).

Discussion

HDV/HBV infection is one of the few remaining difficult-to-treat conditions in viral hepatitis, leading to liver cirrhosis, liver failure, and hepatocellular carcinoma in many patients.[1, 2] Despite the urgent need for novel prophylactic and therapeutic interventions, remarkably little is known about the natural repertoire of virus-specific $CD8^+$ T cells, which are thought to play a key role in immune protection against HDV. In this study, we analyzed HLA class I-associated viral sequence polymorphisms in L-HDAg, revealing new insights into the immunobiology of HDV infection. First, we extended the repertoire of defined HDV-specific $CD8^+$ T-cell epitopes from four, encompassing two restricted by HLA-A*02[4] and two restricted by HLA-B*27,[6] to nine, encompassing one restricted by HLA-B*15, one restricted by HLA-B*18, two restricted by HLA-B*37, and one restricted by HLA-B*41. Second, we demonstrated that viral escape occurs during persistent HDV infection. Third, we uncovered a link between viral escape and infrequent HLA class I alleles, indicating population-level adaptation of HDV. Finally, we demonstrated that HDV-specific $CD8^+$ T cells in patients with resolved or persistent HDV infection display similar non-terminally-exhausted phenotypes, consistent with viral escape in the absence of clearance. In line with this interpretation, HLA-B*15-restricted HDV-specific $CD8^+$ T cells that reproducibly select for viral escape could be expanded from patients with chronic HDV/HBV infection, whereas HDV-specific $CD8^+$ T cells restricted by HLA class I allotypes that less commonly select for viral escape could not be expanded from patients with chronic HDV/HBV infection.

The phenomenon of viral adaptation at the population level may lead to a relative paucity of $CD8^+$ T-cell epitopes restricted by common HLA class I allotypes, which in turn may impede the development of an effective vaccine against HDV. In line with this prediction, viral sequence polymorphisms were not associated with the common HLA alleles A*01, A*02, A*03, A*24, B*07, B*08, B*35, and B*44, and suitable binding motifs for many of the corresponding allotypes were not present in the HDV proteome (L-HDAg).[6] However, these findings need to be confirmed

using direct immunogenicity assays, because prediction algorithms are potentially fallible, and certain epitopes may not be subject to viral escape, particularly if they are located in biologically constrained regions of L-HDAg. It is important to note in this context that our experimental strategy was not designed to identify $CD8^+$ T-cell epitopes that are not subject to viral escape. Moreover, our study cohort did not include patients with acute HDV infection, potentially limiting the detection of novel $CD8^+$ T-cell specificities. An unbiased approach will therefore require to define the full range of immunogenic epitopes derived from HDV.[30] The association between immunogenicity and infrequent allotypes reported here may also be confounded by HLA class I polymorphisms. For example, the viral mutation S170N was found in all patients with HLA-B*15:01, but was not found in patients with HLA-B*15:10 or HLA-B*15:18 (data not shown). Similarly, common polymorphisms in HLA-B*27 can lead to subtype-specific patterns of epitope targeting in the context of other persistent viral infections, including HCV and HIV.[31, 32]

Viral escape has direct implications for immune efficacy and the development of immunotherapeutic strategies in the setting of chronic HDV/HBV infection. Recent vaccine studies in chimpanzees and humans with chronic HCV infection showed good immunogenicity against non-conserved viral epitopes, but the corresponding vaccine-boosted $CD8^+$ T cells were ineffective, because they failed to recognize the circulating virus.[33, 34] Effective therapeutic vaccines will therefore likely need to prime and/or boost virus-specific $CD8^+$ T cells that target conserved epitopes.

The most significant footprint-linked allele in the present study was HLA-B*15, which has also been shown to select for viral escape mutations during persistent HCV infection.[35] In addition, genetic association studies have linked HLA-B*15 and/or its serological equivalent (B62) with better outcomes after infection with HCV.[36-38] However, it remains unclear if HLA-B*15 is associated with spontaneous clearance of HDV, and moreover, any such associations will be extremely difficult to identify, given the clinical heterogeneity of HDV-infected patients and

the requirement for co-infection with HBV. In our study cohort, the presence of HDV-specific $CD8^+$ T-cell responses did not correlate with the outcome of infection (clearance versus persistence). Of note, however, the presence of HDV amino acid variations corresponding to polymorphisms associated with the patient's HLA alleles (supplementary Table S1) displayed a trend for association with both, lower HDV RNA titers (median 20,638 IU/ml [range 62 – 20,098,000,000 IU/ml] in patients with ≥1 variation *versus* median 61,930 IU/ml [range 10 – 9,804,000 IU/ml] in patients without any variation; P = 0.054 [Mann Whitney Test]) and lower ALT levels (median 73 U/L [range 19 – 281 U/L] in patients with ≥1 variation *versus* median 86 U/L [range 31 – 507 U/L] in patients without any variation; P = 0.070 [Mann Whitney Test]), respectively. These results suggest that HLA class I footprint mutations indicate efficient $CD8^+$ T cell pressure, leading to partial control of viremia and liver inflammation. This interpretation is based on the assumption, however, that partial cross-recognition of escape variants occurs and is sufficient to exert pressure on the virus.

It is particularly notable that our findings overlap considerably those reported previously in the setting of persistent infection with HCV or HIV. For example, in all three infections, virus-specific $CD8^+$ T-cell pressure can select escape variants to fixation (HDV: HLA-B*15-associated polymorphism S170N; HCV: HLA-A*01-associated polymorphism Y1444F)[39] and drive viral evolution at the population-level to extinguish commonly restricted epitopes (HDV: HLA-B*18-restricted epitope L-$HDAg_{47–54}$; HCV: HLA-A*01-restricted epitope $NS3_{1436–1444}$;[39] HIV: HLA-B*51-restricted epitope $RT_{128–135}$).[40] Similarly, in all three infections, virus-specific $CD8^+$ T cells are preferentially restricted by HLA-B,[41-44] rare allotypes confer a particular advantage,[45] and viral escape mutations undergo reversion in the absence of selection pressure after transmission.[46, 47] Finally, in all three infections, $CD8^+$ T cells targeting escape variants display a 'memory-like' phenotype, indicating a loss of antigenic drive.[17, 25, 29]

Despite these striking parallels across distinct viruses with respect to adaptive host-pathogen interactions, it is important to note that the immunobiology of HDV infection is complicated by the mandatory requirement for co-infection with HBV. Much work will therefore be required to unravel the parameters that govern clinical outcome in the midst of constantly shifting immune responses that simultaneously target rapidly evolving antigens derived from two different viruses in a genetically diverse host population. The findings reported here are nonetheless informative with respect to the underlying cellular and molecular processes and can be considered as an early step on a challenging journey.

Acknowledgements

We thank the individuals who donated blood for this study, the physicians and nurses involved in patient recruitment and sample acquisition, and the technicians who isolated PBMCs. We are especially grateful to Christin Ackermann, Thomas Schirdewahn, and Ulrich Spengler for providing information from patients.

References

1. Hughes SA, Wedemeyer H, Harrison PM. Hepatitis delta virus. Lancet 2011;378:73-85.
2. Heidrich B, Yurdaydin C, Kabacam G, et al. Late HDV RNA relapse after peginterferon alpha-based therapy of chronic hepatitis delta. Hepatology 2014;60:87-97.
3. Roggendorf M. Perspectives for a vaccine against hepatitis delta virus. Semin Liver Dis 2012;32:256-61.
4. Huang YH, Tao MH, Hu CP, et al. Identification of novel HLA-A*0201-restricted CD8+ T-cell epitopes on hepatitis delta virus. J Gen Virol 2004;85:3089-98.
5. Schirdewahn T, Grabowski J, Owusu Sekyere S, et al. The third signal cytokine interleukin 12 rather than immune checkpoint inhibitors contributes to the functional restoration of hepatitis D virus-specific T cells. J Infect Dis 2017;215:139-149.
6. Karimzadeh H, Kiraithe MM, Kosinska AD, et al. Amino acid substitutions within HLA-B*27-restricted T cell epitopes prevent recognition by hepatitis delta virus-specific CD8+ T cells. J Virol 2018;92:e01891-17.
7. Neumann-Haefelin C. HLA-B27-mediated protection in HIV and hepatitis C virus infection and pathogenesis in spondyloarthritis: two sides of the same coin? Curr Opin Rheumatol 2013;25:426-33.
8. Price DA, Brenchley JM, Ruff LE, et al. Avidity for antigen shapes clonal dominance in CD8+ T cell populations specific for persistent DNA viruses. J Exp Med 2005;202:1349-61.
9. Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. Bioinformatics 2007;23:2947-8.
10. Budeus B, Timm J, Hoffmann D. SeqFeatR for the discovery of feature-sequence associations. PLoS One 2016;11:e0146409.
11. Hoof I, Peters B, Sidney J, et al. NetMHCpan, a method for MHC class I binding prediction beyond humans. Immunogenetics 2009;61:1-13.
12. Nielsen M, Lundegaard C, Worning P, et al. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci 2003;12:1007-17.
13. Rammensee H, Bachmann J, Emmerich NP, et al. SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics 1999;50:213-9.
14. Parker KC, Bednarek MA, Coligan JE. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. J Immunol 1994;152:163-75.
15. Neumann-Haefelin C, McKiernan S, Ward S, et al. Dominant influence of an HLA-B27 restricted CD8+ T cell response in mediating HCV clearance and evolution. Hepatology 2006;43:563-72.
16. Thimme R, Oldach D, Chang KM, et al. Determinants of viral clearance and persistence during acute hepatitis C virus infection. J Exp Med 2001;194:1395-406.
17. **Bengsch B, Seigel B**, Ruhl M, et al. Coexpression of PD-1, 2B4, CD160 and KLRG1 on exhausted HCV-specific CD8+ T cells is linked to antigen recognition and T cell differentiation. PLoS Pathog 2010;6:e1000947.
18. Alanio C, Lemaitre F, Law HK, et al. Enumeration of human antigen-specific naive CD8+ T cells reveals conserved precursor frequencies. Blood 2010;115:3718-25.
19. **Carlson JM, Brumme CJ, Martin E**, et al. Correlates of protective cellular immunity revealed by analysis of population-level immune escape pathways in HIV-1. J Virol 2012;86:13202-16.
20. Kefalakes H, Budeus B, Walker A, et al. Adaptation of the hepatitis B virus core protein to CD8+ T-cell selection pressure. Hepatology 2015;62:47-56.

