

## Analysis of the Evolutionary Forces in an Immunodominant CD8 Epitope in Hepatitis C Virus at a Population Level<sup>∇</sup>

Christoph Neumann-Haefelin,<sup>1</sup> David N. Frick,<sup>2</sup> Jing Jing Wang,<sup>2</sup> Oliver G. Pybus,<sup>4</sup> Shadi Salloum,<sup>3</sup> Gagandeep S. Narula,<sup>2</sup> Anna Eckart,<sup>3</sup> Andrea Biezyński,<sup>3</sup> Thomas Eiermann,<sup>5</sup> Paul Klenerman,<sup>6</sup> Sergei Viazov,<sup>3</sup> Michael Roggendorf,<sup>3</sup> Robert Thimme,<sup>1</sup> Markus Reiser,<sup>7</sup> and Jörg Timm<sup>3\*</sup>

*Department of Medicine II, University of Freiburg, Hugstetter Str. 55, 79106 Freiburg, Germany<sup>1</sup>; Department of Biochemistry and Molecular Biology, New York Medical College, Valhalla, New York<sup>2</sup>; Department of Virology, University of Essen, Hufelandstr. 55, 45122 Essen, Germany<sup>3</sup>; Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom<sup>4</sup>; Department of Transfusion Medicine, University of Hamburg, Martinistr. 52, 20246 Hamburg, Germany<sup>5</sup>; Nuffield Department of Medicine, South Parks Road, Oxford OX1 3SY, United Kingdom<sup>6</sup>; and Paracelsus Hospital Marl, Department of Medicine, Lipper Weg 11, 45770 Marl, Germany<sup>7</sup>*

Received 6 August 2007/Accepted 12 January 2008

**Failure of the adaptive immune response to control infection with the hepatitis C virus (HCV) can result from mutational escape in targeted T-cell epitopes. Recent studies suggest that T-cell immune pressure is an important factor in the evolution of the nonstructural proteins in HCV. The aim of this study was to characterize the forces that contribute to viral evolution in an HLA-A\*01-restricted epitope in HCV NS3. This epitope represents a potentially attractive target for vaccination strategies since it is conserved across all genotypes. In our cohort of subjects with chronic HCV infection (genotype 1b or 3a), it is a frequently recognized CD8 epitope in HLA-A\*01-positive subjects. Viral sequence data reveal that an escape variant is the dominant residue in both genotypes. The predominant Y1444F substitution seemingly impairs binding to the HLA-A\*01 molecule, which may have an important impact on the ability to prime a functional CD8 response upon infection. Interestingly, a case of evolution toward the prototype sequence was observed during chronic infection, possibly because the helicase activity of the protein containing the Y1444F substitution is reduced compared to the prototype sequence. Comparison of HCV sequences from Asia and Europe suggests that the frequency of the HLA-A\*01 allele in a population may influence the frequency of the escape variant in circulating strains. These data suggest a complex interaction of multiple forces shaping the evolution of HCV in which immune pressure both within the individual and also at the population level in addition to functional constraints are important contributing factors.**

A major characteristic of hepatitis C virus (HCV) is the tendency to set up persistent infection. An important role of the cellular immune response for viral control has been highlighted (reviewed in reference 6). Failure of HCV-specific T cells can result from mutational escape in targeted T-cell epitopes (10). Several studies have demonstrated selection of escape mutations during acute HCV infection (18, 39, 41) and suggested T-cell immune pressure as an important driving force for the evolution of the nonstructural proteins in HCV (8). Moreover, population-based studies revealed that at least in some CD8 epitopes selection of escape mutations is reproducible in subjects sharing the same restricting HLA alleles (16, 34, 35, 42). However, functional constraints on the encoded proteins clearly limit the degree of variation of the viral genome (33, 36). Accordingly, reversion of CD8 escape mutations back to the prototype sequence upon transmission to a host who does not express the restricting allele has been demonstrated in human immunodeficiency virus (HIV) and HCV (1, 14, 27, 35, 41). In contrast, some escape mutations can also be stable upon transmission, either if the associated fitness

costs are low or if subsequent compensatory evolution occurs. For HIV it has been described that these stable mutations potentially persist in circulating viruses in a population to an extent where the escape variant appears as the predominant consensus sequence (3, 26).

For the design of effective immunotherapeutic strategies it is important to identify attractive immune targets in the pathogen. HCV epitopes that are conserved across genotypes are particularly valuable because they potentially provide protection against different genotypes. The HLA-A\*01-restricted epitope NS3 1436-1444 (ATDALMTGY) in the helicase portion of NS3 is an attractive CD8 target, since it is conserved across all genotypes. The epitope contains the highly conserved motif V of the HCV helicase domain 2 (reviewed in reference 12). The prototype was defined with a tyrosine in position 9 of the epitope (25)-however, mutational escape in this position to phenylalanine (position 1444 in the polyprotein as aligned to H77) has been documented in a subject with acute HCV genotype 1a infection (8). Moreover, the Y1444F substitution has been observed in association with the HLA-A\*01 allele in subjects with chronic infection (16, 35, 42), indicating that mutational escape is reproducible, even though various frequencies of this substitution have been reported.