21. **Allen TM, Altfeld M**, Yu XG, et al. Selection, transmission, and reversion of an antigen-processing cytotoxic T-lymphocyte escape mutation in human immunodeficiency virus type 1 infection. J Virol 2004;78:7069-78.
22. Milicic A, Price DA, Zimbwa P, et al. CD8+ T cell epitope-flanking mutations disrupt proteasomal processing of HIV-1 Nef. J Immunol 2005;175:4618-26.
23. **Walker A, Skibbe K**, Steinmann E, et al. Distinct escape pathway by hepatitis C virus genotype 1a from a dominant CD8+ T cell response by selection of altered epitope processing. J Virol 2016;90:33-42.
24. Seifert U, Liermann H, Racanelli V, et al. Hepatitis C virus mutation affects proteasomal epitope processing. J Clin Invest 2004;114:250-9.
25. Timm J, Walker CM. Mutational escape of CD8+ T cell epitopes: implications for prevention and therapy of persistent hepatitis virus infections. Med Microbiol Immunol 2015;204:29-38.
26. **Tenzer S, Peters B**, Bulik S, et al. Modeling the MHC class I pathway by combining predictions of proteasomal cleavage, TAP transport and MHC class I binding. Cell Mol Life Sci 2005;62:1025-37.
27. **Nitschke K, Flecken T, Schmidt J**, et al. Tetramer enrichment reveals the presence of phenotypically diverse hepatitis C virus-specific CD8+ T cells in chronic infection. J Virol 2015;89:25-34.
28. Schmidt J, Neumann-Haefelin C, Altay T, et al. Immunodominance of HLA-A2-restricted hepatitis C virus-specific CD8+ T cell responses is linked to naive-precursor frequency. J Virol 2011;85:5232-6.
29. Wieland D, Kemming J, Schuch A, et al. TCF1+ hepatitis C virus-specific CD8+ T cells are maintained after cessation of chronic antigen stimulation. Nat Commun 2017;8:15050.
30. Landahl J, Bockmann JH, Scheurich C, et al. Detection of a broad range of low level MHC class II-restricted Hepatitis Delta Virus (HDV)-specific T cell responses regardless of the clinical status. J Infect Dis 2018 (e-pub ahead of print).
31. Adland E, Hill M, Lavandier N, et al. Differential immunodominance hierarchy of CD8+ T-cell responses in HLA-B*27:05- and -B*27:02-mediated control of HIV-1 infection. J Virol 2018;92:e01685-17.
32. Nitschke K, Barriga A, Schmidt J, et al. HLA-B*27 subtype specificity determines targeting and viral evolution of a hepatitis C virus-specific CD8+ T cell epitope. J Hepatol 2014;60:22-9.
33. Callendret B, Eccleston HB, Satterfield W, et al. Persistent hepatitis C viral replication despite priming of functional CD8+ T cells by combined therapy with a vaccine and a direct-acting antiviral. Hepatology 2016;63:1442-54.
34. **Kelly C, Swadling L**, Capone S, et al. Chronic hepatitis C viral infection subverts vaccine-induced T-cell immunity in humans. Hepatology 2016;63:1455-70.
35. Ruhl M, Knuschke T, Schewior K, et al. CD8+ T-cell response promotes evolution of hepatitis C virus nonstructural proteins. Gastroenterology 2011;140:2064-73.
36. **Fitzmaurice K, Hurst J**, Dring M, et al. Additive effects of HLA alleles and innate immune genes determine viral outcome in HCV infection. Gut 2015;64:813-9.
37. Hraber P, Kuiken C, Yusim K. Evidence for human leukocyte antigen heterozygote advantage against hepatitis C virus infection. Hepatology 2007;46:1713-21.
38. Nitschke K, Luxenburger H, Kiraithe MM, et al. CD8+ T-cell responses in hepatitis B and C: the (HLA-) A, B, and C of hepatitis B and C. Dig Dis 2016;34:396-409.
39. Neumann-Haefelin C, Frick DN, Wang JJ, et al. Analysis of the evolutionary forces in an immunodominant CD8 epitope in hepatitis C virus at a population level. J Virol 2008;82:3438-51.

40. **Kawashima Y, Pfafferott K, Frater J, Matthews P**, et al. Adaptation of HIV-1 to human leukocyte antigen class I. Nature 2009;458:641-5.
41. **Gaudieri S, Rauch A**, Park LP, et al. Evidence of viral adaptation to HLA class I-restricted immune pressure in chronic hepatitis C virus infection. J Virol 2006;80:11094-104.
42. Kiepiela P, Leslie AJ, Honeyborne I, et al. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature 2004;432:769-75.
43. **Neumann-Haefelin C, Timm J**, Spangenberg HC, et al. Virological and immunological determinants of intrahepatic virus-specific CD8+ T-cell failure in chronic hepatitis C virus infection. Hepatology 2008;47:1824-36.
44. Timm J, Li B, Daniels MG, et al. Human leukocyte antigen-associated sequence polymorphisms in hepatitis C virus reveal reproducible immune responses and constraints on viral evolution. Hepatology 2007;46:339-49.
45. **Trachtenberg E, Korber B**, Sollars C, et al. Advantage of rare HLA supertype in HIV disease progression. Nat Med 2003;9:928-35.
46. **Leslie AJ, Pfafferott KJ**, Chetty P, et al. HIV evolution: CTL escape mutation and reversion after transmission. Nat Med 2004;10:282-9.
47. **Timm J, Lauer GM**, Kavanagh DG, et al. CD8 epitope escape and reversion in acute HCV infection. J Exp Med 2004;200:1593-604.

Author names in bold designate shared co-first authorship.

Figure Legends

Figure 1: HLA-B*15-restricted CD8$^+$ T-cell responses specific for L-HDAg$_{170–179}$ drive viral escape. A P values for the association between HDV sequence polymorphisms and the presence of HLA-B*15, plotted for each amino acid residue in the L-HDAg protein. Cut-off for significance was set at P = 0.005 (dotted red line). **B** Predicted HLA-B*15 binding affinity of candidate peptide epitopes in L-HDAg. IC_{50} values were predicted for 8mers, 9mers, and 10-mers using the ANN 3.4 method (www.iedb.org). The best hit corresponding to the confirmed HLA-B*15-restricted epitope L-HDAg$_{170–179}$ is highlighted in green. **C** Viral sequences from individual HLA-B*15$^+$ and HLA-B*15$^–$ patients compared with the consensus sequence derived from all 104 patients (all HLA-B*15$^+$ patients are shown above the line, and the first 20 HLA-B*15$^–$ patients are shown below the line). Dots indicate agreement with consensus, and single-letter amino acid codes indicate variation from consensus. The HLA-B*15-associated polymorphism S170N is highlighted in dark grey, and the confirmed HLA-B*15-restricted epitope L-HDAg$_{170–179}$ is highlighted in grey. **D** PBMCs from HLA-B*15$^+$ patient P13 (2 years after IFN treatment) were expanded in the presence of the wildtype or variant peptide corresponding to the L-HDAg$_{170–179}$ epitope. After 14 days, cells were restimulated in parallel with each peptide and tested for IFN-γ production. Negative controls without peptide restimulation and positive controls stimulated with phorbol myristate acetate/ionomycin are also shown. **E** Equivalent results after restimulation with each peptide in serial dilution assays. PBMCs from patient P13 (2 and 7 years after IFN treatment, respectively) and HLA-B*15$^+$ patient P21 were processed and tested as in (**D**).

Figure 2: Viral sequence polymorphisms are associated with infrequent HLA class I alleles and spare the N- and C-termini of L-HDAg. A Distribution of identified HLA footprints across L-HDAg. Red arrows indicate HLA footprints corresponding to confirmed (predescribed or

newly identified) epitopes. Filled black arrows indicate HLA footprints corresponding to predicted epitopes. Empty black arrows indicate HLA footprints without defined or predicted epitopes. Newly identified epitopes are shown as red bars, and predescribed epitopes are shown as blue bars. **B** HLA-A and HLA-B allele frequencies in the study cohort compared with reference populations from Germany, Italy, and Iran. HLA class I alleles identified in at least two patients in the study cohort are shown (HLA class I alleles present in one patient cannot reach statistical power for the identification of HLA class I-associated viral sequence polymorphisms). Footprint-linked HLA class I alleles are shown in bold and marked with an arrowhead.

Figure 3: HLA-B*18-restricted CD8$^+$ T-cell responses specific for L-HDAg$_{46–54}$ drive viral escape and population-level evolution of HDV. A–**E** Details as per the corresponding panels in Figure 1. In this case, the ancestral peptide is denoted as prototype, and the variant peptide is denoted as consensus, reflecting a negative association between HLA-B*18 and the viral polymorphism D47E. Representative data from patient P3 are shown in (D) and (E).

Figure 4: The rare allotype HLA-B*37 selects for viral escape in the CD8$^+$ T-cell epitopes L-HDAg$_{81–90}$ and L-HDAg$_{100–108}$. A–**E** Details as per the corresponding panels in Figure 1. In (B), the BIMAS score (higher scores indicate higher binding affinities) was used in place of the ANN 3.4 method, which does not include a prediction algorithm for HLA-B*37 binders. Representative data from patient P1 (epitope L-HDAg$_{100–108}$) are shown in (**D**) and (**E**).

Figure 5: Viral escape in the HLA-B*15-restricted epitope L-HDAg$_{170–179}$ after super-infection with HDV. A Clinical course of HDV/HBV super-infection in patient P18 (HLA-A*24:02$^+$, -B*15:01$^+$, -B*35:02$^+$). **B** Viral sequences in patient P18 at early (week 0 and week 2) and late time points (week 71 and week 200) after super-infection with HDV. All longitudinal

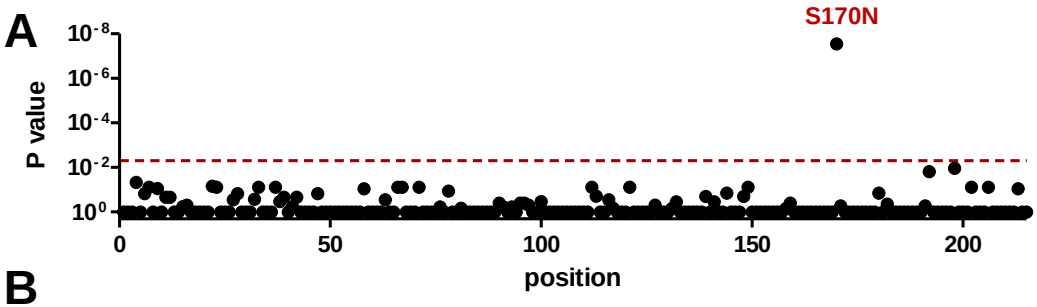
sequence variants are shown (n = 4). Dots indicate agreement with consensus across the study cohort. Residues with *de novo* mutations are shown in red, and residues that revert to wildtype are shown in green. Of note, the *de novo* mutation S170N is located in the confirmed HLA-B*15-restricted epitope L-HDAg$_{170–179}$ and corresponds to the linked footprint (Table 1). Also of note, the F27L reversion co-localizes with an anchor residue for the predescribed HLA-A*02-restricted epitope L-HDAg$_{26–34}$.