The aim of the present study was to characterize the factors influencing the frequency of the Y1444F escape mutation in a population. We analyze here this epitope in two cohorts with

\* Corresponding author. Mailing address: University of Essen, Department of Virology, Hufelandstrasse 55, 45122 Essen, Germany. Phone: 49 201 723 2306. Fax: 49 201 723 5929. E-mail: joerg.timm@uni-due.de.

<sup>∇</sup> Published ahead of print on 23 January 2008.

chronic HCV genotype 1b or 3a infection. We show that this epitope is a frequent target of the CD8 T-cell response in HLA-A\*01-positive subjects infected with HCV. Viral sequence data reveal that the putative escape variant is the dominant residue in both genotypes studied in Western populations, which are rich in HLA-A\*01. This indicates that reproducible immune pressure on this epitope may drive continuous selection and potentially accumulation of the escape variant in populations with a high frequency of HLA-A\*01. We demonstrate that the selected Y1444F substitution impairs binding to the HLA-A\*01 molecule, which may have an important impact on the ability to prime a functional CD8 response against this epitope upon infection.

## MATERIALS AND METHODS

**Subjects.** Subjects with chronic HCV infection were recruited from the hepatology outpatient clinics of the University of Bochum and the University of Freiburg. Additional samples were retrieved from the serum bank of the German Hepatitis Network (Kompetenznetz Hepatitis). Subjects were included if they presented with a positive HCV-RNA test in serum and no clinical evidence for acute infection. Only subjects infected with genotypes 1b and 3a were chosen. The study was approved by the local Institutional Review Board, and all subjects gave written informed consent.

**HLA typing.** HLA-A and -B alleles were serotyped (85 subjects) or genotyped (12 subjects). The analysis was performed by the blood bank service of the University Hospital Freiburg or the Department of Transfusion Medicine at the University of Hamburg.

**Polyclonal expansion of HCV-specific T cells.** A total of  $4 \times 10^6$  peripheral blood mononuclear cells (PBMC) were resuspended in 1 ml of complete medium (RPMI 1640 containing 10% fetal calf serum, 1% streptomycin-penicillin, and 1.5% HEPES buffer [1 mol/liter]) and stimulated with peptide (10  $\mu$ g/ml) and anti-CD28 (0.5  $\mu$ g/ml; BD Pharmingen). On day 3 and 10, 1 ml of complete medium and recombinant interleukin-2 (IL-2; 20 U/ml; Hoffmann-La Roche) were added. On day 7, the cultures were restimulated with the corresponding peptide (10  $\mu$ g/ml) and  $10^6$  irradiated autologous feeder cells. On day 14, the cells were tested for gamma interferon (IFN- $\gamma$ ) secretion after 5 h of stimulation with the corresponding synthetic peptides.

**Intracellular cytokine staining (ICS).** Procedures were performed as described previously (40). Briefly, cells were stimulated with peptides (10  $\mu$ g/ml) or peptide loaded HLA-A\*01+ allogeneic Epstein-Barr virus (EBV) immortalized B-cell lines (B-LCL) in the presence of 50 U of recombinant IL-2/ml and 1  $\mu$ l of brefeldin A (BD Pharmingen)/ml. For comparison of peptides, increasing concentrations of the prototype and variant peptides were used. After 5 h of incubation (37°C, 5% CO<sub>2</sub>), cells were blocked with immunoglobulin G1 antibodies and stained with antibodies to CD8. After permeabilization with Cytofix/Cytoperm (BD Pharmingen) the cells were stained with antibodies to IFN- $\gamma$  (BD Pharmingen) and fixed in 100  $\mu$ l of CellFIX (BD Pharmingen) per well before fluorescence-activated cell sorting analysis.

**Amplification, sequencing, and cloning of HCV RNA.** Viral RNA was extracted from plasma samples by using a Qiagen (Hilden, Germany) vRNA extraction kit and amplified with genotype-specific primers in a nested PCR. Specific primers were designed for genotypes 1b and 3a based on alignments of available sequences from a public HCV database (<http://hcv.lanl.gov>) (21). In a combined reverse transcription (RT) and first-round PCR step, a 1,361-bp fragment for genotype 1b (primers HCV1b 4a-F 5'-ATGGAAGTACTATGCGG, nucleotides [nt] 3942 to 3959 as aligned to H77, and HCV1b 4d-R, 5'-CCAGG TGCTVGTGACGACC, nt 5303 to 5321) and a 1,268-bp fragment for genotype 3a (primers HCV3a 4044-F, 5'-CATGCCCCGACCGCAGCG, nt 4026 to 4044, and HCV3a 5314-R, 5'-GCTGACATGCATGCCATG, nt 5276 to 5293) were amplified. Nested primers amplifying an internal fragment were used in a second-round PCR (for genotype 1b, HCV1b-4b-F, 5'-AAGGACCATCACCA CGGG, nt 4187 to 4204, and HCV1b-4c-R, 5'-GGTGTATTTAGGTAAGC CCG, nt 4959 to 4978, and for genotype 3a, HCV3a-4420-F, 5'-CCAGGCAGC ATCACKGTGCC, nt 4395 to 4414, and HCV3a-4902-R, 5'-CCCAGCTCCTA GCACTCAC, nt 4894 to 4912). Using the Qiagen One-Step RT-PCR kit, the RT-PCR cycling conditions were as follows: 50°C for 45 min and 95°C for 15 min, followed by 35 cycles of 60 s at 94°C, 60 s at 54°C, and 90 s at 72°C, with a final extension of 72°C for 10 min. Nested PCR conditions consisted of 35 cycles of 60 s at 94°C, 60 s at 58°C, and 120 s at 72°C, with a final extension of 72°C for 10