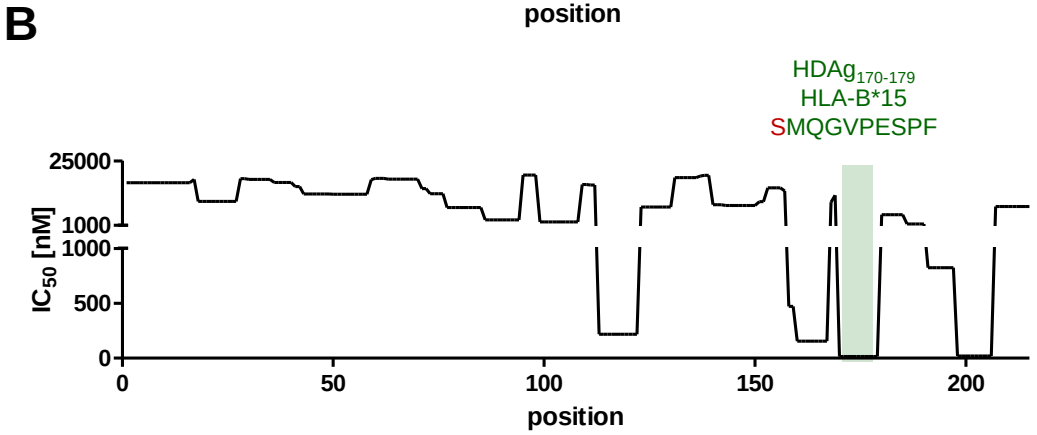
Figure 6: Characterization of HLA-B*15-restricted HDV-specific CD8$^+$ T cells in patients with resolved or persistent HDV infection. A HLA-B*15-restricted CD8$^+$ T cells specific for L-HDAg$_{170–179}$ were tetramer-enriched from PBMCs isolated from patients with chronic HDV/HBV infection. Representative tetramer stainings from patient P13 (relatively large tetramer$^+$ population) and patient P17 (relatively small tetramer$^+$ population) are shown before (upper panels) and after enrichment (lower panels). **B** Frequency of tetramer$^+$ cells among total CD8$^+$ T cells. HDV-specific CD8$^+$ T cells from seven patients with chronic HDV/HBV infection and one patient with resolved HDV infection (P5, highlighted in red) were compared with HCV-specific CD8$^+$ T cells targeting epitopes with conserved (wildtype, wt) or escaped viral sequences in chronic HCV infection. Autologous viral sequences corresponding to the L-HDAg$_{170–179}$ epitope are displayed next to the patient codes. S170N indicates the presence of the viral escape sequence, and S170S/N indicates presence of the prototype and the viral escape sequence, respectively. ND: not done (sequences were not obtained from patients with low levels of HDV RNA). **C**–**H** Enriched HDV-specific CD8$^+$ T cells were characterized by flow cytometry. Grey histograms indicate non-naive bulk CD8$^+$ T cells. Representative plots are derived from patient P21. **C** Distribution among naive (CD45RA$^+$CD27$^+$CCR7$^+$), central-memory (CM, CD45RO$^+$CCR7$^+$), effector-memory (EM, CD45RO$^+$CCR7$^–$), and terminally differentiated effector-memory subsets (TEMRA, CD45RO$^–$CCR7$^–$). **D** Expression of the exhaustion markers

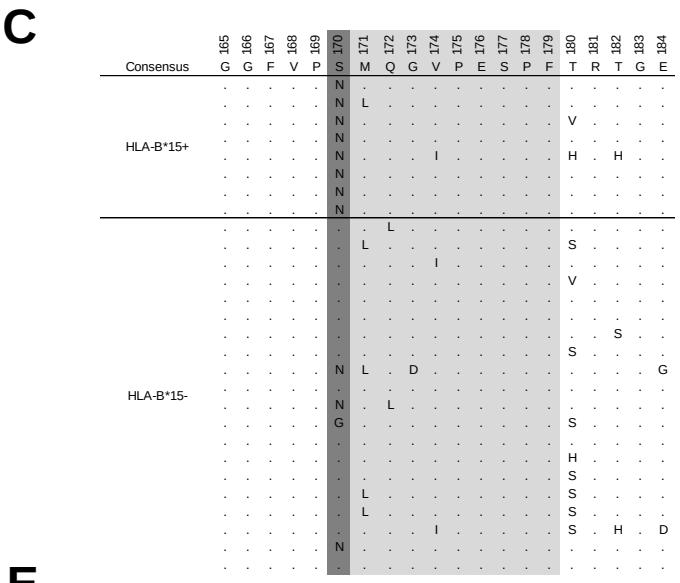
PD-1 and KLRG1. **E** Expression of the transcriptions factors T-bet and Eomes. **F** Expression of the activation marker CD38 compared with HCV-specific $CD8^+$ T cells targeting epitopes with conserved (wildtype, wt) or escaped viral sequences in chronic HCV infection. **G** Distribution among subsets defined by expression of PD-1 and CD127. **H** Expression of the transcription factors TCF1 and BCL2. Dotted lines indicates fluorescence minus one (FMO) controls. MFI: median fluorescence intensity.

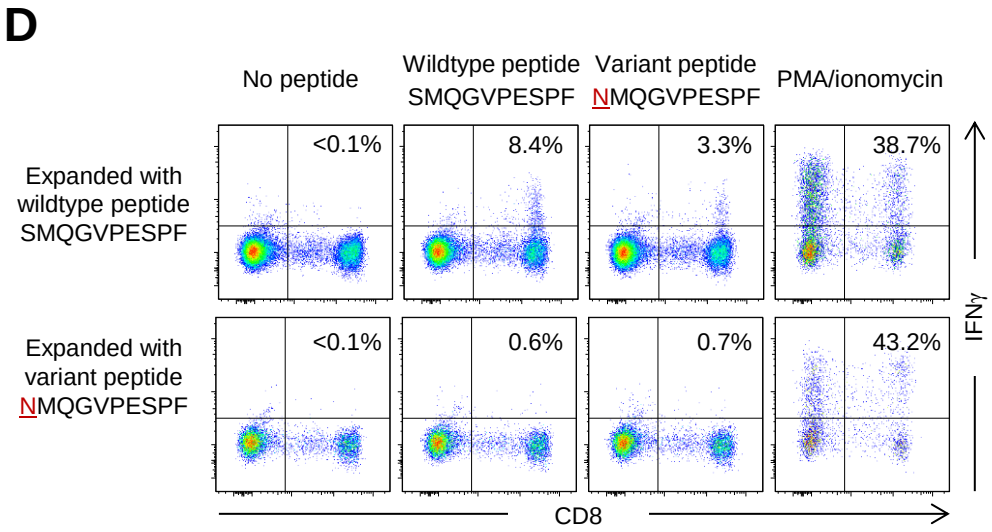
Table 1: HLA class I-associated viral sequence polymorphisms and HDV-specific CD8$^+$ T cell epitopes. Confirmed epitopes are displayed in bold and red. nd: not done.

Nr.	HLA	AA position	Mutation	P value	AA sequence epitope candidate	AA range	Positive responses / number of patients tested
1	**B*15**	**170**	**S170N**	**2.9 x 10^{-8}**	**SMQGVPESPF**	**170–179**	**10 / 14**
2	B*13	33	D33E	0.0001	DLRKVKKKI	33–41	0 / 5
3	**B*37**	**101**	**D101E**	**0.0002**	**QDHRRRKAL**	**100–108**	**1 / 1**
4	A*29	63	K63R	0.0002	–	–	–
5	A*30	47	D47E	0.0010	KVKKKIKK	36–43	0 / 3
6	**B*37**	**89**	**P89T/I**	**0.0011**	**VDSGPRKRPL**	**81–90**	**1 / 1**
7	B*49	37	V37A/T	0.0015	EELERDLRKV	28–37	nd
8	B*13	100	Q100K	0.0018	RQDHRRRKAL	99–108	0 / 5
9	B*51	81	V81I	0.0019	–	–	–
10	B*13	43	K43R	0.0021	DLRKVKKKI	33–41	0 / 5
11	B*41	158	G158A/D/M	0.0023	–	–	–
12	**B*18**	**47**	**E47D**	**0.0027**	**DENPWLGNI**	**46–54**	**1 / 5**
13	A*33	37	V37T/A	0.0028	EELERDLR	28–35	nd
14	B*14	107	A107T	0.0028	DHRRRKAL	101–108	0 / 1
15	A*30	49	P49L/S	0.0031	KVKKKIKK	36–43	0 / 3
16	**B*41**	**139**	**R139K**	**0.0034**	**RERRVAGPPV**	**140–149**	**1 / 1**
17	B*13	96	D96E	0.0035	RQDHRRRKAL	99–108	0 / 5
18	**B*27**	**105**	**R105K**	**0.0039**	**RRKALENKK**	**104–112**	**previously described[6]**
					RRDHRRRKAL	**99–108**	**previously described[6]**
19	B*38	131	K131G	0.0039	–	–	–
20	B*13	113	K113R	0.0043	RQDHRRRKAL	99–108	0 / 5
					KQLSAGGKNL	113–122	0 / 5
21	A*68	134	T134A	0.0045	LTEEDERR	133–140	0 / 5


A
S170N
P value
position


B
HDAg170-179
HLA-B*15
SMQGVPESPF
IC50 [nM]
position


C
Consensus
HLA-B*15+
HLA-B*15-


D
No peptide
Wildtype peptide SMQGVPESPF
Variant peptide NMQGVPESPF
PMA/ionomycin
Expanded with wildtype peptide SMQGVPESPF
Expanded with variant peptide NMQGVPESPF
<0.1%
8.4%
3.3%
38.7%
<0.1%
0.6%
0.7%
43.2%
IFNγ
CD8

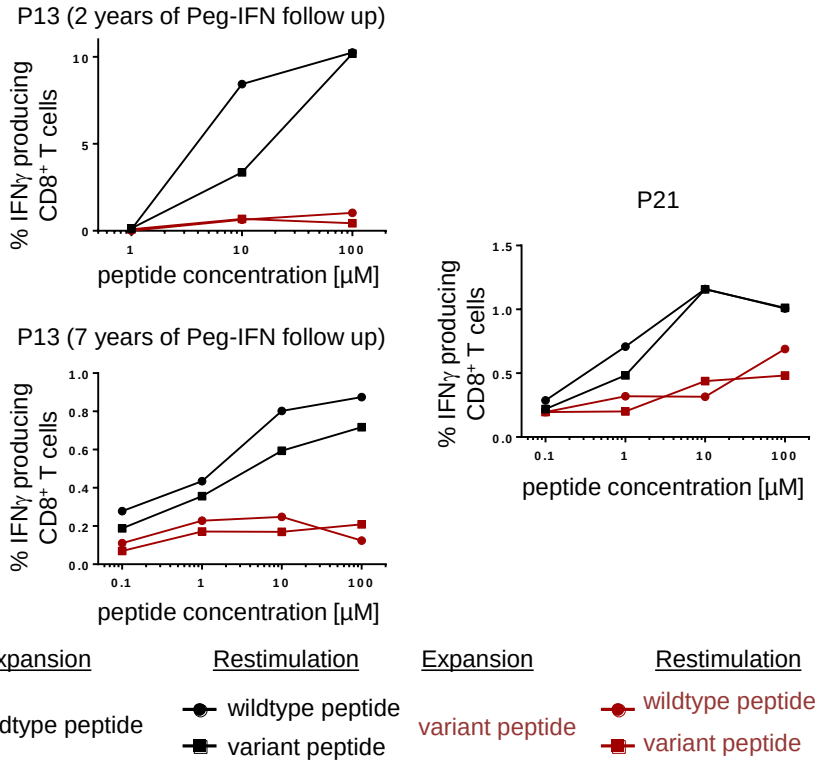

E
P13 (2 years of Peg-IFN follow up)
P13 (7 years of Peg-IFN follow up)
P21
% IFNγ producing CD8+ T cells
peptide concentration [μM]
Expansion
Restimulation
wildtype peptide
variant peptide

Figure 1

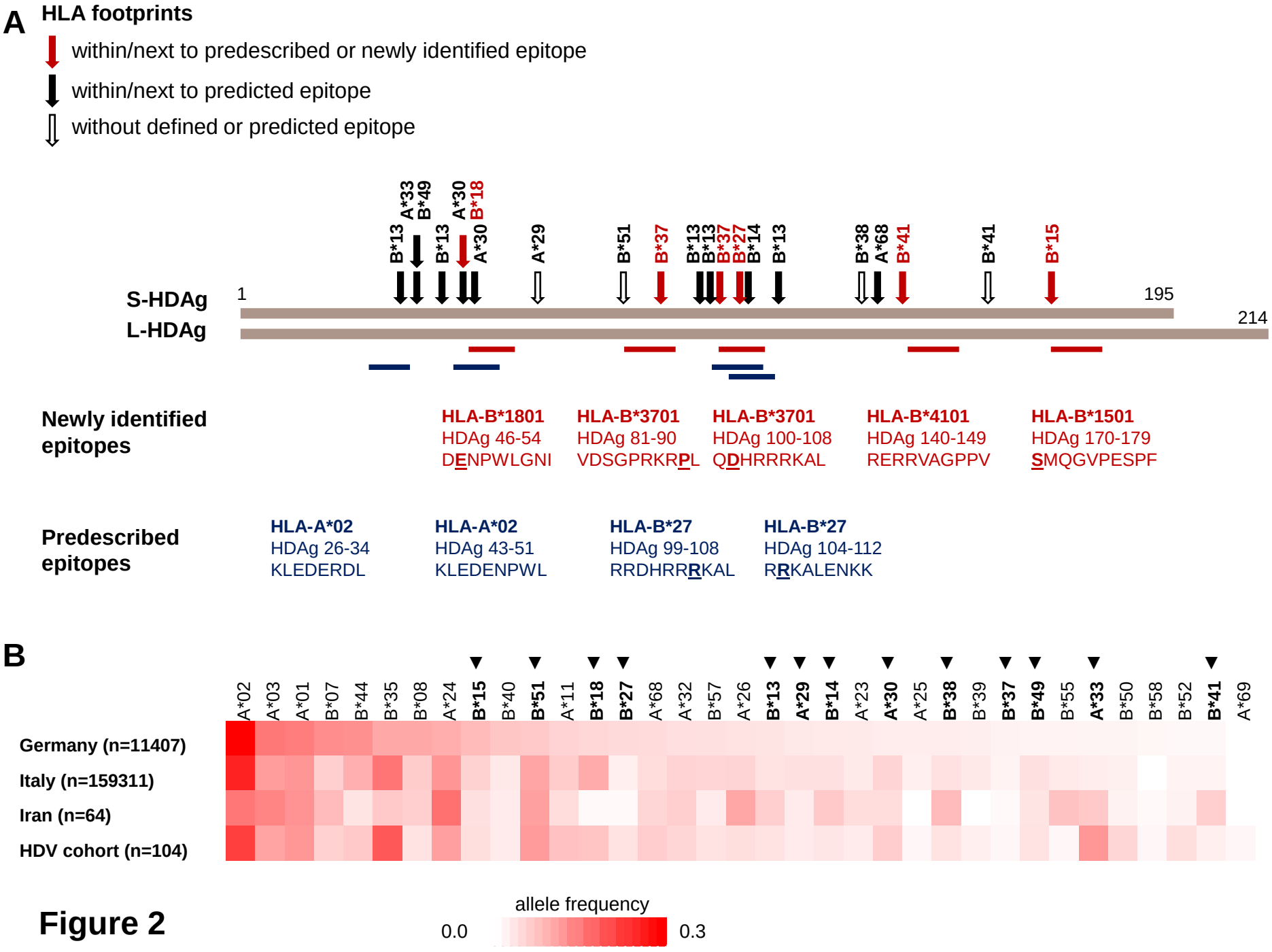
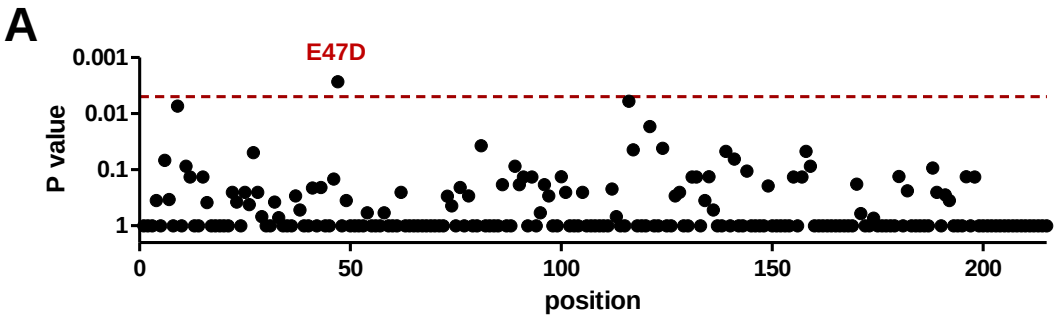
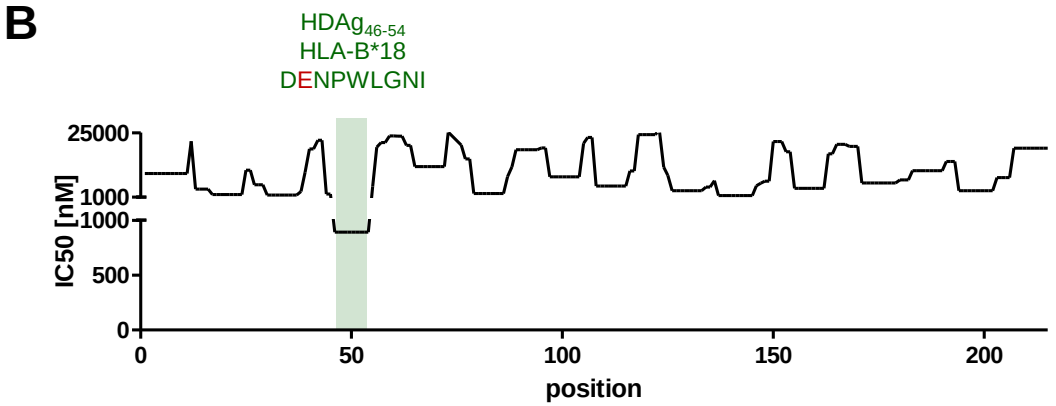
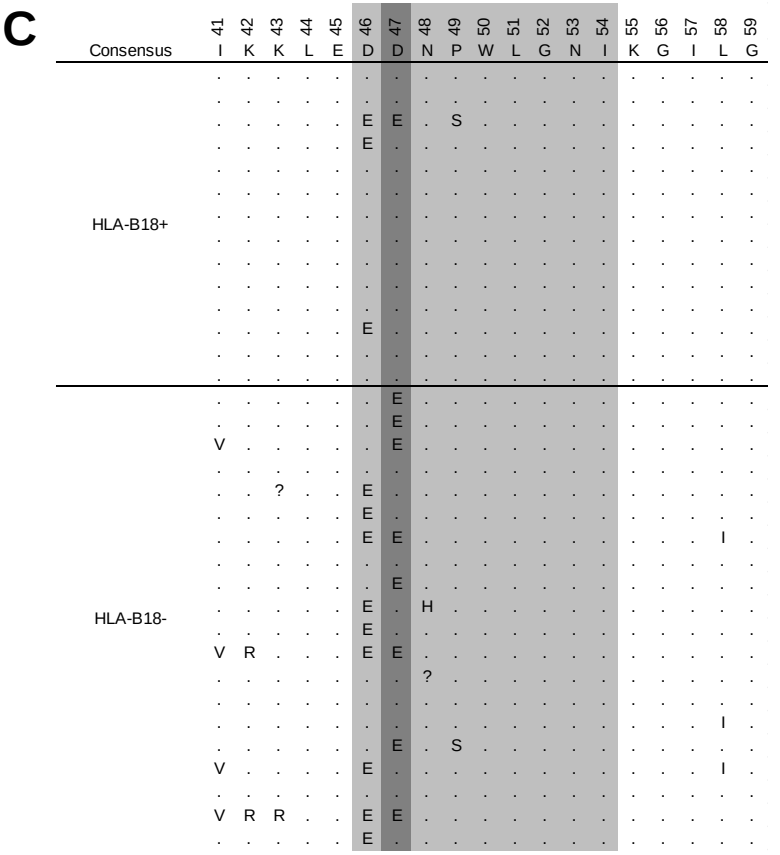
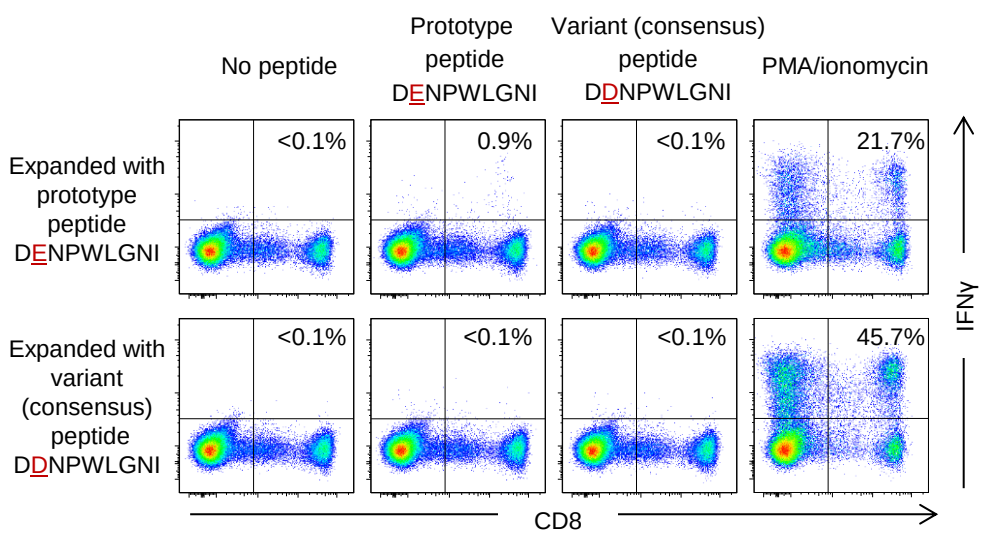