min (GoTaq; Promega). PCR products were then extracted from a 1% agarose gel (Qiagen gel extraction kit) or directly purified utilizing a QIAquick PCR purification kit (Qiagen) and population sequenced on an ABI 3730 XL automated sequencer. Some PCR products were also cloned (TOPO TA; Invitrogen) and sequenced. Sequences were aligned and edited by using CodonCode Aligner (Dedham, MA) and Se-Al (<http://evolve.zoo.ox.ac.uk>). Sequence data are available at GenBank under accession numbers EU078744 to EU078840.

**Phylogenetic analysis.** All available reference sequences >300 bp in length which spanned the sequenced region were identified and downloaded from the HCV sequence database (21). Sequences with known countries of origin were retained. All sample sequences and reference strains were aligned by hand using Se-Al. Phylogenies were estimated by using the maximum-likelihood approach implemented in PAUP\*4 (38). The HKY model (19) with gamma-distributed among-site rate heterogeneity was used. A heuristic search for the maximum-likelihood tree was performed by using the tree-branch-reconnection and nearest-neighbor-interchange algorithms. The resulting estimated trees were midpoint rooted, and branches were colored according to the F/Y polymorphism in the epitope using FigTree 1.0 (<http://evolve.zoo.ox.ac.uk>). A neighbor-joining bootstrap analysis was performed with distances calculated by using maximum likelihood under the "HKY+gamma" model, as described above. Bootstrap scores of >60% are shown in boxes on the appropriate branches.

**Expression of the HCV genotype 3a helicase, mutagenesis, and purification.** HCV cDNA derived from patient 3a-Ch7 (accession no. EU287946) was amplified in PCRs using the primers Nhe-3A (5'-GCG CGC GCT AGC CTA CAA TTC ATA CCA GTG GAA ACC C-3') and Bam-3A (5'-GCG CGC GGA TCC CCA AGC AAC ACC CAG GTG CTG GTG GTC-3'). The PCR product was digested with BamHI and NheI and ligated into a similarly treated pET24 vector (Novagen). The resulting plasmid (p24Hel-3a Y1444F) contained the Y1444F escape variant. To generate the ATDALMTGY prototype p24Hel-3a, Y1444F was subjected to site-directed mutagenesis using a QuikChange kit (Stratagene) and the primers F1444Y(+) (5'-GCG CCT CAT GAG TGG ATA CAC CGG AGA CTT CGA CTC-3') and F1444Y(-) (5'-GAG TCC AAG TCT CCG GTG TAT CCA GTC ATG AGG GCG-3'). The sequence of the resulting plasmid (p24Hel-3a proto) was confirmed, and both p24Hel-3a proto and p24Hel-3a Y1444F were used to transform Rosetta (DE3) cells (Novagen) and to express a 53-kDa protein containing NS3 amino acids 116 to 631 (the complete HCV helicase) flanked by an N-terminal MAS sequence and a C-terminal STWVLL GDPNSSSVDKLAALAEHHHHHHH sequence (His tag).

Proteins were purified after suspension in 5 volumes of buffer A (20 mM Tris [pH 8], 0.5 M NaCl, 5 mM imidazole), sonicated, and centrifuged. The lysate was applied to a 5 ml of nickel-NTA column (Novagen) and eluted with a linear imidazole gradient. Fractions containing the helicase were precipitated with 60% ammonium sulfate. The resulting pellet was dissolved in buffer B (20 mM Tris [pH 8], 50 mM NaCl, 1 mM EDTA, 0.1 mM dithiothreitol) and subjected to gel filtration chromatography, followed immediately by ion-exchange chromatography using DEAE-Sepharose (Pharmacia). Protein concentrations were determined from calculated extinction coefficients at  $A_{280}$ .

The full-length NS3 protein (NS3 [con1]) was isolated as an N-terminally His-tagged protein by amplifying cDNA from the plasmid pBDL429P+S+ (29) using the PCR primers NS3(+) (5'-GCG CGC GCT AGC GCG CCT ATT ACG GCC TAC) and NS3(-) (5'-GCG CGC GAA TTC GGT CAC GTG ACG ACC TCC AG). The PCR product was purified, digested with NheI and EcoRI (sites underlined), and ligated into a similarly treated pET28 vector (Novagen). The NS3 protein was expressed and purified as described previously for truncated NS3 except that the final DEAE column was substituted with a Fractogel EMD COO- column (EMD Biosciences).

The prototype Y1444 sequence was generated in a con1 background by subjecting plasmids expressing either Hel-1b (con1) or NS3 (con1) to site-directed mutagenesis using the primers F1444Y-Con1(+) (5'-GAC GCT CTA ATG ACG GGC TAT ACC GGC GAT TTC GAC-3') and F1444Y-Con1(-) (5'-GTC GAA ATC GCC GGT ATA GCC CGT CAT TAG ACG GTC-3'). The resulting proteins Hel-1b (F1444Y) and NS3 (F1444Y) were purified as described above.