A
HLA footprints
within/next to predescribed or newly identified epitope
within/next to predicted epitope
without defined or predicted epitope
B*13
A*33
B*49
B*13
A*30
B*18
A*30
A*29
B*51
B*37
B*13
B*13
B*37
B*27
B*14
B*13
B*38
A*68
B*41
B*41
B*15
S-HDAg
L-HDAg
1
195
214
Newly identified epitopes
HLA-B*1801
HDAg 46-54
DENPWLGNI
HLA-B*3701
HDAg 81-90
VDSGPRKRPL
HLA-B*3701
HDAg 100-108
QDHRRRKAL
HLA-B*4101
HDAg 140-149
RERRVAGPPV
HLA-B*1501
HDAg 170-179
SMQGVPESPF
Predescribed epitopes
HLA-A*02
HDAg 26-34
KLEDERDL
HLA-A*02
HDAg 43-51
KLEDENPWL
HLA-B*27
HDAg 99-108
RRDHRRRKAL
HLA-B*27
HDAg 104-112
RRKALENKK
B
A*02
A*03
A*01
B*07
B*44
B*35
B*08
A*24
B*15
B*40
B*51
A*11
B*18
B*27
A*68
A*32
B*57
A*26
B*13
A*29
B*14
A*23
A*30
A*25
B*38
B*39
B*37
B*49
B*55
A*33
B*50
B*58
B*52
B*41
A*69
Germany (n=11407)
Italy (n=159311)
Iran (n=64)
HDV cohort (n=104)
allele frequency
0.0
0.3

Figure 2


A
E47D
P value
0.001
0.01
0.1
1
0
50
100
150
200
position


B
HDAg46-54
HLA-B*18
DENPWLGNI
IC50 [nM]
25000
1000
1000
500
0
0
50
100
150
200
position


C
Consensus
41 I
42 K
43 K
44 L
45 E
46 D
47 D
48 N
49 P
50 W
51 L
52 G
53 N
54 I
55 K
56 G
57 I
58 L
59 G
HLA-B18+
HLA-B18-


D
No peptide
Prototype peptide DENPWLGNI
Variant (consensus) peptide DDNPWLGNI
PMA/ionomycin
Expanded with prototype peptide DENPWLGNI
Expanded with variant (consensus) peptide DDNPWLGNI
<0.1%
0.9%
<0.1%
21.7%
<0.1%
<0.1%
<0.1%
45.7%
IFNγ
CD8

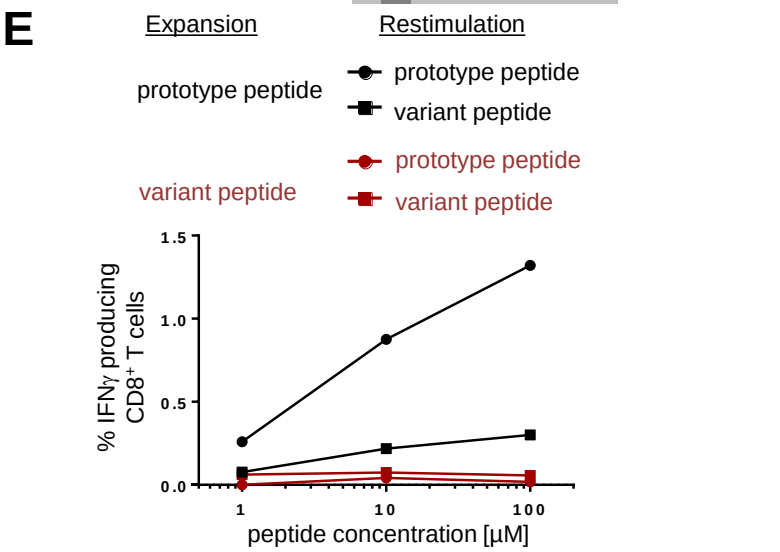

E
Expansion
Restimulation
prototype peptide
prototype peptide
variant peptide
variant peptide
prototype peptide
variant peptide
% IFNγ producing CD8+ T cells
1.5
1.0
0.5
0.0
1
10
100
peptide concentration [µM]

Figure 3

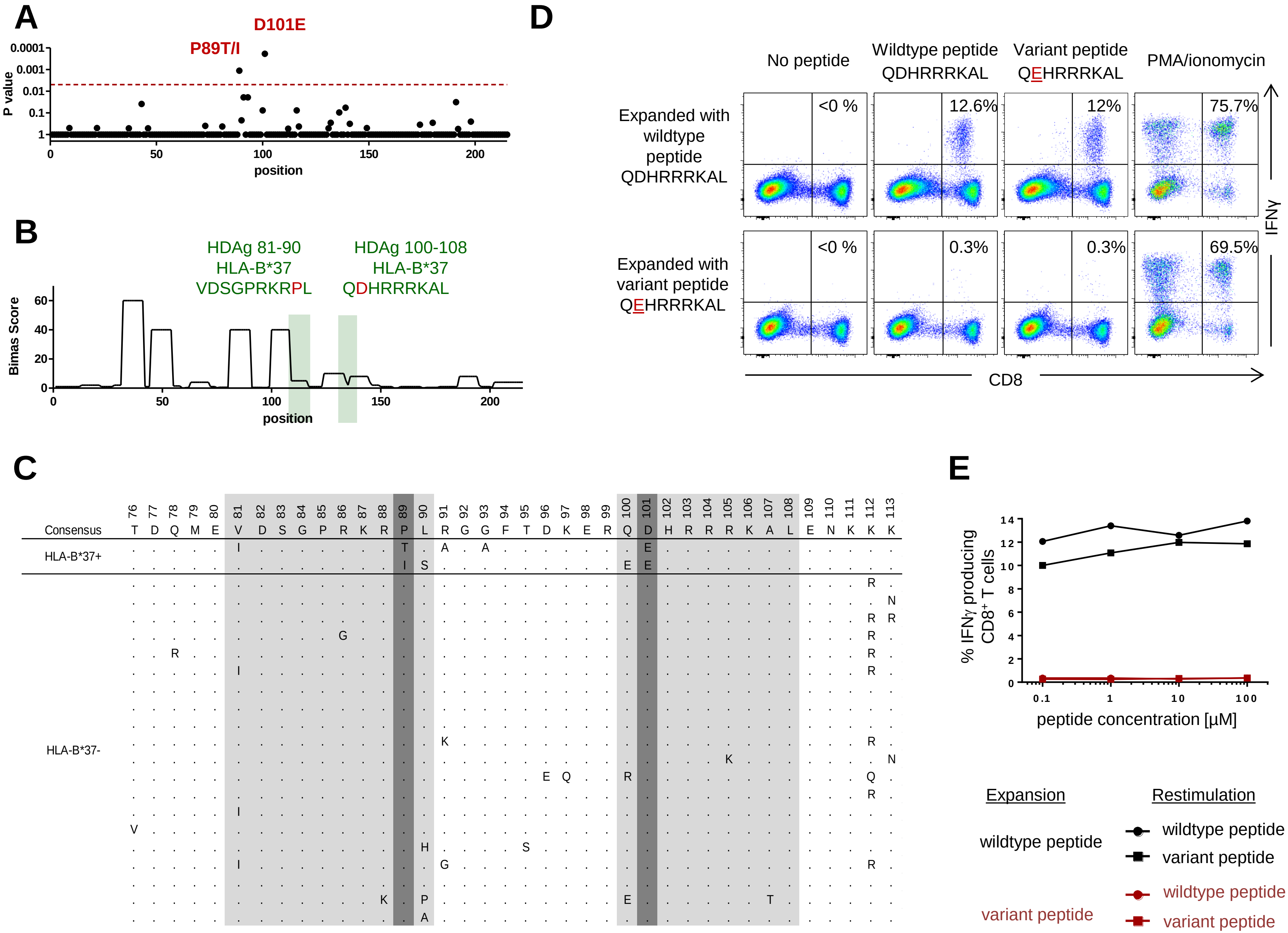
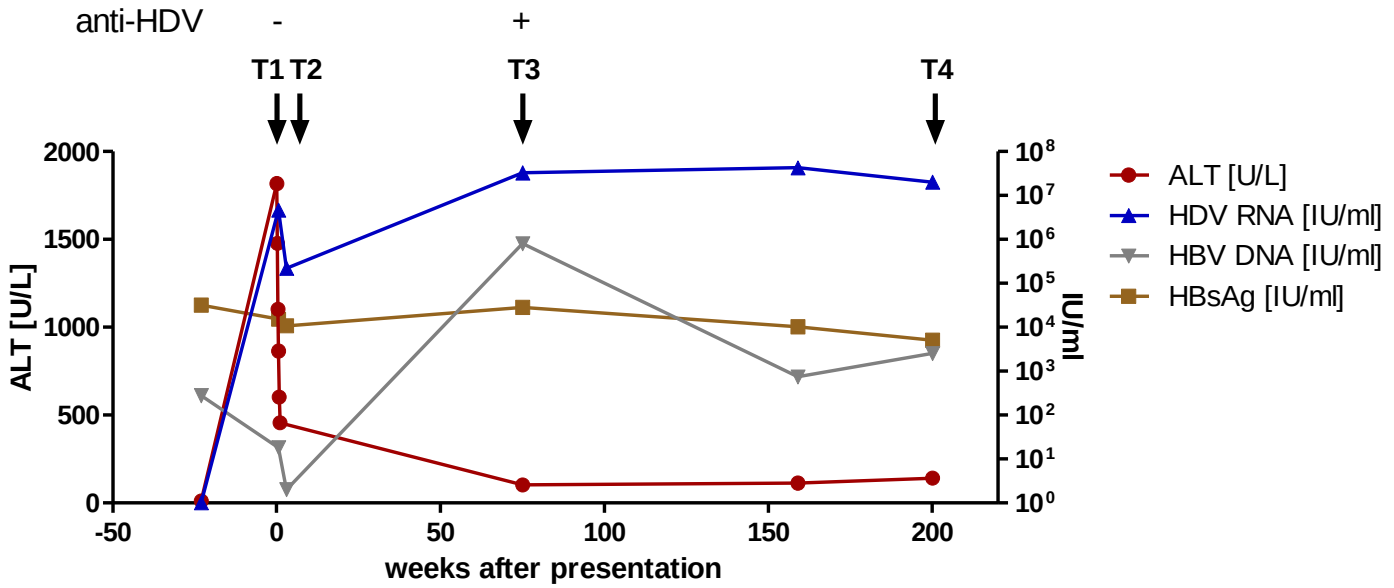

A
D101E
P89T/I
P value
position
B
HDAg 81-90
HLA-B*37
VDSGPRKRPL
HDAg 100-108
HLA-B*37
QDHRRRKAL
Bimas Score
position
C
Consensus
HLA-B*37+
HLA-B*37-
D
No peptide
Wildtype peptide QDHRRRKAL
Variant peptide QEHRRRKAL
PMA/ionomycin
Expanded with wildtype peptide QDHRRRKAL
Expanded with variant peptide QEHRRRKAL
<0 %
12.6%
12%
75.7%
0.3%
69.5%
CD8
IFNγ
E
% IFNγ producing CD8+ T cells
peptide concentration [µM]
Expansion
Restimulation
wildtype peptide
variant peptide