**Helicase assay.** HCV helicase-catalyzed DNA unwinding was monitored by using a novel fluorescent "self-trapping" unwinding assay. This assay monitors the displacement of a dual-labeled DNA oligonucleotide from a longer complementary oligonucleotide. Upon displacement, both strands form hairpin structures, making the reaction essentially irreversible. Hairpin formation leads to a fluorescence decrease because a Cy5 fluorophore approaches an Iowa Black quencher moiety (Integrated DNA Technologies). To generate the substrate, 25  $\mu$ M concentrations of the labeled short strand (Cy5-GCT CCC CAA TCG ATG AAC GGG GAG C-IBQ) and long strand (3'-TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT CGA GGG GTT AGC TAC TTG CCC CTCG-5') were

TABLE 1. HLA-A1-positive subjects included for the analysis of ATDALMTGY<sub>1436-1444</sub>-specific T cells

Genotype	Subject	Age (yr)	Sex <sup>a</sup>	ALT (U/liter)	Viral load (IU/ml)	HLA alleles	% IFN- $\gamma$ <sup>+</sup> /CD8 <sup>+</sup> ATDALMTGY <sub>1436-1444</sub> <sup>b</sup>
1b	1b-Ch1	72	F	91	1,500,000	A1, A23, B35, B49	Neg <sup>d</sup>
	1b-Ch2	32	M	104	375,000	A1, A11, B7, B8	0.41
	1b-Ch3	38	M	68	1,140,000	A1, A11, B52, B61	6.62
	1b-Ch4	19	M	88	30,000	A1, A31, B60	Neg
	1b-Ch5	54	M	76	22,000	A1, B8	20.53
	1b-Ch6	41	M	55	>8,000,000	A1, B8, B17	0.59
	1b-Ch7	64	F	120	566,000	A1, A29, B8, B44	Neg
	1b-Ch8	40	M	238	340,000	A1, A3, B35, B63	8.58
	1b-Ch9	22	M	55	580,000	A1, A3, B8, B35	Neg
	1b-Ch10	30	M	197	2,800,000	A1, A24, B8, B35	Neg
3a	3a-Ch1	47	M	141	1,000,000	A1, A3, B7, B35	0.33
	3a-Ch2	18	M	450	2,000,000	A1, A3, B44	Neg
	3a-Ch3	37	F	58	61,000	A1, A28, B13, B51	Neg
	3a-Ch4	32	F	142	13,000	A1, A2, B7, B61	Neg
	3a-Ch5	41	F	139	188,000	A1, A26, B8, 35	1.23
	3a-Ch6	41	M	170	5,800	A1, A24, B7, B8	Neg
	3a-Ch7	41	F	61	220,000	A1, A3, B8, B44	6.37
	3a-Ch8	34	F	78	33,000	A1, A2, B7, B13	Neg
	3a-Ch9	48	M	24	3,080,000	A1, A32 <sup>c</sup>	Neg
	3a-Ch10	21	M	184	1,100,000	A1, A2, B8, B44	4.27

<sup>a</sup> F, female; M, male.

<sup>b</sup> After 14 days of in vitro stimulation.

<sup>c</sup> B alleles were not determined.

<sup>d</sup> Neg, negative.

heated to 95°C and then cooled to room temperature. The annealed substrate was diluted to 5 nM in reaction buffer (25 mM MOPS [pH 6.5], 2 mM MgCl<sub>2</sub>, 0.1% Tween 20), and fluorescence was monitored continually at 37°C. The helicase protein was then added (final concentration 5 to 100 nM) and equilibrated with the substrate for 2 min before initiation of the unwinding reaction by adding ATP (4 mM). Initial unwinding rates were measured by using linear regression over the initial phase of the reaction (0.5 to 2 min).

**ATPase assays.** The initial rates of HCV helicase-catalyzed ATP hydrolysis were measured by using a modified colorimetric "malachite green" assay. Reactions were performed with catalytic amounts of HCV helicase (5 to 50 nM) at 37°C in 25 mM morpholinepropanesulfonic acid (pH 7.5). After 10 to 20 min, 25  $\mu$ l of a reaction was added to 200  $\mu$ l of a freshly prepared color reagent (3 parts 0.045% malachite green, 1 part 4.2% ammonium molybdate in 4.0 N HCl, 0.01 part 10% Tween 20). Within 1 min, solutions were quenched with 25  $\mu$ l of a 34% sodium citrate and incubated for 20 min at room temperature, and the resulting absorbance was read at 625 nm. Phosphate concentrations were determined from a standard curve generated with the same batch of reagents.

## RESULTS

**CD8 epitope ATDALMTGY<sub>1436-1444</sub> is frequently targeted in subjects infected with HCV genotype 1b or 3a.** First, we determined the frequency of recognition of the HLA-A\*01-restricted epitope ATDALMTGY<sub>1436-1444</sub> during chronic HCV infection. PBMC from HLA-A\*01-positive subjects infected with genotype 1b or genotype 3a were stimulated with the synthetic peptide to expand antigen-specific T cells. In 5 of 10 subjects (50%) infected with genotype 1b and in 4 of 10 subjects (40%) infected with genotype 3a specific T cells targeting ATDALMTGY<sub>1436-1444</sub> were detectable (Table 1). In one subject (1b-Ch5) the number of specific T cells targeting this single epitope could be expanded up to 20% of all CD8 T cells. Of note, this response was detectable in only 3 of 10 subjects infected with genotype 1b before antigen-specific expansion (data not shown), suggesting that some low-level responses may be missed if only ex vivo techniques are used. In summary, 9 of 20 HLA-A\*01-positive subjects with HCV genotype 1b or

3a infection had a detectable CD8 response against the ATDALMTGY<sub>1436-1444</sub> epitope, indicating that it is a frequent target even during chronic infection.