Figure 4

A


anti-HDV
-
+
T1 T2
T3
T4
2000
1500
1000
500
0
ALT [U/L]
-50
0
50
100
150
200
weeks after presentation
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^1
10^0
IU/ml
ALT [U/L]
HDV RNA [IU/ml]
HBV DNA [IU/ml]
HBsAg [IU/ml]

B

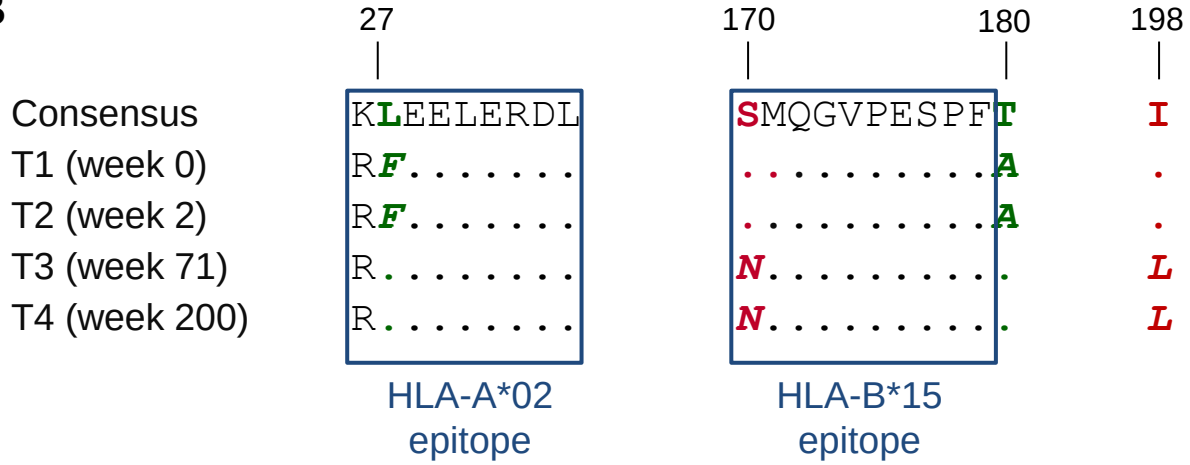

27
170
180
198
Consensus KLEELERDL SMQGVPESPFT I
T1 (week 0) RF.......A .
T2 (week 2) RF.......A .
T3 (week 71) R.......N.....L
T4 (week 200) R.......N.....L
HLA-A*02 epitope
HLA-B*15 epitope

Figure 5

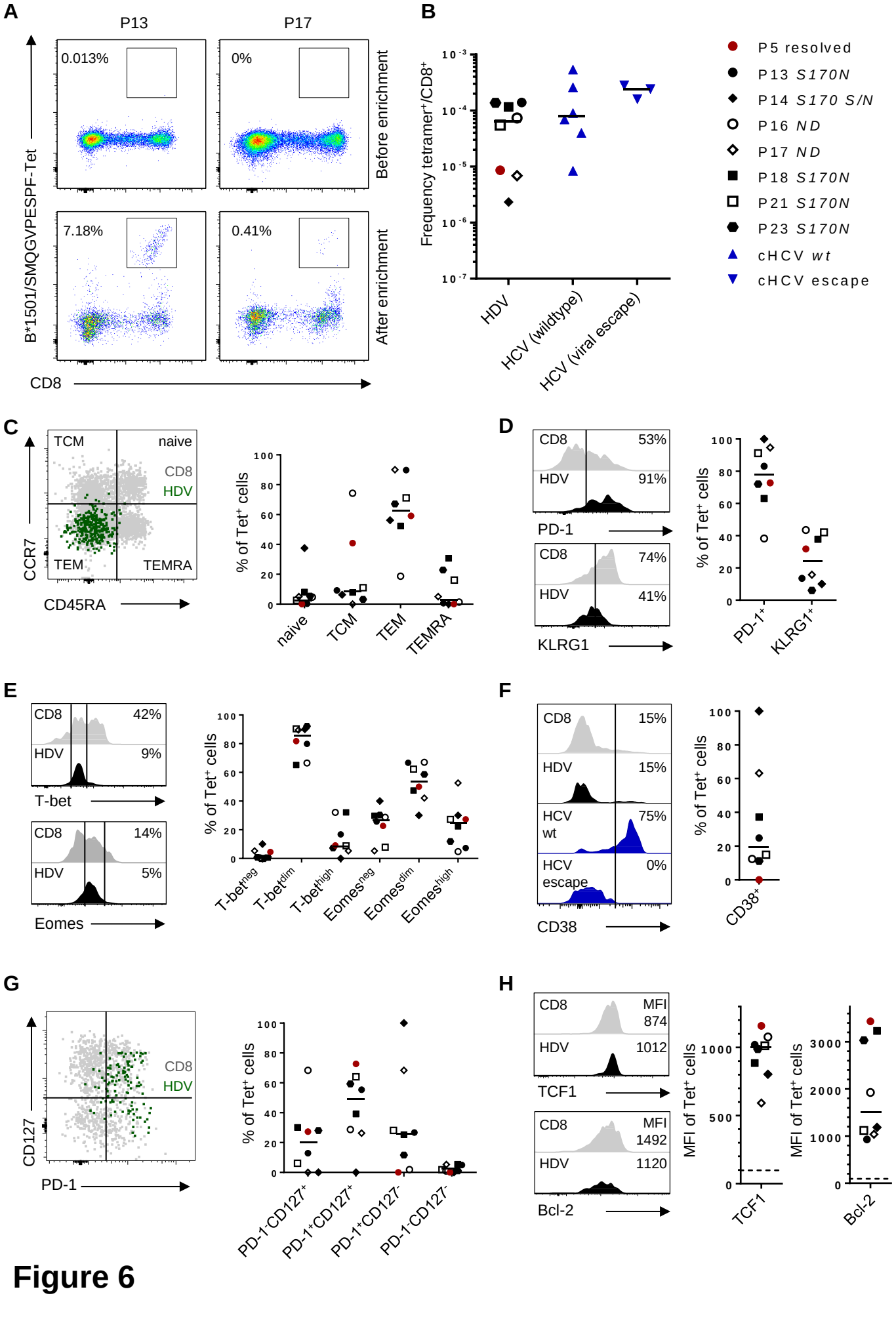

A
P13
P17
0.013%
0%
7.18%
0.41%
Before enrichment
After enrichment
B*1501/SMQGVPESPF-Tet
CD8
B
Frequency tetramer+/CD8+
HDV
HCV (wildtype)
HCV (viral escape)
P5 resolved
P13 S170N
P14 S170 S/N
P16 ND
P17 ND
P18 S170N
P21 S170N
P23 S170N
cHCV wt
cHCV escape
C
TCM
naive
CD8
HDV
TEM
TEMRA
CCR7
CD45RA
% of Tet+ cells
naive
TCM
TEM
TEMRA
D
CD8
53%
HDV
91%
PD-1
CD8
74%
HDV
41%
KLRG1
% of Tet+ cells
PD-1+
KLRG1+
E
CD8
42%
HDV
9%
T-bet
CD8
14%
HDV
5%
Eomes
% of Tet+ cells
T-betneg
T-betdim
T-bethigh
Eomesneg
Eomesdim
Eomeshigh
F
CD8
15%
HDV
15%
HCV wt
75%
HCV escape
0%
CD38
% of Tet+ cells
CD38+
G
CD8
HDV
CD127
PD-1
% of Tet+ cells
PD-1-CD127+
PD-1+CD127+
PD-1+CD127-
PD-1-CD127-
H
CD8
MFI 874
HDV
1012
TCF1
CD8
MFI 1492
HDV
1120
Bcl-2
MFI of Tet+ cells
TCF1
MFI of Tet+ cells
Bcl-2

Figure 6

Supplementary Table S1: Patients' characteristics (patients included in analysis of HLA-associated viral sequence polymorphisms). 3TC, Lamivudine; ETV, Entecavir; IFN, interferon; na, not available; NUC, nucleos(t)ide analogon (not specified); pegIFN, pegylated interferon; TDF, Tenofovir. *Number of HDV amino acid variations corresponding to polymorphisms associated with the patient's HLA-A or HLA-B alleles (compare Table 1).

No	Age	Sex	HLA-A	HLA-B	Escape mutations [n]*	Current therapy	Previous Therapy	ALT [U/L]	Bilirubin [mg/dl]	INR	Platelets [G/L]	Cirrhosis [yes/no]	HDV load [IU/ml]	HBV load [IU/ml]
1	51	M	A26, A30	B14, B41	3	ETV	IFN	116	1.3	1.02	204	no	21,48,763	38
2	52	M	A03, A33	B08, B14	0	TDF	-	190	2.1	1.38	39	yes	3,00,001	<20
3	45	M	A02, A30	B13, B44	3	TDF	-	281	0.6	1.25	125	yes	45,82,744	40
4	45	M	A24, A24	B35, B44	0	TDF	IFN	86	1.1	1.29	84	yes	2,57,842	24
5	46	M	A03, A11	B07, B35	0	TDF	-	131	0.9	1.12	171	no	35,796	neg
6	36	M	A24, A32	B35, B51	1	TDF	-	93	1.1	1.57	55	yes	1,71,690	59
7	48	M	A01, A03	B35, B57	0	ETV	IFN	111	0.5	1.12	95	no	1,68,068	<20
8	44	M	A02, A33	B14, B49	0	ETV	IFN	259	0.6	1.25	293	no	93,530	28
9	63	M	A01, A02	B15, B56	1	TDF	IFN	72	0.9	1.20	57	yes	20,91,055	<20
10	52	M	A02, A23	B49, B50	1	-	-	40	0.3	1.09	196	no	80,928	15,000
11	37	F	NA	B18, B27	2	ETV	-	42	0.3	1.05	186	no	39,044	910
12	39	F	A01, A02	B15, B57	1	ETV	-	232	2.6	1.28	133	yes	2,768	<20
13	32	M	A23, A24	B35, B62	0	TDF	-	133	13.1	1.58	111	yes	3,694	<20
14	36	M	A01, A02	B07, B08	0	TDF	-	56	2.4	1.07	156	yes	<10	<20
15	26	F	A02, A03	B35, B35	0	-	IFN; pegIFN + TDF	44	0.4	na	173	no	2,367	neg
16	43	M	A01, A03	B18, B50	1	TDF	pegIFN + 3TC	55	1.5	1.20	115	yes	765	<20
17	29	M	A24	B44, B51	1	-	pegIFN + TDF	24	0.4	0.90	50	no	703	neg
18	46	M	A02, A24	B35, B52	0	TDF	-	60	1.7	1.22	51	yes	13,929	neg
19	38	F	A24, A68	B14, B35	1	ETV	-	172	19.1	1.55	82	yes	24,555	neg
20	71	M	A02	B41, B51	1	ETV	-	30	1.4	1.71	51	yes	3,92,837	neg
21	34	M	A02, A68	B15	1	-	pegIFN + ETV	121	0.5	na	47	yes	7,78,038	neg
22	16	M	A01, A24	B35, B58	0	3TC	-	507	1.1	1.11	235	no	5,57,336	<20
23	56	M	A01, A03	B08, B35	0	TDF	-	31	0.4	0.95	180	yes	6,506	neg
24	32	M	A02	B55, B60	0	TDF + ETV	ETV	192	3.1	1.07	87	yes	7,70,954	<20
25	44	M	A02, A26	B38, B51	0	TDF	IFN	283	0.7	1.00	42	yes	98,04,320	neg
26	52	F	A02, A11	B51, B55	0	Telbivudine	-	42	0.9	1.14	47	yes	37,033	neg

27	46	M	A01, A11	B35, B51	0	TDF	Telbivudine	166	0.4	0.99	28	yes	<10	<20
28	40	M	A03, A11	B35, B51	0	ETV	-	249	1.0	1.15	193	no	8,241	neg
29	49	M	A11, A30	B18, B35	2	ETV	-	45	0.8	1.28	103	yes	13,270	<20
30	51	M	A02, A03	B18, B41	3	ETV	-	36	0.5	1.04	211	no	1,486	neg
31	59	M	A02	B50, B52	0	TDF	IFN + 3TC	116	0.8	1.11	125	yes	4,85,646	<20
32	22	M	A24, A31	B35, B51	0	ETV + TDF	ETV	175	0.4	1.06	132	no	32,47,316	<20
33	70	M	A03, A24	B18, B35	1	pegIFN + TDF	-	53	0.5	0.96	65	yes	7,061	neg
34	54	M	A03, A32	B18, B51	2	-	pegIFN	105	10.9	1.65	144	yes	29,842	neg
35	28	F	A02, A24	B38, B51	1	-	-	99	0.2	na	343	na	1,166	neg
36	60	F	A03	B07, B40	0	IFN + TDF	-	42	0.7	1.10	17	yes	2,572	neg
37	31	M	A02, A29	B15, B44	1	na	na	na	na	na	na	na	2,216	neg
38	33	M	A01, A11	B35, B40	0	ETV	None	na	na	na	167	no	18,81,647	267
39	27	M	A66, A68	B52, B57	0	na	na	na	na	na	na	na	46,485	neg
40	44	M	A24	B18, B51	2	pegIFN + TDF	IFN	77	0.4	1.10	70	no	125	neg
41	45	M	A02	B44	0	-	-	84	2.5	1.42	55	yes	1,36,571	neg
42	41	F	A03, A32	B15, B52	1	TDF		47	0.5	1.09	115	yes	62	neg
43	29	F	A02, A03	B07, B35	0	-	-	146	1.4	1.10	187	no	49,600	36
44	34	M	A02, A24	B38, B73	0	-	IFN	72	0.2	0.98	195	no	1,91,000	3,800
45	24	M	A30, A32	B13, B35	7	-	-	108	0.5	1.04	179	yes	1,16,000	68
46	49	F	A11, A23	B49, B50	1	-	IFN	52	0.4	1.16	91	no	1,99,000	20
47	34	M	A03, A30	B13, B51	8	-	-	189	0.2	1.12	152	no	#########	83,900
48	44	M	A01, A30	B38, B40	1	-	-	47	0.6	1.13	102	yes	2,668	1,130
49	42	M	A26, A29	B07, B44	0	-	-	75	0.4	1.15	190	no	1,77,000	5,670
50	22	F	A01, A29	B27, B44	1	-	IFN	104	0.5	na	na	no	18,600	1,970
51	57	M	A68, A69	B35, B51	1	-	-	224	1.9	1.26	120	yes	17,510	neg
52	42	F	A02, A26	B07, B15	1	-	-	36	0.5	1.13	176	yes	490	136
53	51	F	A03, A24	B08, B51	0	-	-	36	0.4	1.16	93	no	61,930	20
54	46	F	A02	B40	0	-	-	63	0.5	1.14	176	no	72,430	59
55	32	M	A03, A30	B13, B51	7	-	-	83	0.7	1.16	129	no	70,850	678
56	27	M	A29, A68	B15, B47	2	-	IFN	102	1.6	1.18	128	no	1,40,300	neg
57	51	F	A03, A11	B27, B52	0	-	-	99	0.4	1.18	331	no	7,552	46,900
58	36	M	A02, A32	B07, B18	1	-	-	46	0.7	1.26	96	yes	2,990	20
59	49	M	A01, A02	B41, B44	2	-	IFN	107	0.2	1.09	128	no	7,809	34,300

60	49	M	A02, A32	B07, B18	1	-	-	126	0.4	1.11	116	yes	361	914
61	54	M	A03, A24	B35, B38	1	-	-	115	0.6	1.04	140	no	2,87,500	neg
62	31	M	A11, A69	B35, B52	0	-	-	64	0.2	0.95	223	no	14,940	neg
63	38	M	A01, A02	B07, B08	0	-	-	121	0.4	1.18	242	yes	12,07,000	33
64	48	M	A01, A02	B51	0	-	-	80	0.8	na	223	no	14,820	479
65	55	M	A02, A32	B14, B18	1	-	-	40	1.4	na	117	yes	55,420	55
66	31	F	A02, A68	B07, B35	0	-	IFN	44	0.5	1.16	159	yes	8,10,000	238
67	56	F	A02, A24	B50, B51	1	-	-	83	0.5	1.16	233	no	1,49,000	30,13,000
68	44	M	A02, A03	B35, B51	1	-	IFN	37	1.2	1.20	80	yes	611	51
69	31	M	A02, A23	B35, B44	0	IFN	-	42	1.1	1.07	160	no	16,160	33
70	37	M	A02, A25	B18, B39	1	IFN + NUC	-	253	1.3	1.25	33	yes	17,130	80
71	61	M	A02, A24	B27, B50	0	NUC	-	72	6.8	1.15	57	yes	731	12,900
72	39	M	A01, A02	B40, B57	0	NUC	IFN + NUC	97	0.7	1.19	110	yes	23,060	45
73	53	M	A02, A26	B39, B51	0	IFN + NUC	-	141	2.2	1.18	260	yes	21,280	98
74	36	M	A02, A11	B08, B44	0	NUC	-	64	2.3	1.10	149	yes	6,11,000	<20
75	38	F	A01, A02	B35, B51	1	NUC	-	136	2.7	1.51	40	yes	20,120	<20
76	35	F	A02, A03	B13, B35	1	-	IFN	19	0.8	1.16	233	no	33,242	27
77	62	M	A02	B35, B52	0	-	IFN	70	1.2	1.13	166	yes	3,55,057	41
78	56	M	A01, A68	B44, B57	1	-	IFN	109	0.8	1.12	162	yes	1,647	<20
79	26	M	A03, A25	B18, B50	1	-	-	47	2.4	1.40	267	yes	1,558	<20
80	41	M	A02, A68	B18, B37	3	-	IFN	123	0.7	0.91	163	no	15,495	<20
81	57	M	A26, A30	B27, B57	1	-	IFN	241	1.2	1.04	79	yes	4,15,808	32
82	40	M	A11, A24	B07, B44	0	-	IFN	137	0.8	1.07	186	yes	43,14,880	<20
83	56	M	A02, A30	B18, B58	1	-	IFN	204	0.9	1.20	151	yes	39,623	159
84	38	F	A01	B35, B37	2	-	-	37	2.7	1.39	74	yes	988	93
85	47	M	A32, A68	B35	0	-	IFN	66	0.6	1.02	267	yes	2,403	64
86	41	M	A01, A30	B13, B50	4	-	IFN	251	0.8	1.04	95	yes	21,155	<20
87	50	F	A11, A68	B35, B39	1	-	IFN	21	0.6	1.00	113	yes	5,81,197	<20
88	56	M	A23, A30	B42, B49	1	-	-	73	1.5	1.25	58	yes	1,75,704	22
89	46	M	A01, A24	B49, B51	1	-	-	78	1.2	1.30	59	yes	1,25,908	<20
90	54	M	A01, A11	B08, B15	1	-	IFN	222	1.3	1.19	192	no	3,02,564	<20
91	26	M	A01, A02	B50, B51	0	-	IFN	45	1.4	1.22	61	yes	46,405	<20
92	24	M	A02, A68	B13, B35	0	-	IFN	481	0.8	0.90	91	no	10,28,386	neg

93	53	F	A30, A68	B27, B35	1	-	-	46	1.3	na	209	no	3,222	30
94	32	M	A02, A29	B27, B45	2	-	pegIFN	29	0.9	1.10	162	no	5,269	280
95	47	M	A02, A11	B35, B39	0	-	IFN	149	2.3	1.10	98	no	3,19,973	neg
96	71	M	A01, A24	B44, B57	0	-	-	106	0.7	1.20	117	yes	pos	neg
97	64	F	A26, A32	B35, B38	1	TDF	IFN	30	0.9	1.00	131	yes	pos	4,480
98	54	M	A01, A24	B35, B49	1	-	Adefovir	55	1.3	1.50	70	yes	pos	neg
99	53	F	A32, A33	B35, B38	1	TDF	-	32	1.2	1.00	80	yes	pos	neg
100	51	M	A26, A33	B35, B51	2	na	na	na	na	na	na	na	pos	neg
101	53	M	A11, A33	B52	1	TDF	-	49	2.1	1.10	118	yes	pos	neg
102	34	M	A03, A24	B35, B50	0	-	-	70	0.8	1.30	150	yes	pos	neg
103	49	M	A24, A33	B14, B35	1	TDF	-	38	1.1	1.30	152	yes	pos	neg
104	56	F	A01	B49, B51	1	TDF	-	58	1.8	1.40	60	yes	pos	10,432

SupplementaryTable S2: Patients' characteristics (patients with cellular assays). 3TC, Lamivudine; ETV, Entecavir; pegIFN, pegylated interferon; TDF, Tenofovir.