**Impact of Y1444F variant on T-cell targeting.** Selection of the Y1444F substitution has been described in a subject during acute HCV genotype 1a infection (8). The Y1444F substitution in the HLA-A\*01 epitope is located in the anchor position for MHC class I binding (9), suggesting that the mutation from tyrosine to phenylalanine would affect peptide binding to the HLA-A\*01 molecule and, consequently, efficient targeting by specific T cells. Interestingly, in the previous escape study no differential impact on T-cell recognition between the prototype and the variant was observed when tested ex vivo or after antigen-specific expansion of T cells (8). To further evaluate this finding, we determined the impact of the Y1444F substitution on T-cell recognition in all nine subjects targeting this epitope. In line with the previous observation there was substantial cross-recognition between the prototype and the variant in most cases (Fig. 1). Only T cells expanded from subject 1b-Ch6 did not target the variant sequence. In three cases (1b-Ch2, 1b-Ch5, and 1b-Ch8) the secretion of IFN- $\gamma$  was diminished compared to the prototype. However, in all remaining subjects no difference between the prototype and the Y1444F variant was observed. Interestingly, when PBMC were cultivated in the presence of the variant peptide no or only a minor fraction of specific T cells were detectable against both the prototype sequence and the Y1444F variant. As a possible mechanism, we suggest that the Y1444F variant does not stimulate expansion of antigen-specific T cells against this epitope in vitro.

One possible explanation for this observation is an accelerated off-rate of the variant peptide from the major histocompatibility complex (MHC) molecule that is not unmasked in a

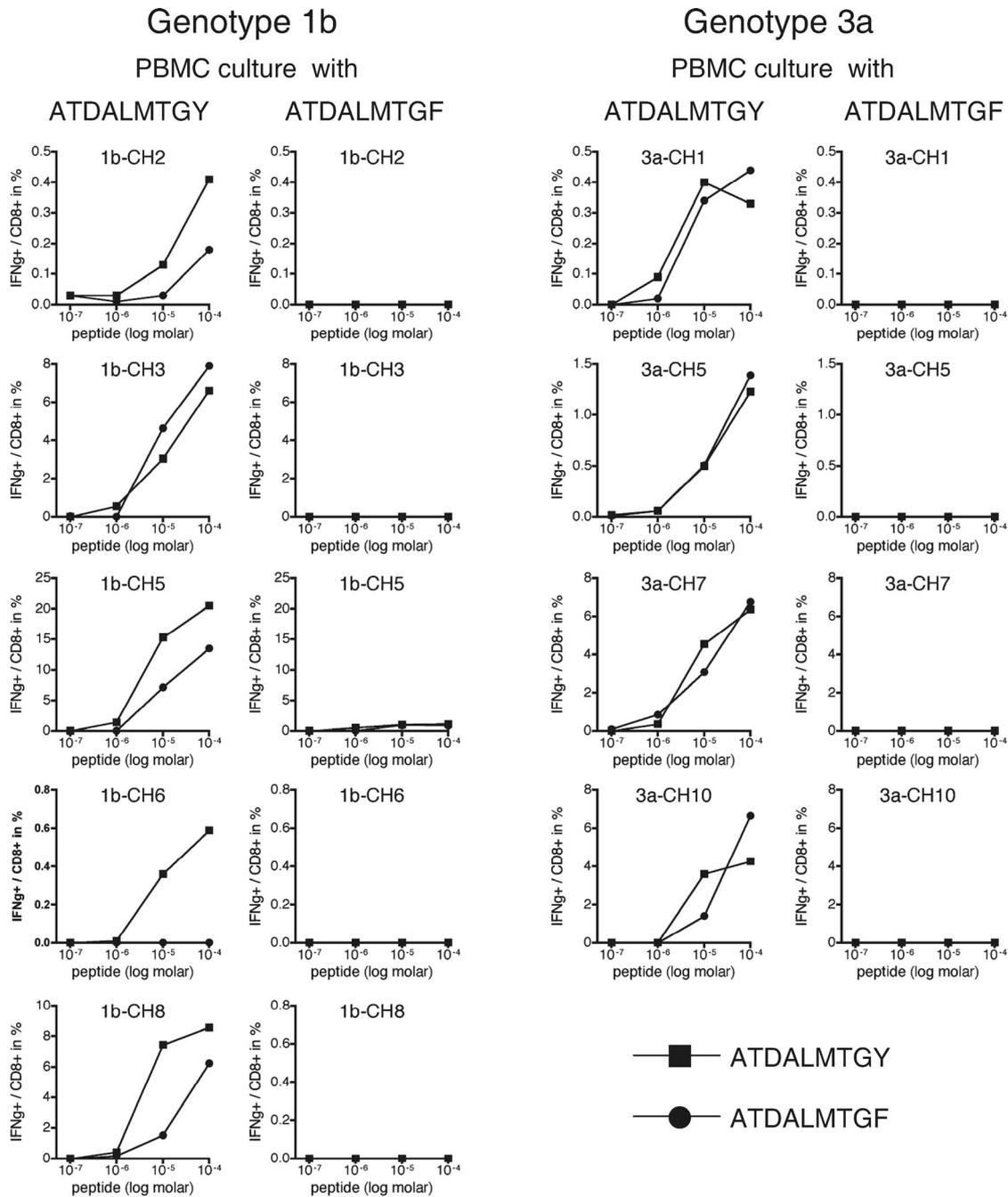


FIG. 1. Impact of variant peptides on IFN- $\gamma$  secretion. PBMC from HLA-A\*01-positive subjects infected with genotype 1b (left panels) or genotype 3a (right panels) were cultivated in the presence of the prototype peptide (ATDALMTGY) or the Y1444F variant (ATDALMTGF) as indicated. After 14 days cells were tested after 5 h of restimulation with the prototype peptide (■) and Y1444F variant (●) in serial dilutions in a standard ICS assay for IFN- $\gamma$ . The percentage of IFN- $\gamma$ -positive cells of all CD8-positive cells is shown.

standard ICS assay, where peptide is present continuously. To address this, we loaded the variant or prototype peptide on allogeneic HLA-A\*01-positive B cells, washed the cells extensively, and used them in an ICS assay as targets for specific T cells that have been expanded in the presence of the prototype (Fig. 2). B cells pulsed with the prototype peptide were efficiently targeted. In contrast, B cells pulsed with the variant peptide were not targeted after the washing step. In this case

the stability of the MHC-peptide complex may not have been sufficient, and the variant peptide was removed during the washing procedure. In summary, the Y1444F substitution represents a true escape mutation in the HLA-A\*01 restricted epitope that may act through a decrease in binding to the MHC class I molecule.

**Putative escape variant Y1444F is predominant in two cohorts infected with HCV genotype 1b or 3a.** To determine the

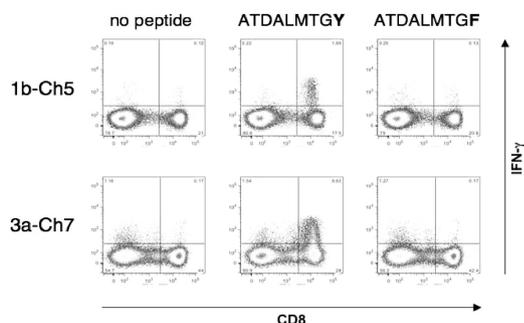


FIG. 2. Accelerated off-rate of the Y1444F variant. HLA-A\*01-positive allogeneic EBV immortalized B-cell lines were incubated without peptide (left panels) or in the presence of the prototype peptide (middle panels) or the Y1444F variant (right panel). After six washes with medium the B cells were used as targets for polyclonally expanded T cells (stimulated with ATDALMTGY) in an ICS assay for IFN- $\gamma$ . Dot plots from two subjects (1b-Ch5 and 3a-Ch7) are shown.

frequency of the Y1444F escape mutation in circulating viruses in our cohort, the region of NS3 containing this epitope was sequenced from 49 subjects infected with HCV genotype 1b and 48 subjects infected with genotype 3a (Fig. 3). In both cohorts the most common residue at position 1444 is the Y1444F substitution, and this putative escape variant represents the consensus sequence. The prototype sequence ATDALMTGY is only present in 9 of 49 (18.4%) subjects infected with HCV genotype 1b and 12 of 48 (25.0%) subjects infected with genotype 3a. All HLA-A\*01-positive subjects harbor a virus with the escape sequence (ATDALMTGF), including the subjects that were previously tested for the presence of the corresponding CD8 response (Table 1). Interestingly, the majority of HLA-A\*01-negative subjects are also infected with the escape variant: 22 of 31 (71.0%) of genotype 1b patients and 19 of 31 (61.3%) of genotype 3a patients. This indicates that the escape variant is relatively stable also in the absence of HLA-A\*01-associated immune pressure and persists in these populations. Of note, we tested PBMC from 10 HLA-A\*01-negative subjects before and after antigen-specific stimulation with the prototype, variant, or overlapping peptides covering residue 1444 and were not able to detect a response targeting this region in a single case (data not shown). Other epitopes overlapping with this residue have not been published. It therefore seems unlikely that a second overlapping epitope restricted by another allele accounts for this high frequency of the Y1444F substitution in HLA-A\*01-negative subjects. Despite this high frequency of the Y1444F substitution, the association between the escape sequence and presence of HLA-A\*01 remains statistically significant in both cohorts (Fig. 3, genotype 1b,  $P = 0.0098$ ; genotype 3a,  $P = 0.0020$  [Fisher exact  $t$  test]).

To determine whether minor frequencies of sequence variations at this site were missed in the quasispecies population of individual patients, the PCR product containing the ATDALMTGY<sub>1436-1444</sub> epitope from a total of eight subjects was cloned. The analysis includes four HLA-A\*01-positive subjects (two with a strong and two without a detectable CD8 response) and four HLA-A\*01-negative subjects (two containing the prototype sequence and two containing the Y1444F substitution). A total of 84 clones (8 to 12 clones per subject) were sequenced. In all eight subjects no sequence variation was iden-

tified in position 1444 in the quasispecies population (data not shown), suggesting that this residue had achieved fixation within the virus populations infecting these eight individuals.