Status	No.	Sex	Age	HLA-A	HLA-B	HBsAg	HBV DNA [IU/ml]	HDV RNA [IU/ml]	Current therapy	Previous therapy	ALT [U/L]	Billirubin [mg/dl]	Platelets [G/L]	INR	Cirrhosis	Epitopes tested	positive
Spontaneously resolved HDV	P1	M	57	A*02:01, A*24:02	B*35:02, B*37:01	pos	neg	neg	TDF	-	89	0.4	271	1.01	no	B37	B37 (2 epitopes)
	P2	M	43	A*23:01, A*26:01	B*41:02, B*44:03	neg	neg	neg	-	-	19	0.4	243	1.02	no	B41	B41 (1 epitope)
	P3	F	34	A*02:01, A*74:01	B*18:01, B*44:03	pos	738	neg	-	-	23	0.4	272	1.03	no	B18	B18
	P4	M	40	A*02:01, A*24:02	B*13:02, B*51:01	pos	neg	neg	-	-	95	0.7	258	0.95	no	B13	-
	P5	M	65	A*30:01, A*68:01	B*13:02, B*15:01	pos	neg	neg	TDF	pegIFN	71	0.3	202	1.00	no	A30, A68, B13, B15	B15
	P6	M	46	A*01, A*68	B*15:17, B*52	pos	neg	neg	TDF	3TC	30	1.3	121	1.4	yes	A68, B15	-
	P7	M	62	A*30:02, A*33:03	B*53:01, B*57:03	pos	1100	neg	-	-	19	1.5	131	1.08	no	A30	-
	P8	M	33	A*02:01, A*03:01	B*18:03, B*35:03	pos	21000	neg	-	-	19	0.8	292	1.09	no	B18	-
IFN-induced HDV resolution	P9	M	32	A*02:01, A*24:02	B*13:02, B*40:02	pos	neg	neg	TDF	pegIFN	42	0.2	238	1.00	no	B13	-
	P10	F	66	A*03	B*15, B*40	pos	<10	neg	TDF	pegIFN	137	0.9	145	1.3	no	B15	B15
	P11	F	39	A*23, A*30	B*15, B*53	pos	240	neg	-	pegIFN	36	0.7	150	1.00	no	A30, B15	-
	P12	M	40	A*02:01, A*24:02	B*15:01, B*51:01	pos	3400	neg	-	pegIFN	31	0.8	128	0.96	no	B15	B15
Chronic HDV infection	P13	M	42	A*01:01, A*02:01	B*08:01, B*15:01	pos	71	10	-	pegIFN, 3TC, ETV, TDF	31	1.5	187	1.05	yes	B15	B15
	P14	F	58	A*02:01	B*07:02, B*15:01	pos	neg	1981	ETV	pegIFN	120	1	90	1.00	yes	B15	B15 (tet only)
	P15	M	36	A*03:01, A*68:01	B*07:02, B*15:10	pos	neg	12	ETV	pegIFN	24	0.3	252	1.01	no	A68, B15	-
	P16	M	48	A*02:01, A*03:01	B*07:02, B*15:01	pos	10	1698	-	-	73	0.5	207	1.16	no	B15	B15
	P17	F	55	A*01:01, A*32:01	B*14:01, B*15:01	pos	46	7264	-	-	26	0.2	288	0.96	no	B14, B15	B15 (tet only)
	P18	F	31	A*24:02	B*15:01, B*35:02	pos	neg	1501664	TDF	pegIFN	112	0.6	212	1.07	no	B15	B15
	P19	M	51	A*02:01, A*24:02	B*18:01, B*40:02	pos	neg	37364	TDF	3TC	73	0.6	180	1.06	yes	B18	-
	P20	F	57	A*01:01, A*02:01	B*13:02, B*35:03	pos	<10	95283	TDF	-	37	4.1	47	1.44	yes	B13	-
	P21	F	36	A*01:01, A*02:01	B*08:01 B*15:01	pos	<10	49034	-	-	34	0.5	210	1.00	no	B15	B15
	P22	M	38	A*24:02, A*68:01	B*18:01, B*27:02	pos	neg	18742	TDF	-	48	3.9	116	1.44	yes	A68, B18	-
	P23	M	25	A*02, A*68	B*13, B*15:01	pos	86	52×10^6	-	pegIFN	570	0.7	169	1.2	no	A68, B13, B15	B15
	P24	M	70	A*01, A*02	B*08, B*15	pos	<10	170×10^6	-	-	227	0.5	113	0.99	yes	B15	-
	P25	F	27	A*02:01, A*25:01	B*18:01	pos	90	15000	-	-	70	0.4	95	1.12	yes	B18	-

Supplementary Table S3: HLA class I associated HDV polymorphisms, predicted and confirmed HDV-specific CD8+ T cell epitopes. Confirmed epitopes are displayed in bold and red. nd: not done.

Nr.	HLA type	Position	Mutation	P value	Aa seqeunce epitope candiate	Aa position	ANN IC50 [nM]	ANN rank	NetMHCPan IC50 [nM]	NetMHCPan rank	Syfpeithi score	Syfpeithi rank	Bimas score	Bimas rank	Positive responses / number of patients tested
1	**B*15**	**170**	**S170N**	**2.9×10^{-8}**	**SMQGVPESPF**	**170-179**	**16**	**1**	**38.76**	**2**	**<20**	**-**	**<20**	**-**	**10 / 14**
2	B*13	33	D33E	0.0001	DLRKVKKKI	33-41	NA	-	>1000	-	21	3	NA	-	0 / 5
3	**B*37**	**101**	**D101E**	**0.0002**	**QDHRRRKAL**	**100-108**	**NA**	**-**	**>1000**	**-**	**26**	**1**	**40**	**2**	**1 / 1**
4	A*29	63	K63R	0.0002	-	-	>1000	-	NA	-	NA	-	NA	-	-
5	A*30	47	D47E	0.0010	KVKKKIKK	36-43	5	1	12.75	1	NA	-	NA	-	0 / 3
6	**B*37**	**89**	**P89T/I**	**0.0011**	**VDSGPRKRPL**	**81-90**	**NA**	**-**	**>1000**	**-**	**NA**	**-**	**40**	**2**	**1 / 1**
7	B*49	37	V37A/T	0.0015	EELERDLRKV	28-37	NA	-	>1000	-	25	1			nd
8	B*13	100	Q100K	0.0018	RQDHRRRKAL	99-108	NA	-	>1000	-	22	1	NA	-	0 / 5
9	B*51	81	V81I	0.0019	-	-	>1000	-	NA	-	NA	-	NA	NA	-
10	B*13	43	K43R	0.0021	DLRKVKKKI	33-41	NA	-	>1000	-	21	3	NA	-	0 / 5
11	B*41	158	G158A/D/M	0.0023	-	-	NA	-	>1000	-	<20	-	NA	-	-
12	**B*18**	**47**	**E47D**	**0.0027**	**DENPWLGNI**	**46-54**	**891**	**1**	**906.16**	**2**	**<20**	**-**	**NA**	**-**	**1 / 5**
13	A*33	37	V37T/A	0.0028	EELERDLR	28-35	261	5	>1000	-	NA	-	NA	-	nd
14	B*14	107	A107T	0.0028	DHRRRKAL	101-108	65	1	925.81	1	<20	-	NA	-	0 / 1
15	A*30	49	P49L/S	0.0031	KVKKKIKK	36-43	5	1	12.75	1	NA	-	NA	-	0 / 3
16	**B*41**	**139**	**R139K**	**0.0034**	**RERRVAGPPV**	**140-149**	**NA**	**-**	**43.01**	**1**	**<20**	**-**	**NA**	**-**	**1 / 1**
17	B*13	96	D96E	0.0035	RQDHRRRKAL	99-108	NA	-	>1000	-	22	1	NA	-	0 / 5
18	**B*27**	**105**	**R105K**	**0.0039**	**RRKALENKK**	**104-112**	**79**	**3**	**118.78**	**3**	**28**	**3**	**6000**	**1**	**previously described[6]**
					RRDHRRRKAL	**99-108**	**59**	**2**	**317.13**	**8**	**24**	**10**	**1800**	**10**	**previously described[6]**
19	B*38	131	K131G	0.0039	-	-	>1000	-	NA	-	NA	-	NA	-	-
20	B*13	113	K113R	0.0043	RQDHRRRKAL	99-108	NA	-	>1000	-	22	1	NA	-	0 / 5
					KQLSAGGKNL	113-122	NA	-	>1000	-	20	4	NA	-	0 / 5
21	A*68	134	T134A	0.0045	LTEEDERR	133-140	210	5	>1000	-	NA	-	NA	-	0 / 5

Search for epitope 15 aa upstream and downstream of mutation (31 aa)
Search strategy:
1. If available ANN (8-10mers; cut-off 1000 nM)
2. If ANN not available: NetMHCpan (8-10mers, cut-off IC50=1000 nM) + Syfpeithi (8-10mers, cut-off score=20) + Bimas (8-10mers, cut off score=20)

Supplementary Table S4: *In silico* prediction of impact of mutations in flanking regions on processing of HDV-specific candidate epitopes as well as HCV- and HIV-specific epitopes with described impact of flanking mutations. The antigen processing algorithm on www.iedb.org was used for prediction.

Footprint number from Table 1 or reference	HLA type	Antigen	Position	Mutation	P value	Aa seqeunce epitope candiate	Aa position	Variant	Proteasome score	TAP score	Processing Score
5	A*30	HDV HDAg	47	D47E	0.001	KVKKKIKK	36-43	D47	0.74	0.37	1.11
								E47	0.72	0.37	1.09
10	B*13	HDV HDAg	43	K43R	0.0021	DLRKVKKKI	33-41	K43	1.14	0.18	1.32
								R43	1.15	0.18	1.32
13	A*33	HDV HDAg	37	V37T/A	0.0028	EELERDLR	28-35	V37	0.90	0.53	1.43
								T37	1.00	0.53	1.53
								A37	0.94	0.53	1.47
15	A*30	HDV HDAg	49	P49L/S	0.0031	KVKKKIKK	36-43	P49	0.74	0.37	1.11
								L49	0.74	0.37	1.11
								S49	0.74	0.37	1.11
16	B*41	HDV HDAg	139	R139K	0.0034	RERRVAGPPV	140-149	R139	0.73	0.26	0.99
								K139	0.73	0.25	0.97
17	B*13	HDV HDAg	96	D96E	0.0035	RQDHRRRKAL	99-108	D96	1.52	0.41	1.92
								E96	1.52	0.41	1.92
20	B*13	HDV HDAg	113	K113R	0.0043	RQDHRRRKAL	99-108	K113	1.52	0.41	1.92
								R113	1.52	0.41	1.92
Walker et al., 2016[23]	B*51	HCV NS3	1368	S1368P	0.00002	IPFYGKAI	1373-1380	S1368	1.17	0.09	1.26
								P1368	1.17	0.09	1.26
Timm et al., 2004[47]	B*08	HCV NS3	1403	L1403V/F	ND	HSKKKCDEL	1395-1403	L1403	1.47	0.39	1.86
								V1403	1.03	0.11	1.13
								F1403	1.31	1.07	2.39
Seifert et al., 2004[24]	A*02	HCV NS3	1082	Y1082F	0.07	CVNGVCWTV	1073-1081	Y1082	1.36	0.14	1.50
								F1082	1.36	0.14	1.45
Allen et al., 2004[21]	A*03	HIV p17 Gag	28	K28Q	ND	KIRLRPGGK	18-26	K28	0.71	0.25	0.97
								Q28	0.82	0.25	1.07
Milicic et al., 2005[22]	B*35	HIV Nef	69; 81; 87	R69K; A81G; H87R	ND	VPLRPMTY	74-81	R69; A81; H87	1.51	1.15	2.67
								K69; G81; R87	1.51	1.15	2.65