Next we investigated whether longitudinal analysis of HCV sequences reveals ongoing evolution in the ATDALMTGY<sub>Y1436-1444</sub> epitope during chronic infection. Follow-up samples were therefore collected and sequenced. Most subjects received treatment after their first presentation and were therefore excluded from longitudinal analysis. Follow-up samples from four untreated HLA-A\*01-negative subjects between 32 and 78 weeks after the first time point were available. No sequence evolution was observed in two subjects (Fig. 4). A third subject infected with genotype 1b showed one variant (G1443D) in the epitope in 1 of 11 clones; however, position 1444 was fixed and did not evolve between these two time points. A fourth subject, 3a-Ch37, harbored a virus with the Y1444F substitution at the first time point. Interestingly, sequence analysis at follow-up 40 weeks later revealed a mixed base in the bulk sequence coding for both the prototype sequence and the Y1444F polymorphism, suggesting that either neutral evolution or selectively driven reversion of the Y1444F site was under way. Analysis of the clonal sequences confirmed the presence of a mixed quasispecies population with 7 of 12 clones harboring the prototype sequence in the second time point (Fig. 4). Even though the prototype sequence might have been present in a very low frequency at the previous time point the shift in the clonal frequency indicates that evolution toward the prototype sequence can occur in genotype 3a.

**Y1444F substitution is associated with decreased NS3 helicase function.** The Y1444F substitution is located in the helicase domain of the NS3 protein. Since the amino acid substitution is near the presumed ATP and RNA binding sites of the protein, such genetic variation could impair its functional performance. To test this hypothesis, we cloned NS3/4A from patient 3a-Ch7 infected with HCV genotype 3a and subjected it to site-directed mutagenesis to generate both variants. The helicase domain of NS3 containing either the prototype (Hel-3a proto) or the Y1444F substitution (Hel-3a Y1444F) was then expressed as a recombinant protein and tested in a DNA unwinding assay, which monitors the displacement of a dual-labeled DNA oligonucleotide from a longer complementary oligonucleotide. Upon displacement both strands form hairpin structures, making the reaction essentially irreversible. Hairpin formation is monitored by a fluorescence decrease of a Cy5 fluorophore (Fig. 5A). The unwinding rates were compared to rates obtained using similar HCV helicase proteins from other genotypes (22). An alignment of the proteins used, showing the region surrounding the ATDALMTGY epitope, is shown in Fig. 5B. Notably, helicase containing the escape mutant (Hel-3a Y1444F) unwound the DNA substrate about five times slower than the prototype sequence (Hel-3a proto), suggesting that the Y1444F variant results in less active helicase (Fig. 5C). This difference in helicase activity was reproducible at different concentrations. At each concentration Hel-3a proto unwound DNA faster than Hel-3a Y1444F (Fig. 5D). To exclude that this observed difference is due to variations in the amount of active protein in either preparation the HCV helicase catalyzed ATP hydrolysis was determined for both helicases under optimal conditions. The maximum rate of ATP hydrolysis divided by the total protein concentration revealed

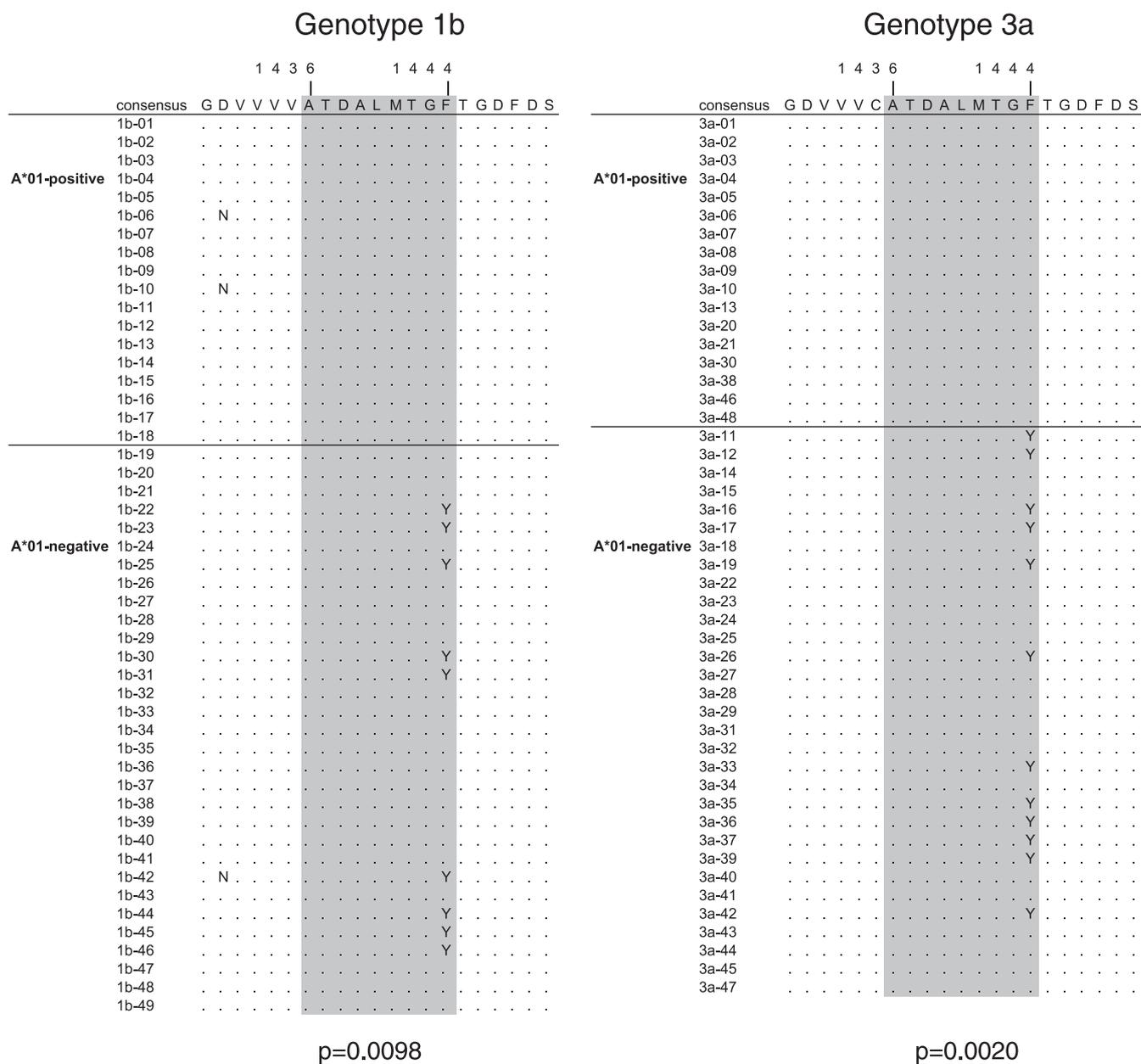


FIG. 3. Amino acid alignment of the ATDALMTGY<sub>1436-1444</sub> epitope. Sequences spanning the ATDALMTGY<sub>1436-1444</sub> epitope (highlighted in gray) from subjects with chronic HCV genotype 1b or 3a infection are aligned to the consensus sequence for each genotype. Sequences derived from HLA-A\*01-positive subjects are shown above the line, while those from HLA-A\*01-negative subjects are shown below the line. Note that the escape variant (Y1444F) with Phe in position 1444 (as aligned to H77) is the consensus residue in both genotypes.

that Hel-3a proto was slightly more active than Hel-3a Y1444F ( $k_{cat} = 37 \pm 13$  versus  $25 \pm 15$ ). However, this difference was not statistically significant. When ATP hydrolysis was measured at various substrate concentrations, nearly identical results were obtained for Hel-3a proto and Hel-3a Y1444F ( $K_m = 188 \pm 22$  versus  $193 \pm 26$ ). Thus, both the turnover rate ( $k_{cat}$ ) and the catalytic efficiency ( $k_{cat}/K_m$ ) of HCV helicase-catalyzed ATP hydrolysis appeared to be unchanged by the escape mutation.

To examine the impact of the escape mutation in the genotype 1b background, the plasmid encoding Hel-1b (con1), which contains the escape mutation Y1444F, was subjected to

site-directed mutagenesis to generate the prototype sequence. The resulting protein (Hel-1b [F1444Y]) was then compared to the other helicases studied here (Fig. 5C and D). The con1 helicase with the prototype sequence (Hel-1b [F1444Y]) was slightly more active than the helicase with the escape mutation (Hel-1b [con1]). However, the difference in genotype 1b was minor compared to genotype 3a. Hel-3a (proto) has a specific activity (i.e., nM DNA unwound/min/nM protein) five times higher than that of the escape mutant Hel-3a (Y1444F), but Hel-1b (F1444Y) is only ca. 20% (1.2-fold) more active than the same protein with the escape mutation (Hel-1b [con1]). Similar differences between the proteins were observed when

		1 4 3 0										1 4 4 0										1 4 5 0												
		P T T G D V V V C										A T D A L M T G F										T G D F D S V I D												
<b>bulk</b>	ID	GT3a cons																														viral load		
	3a-Ch19	T1	. A . . . . .																														1,029 kIU/ml	
		week 54	. . . . .																														1,238 kIU/ml	
	3a-Ch22	T1	. . . . .																														425 kIU/ml	
		week78	. . . . .																														ND	
	3a-Ch37	T1	. . S . . . . .																														617 kIU/ml	
		week 40	. . S . . . . .																														718 kIU/ml	
		GT1b cons																																
	1b-CH37	T1	. A . . . . .																														ND	
		week 32	. A . . . . .																														100 kIU/ml	
<b>clones</b>	ID	GT3a cons																														# clones		
	3a-Ch37	T1	. . S . . . . .																														10/10	
		3a-Ch37	week 40	. . S . . . . .																														7/12
			week 40	. . S . . . . .																														5/12
			GT1b cons																															
		1b-CH37	T1	. A . . . . .																														10/10
			week 32	. A . . . . .																														9/11
		week 32	. A . . . . .																														1/11	
		week 32	. A . . . . .																														1/11	

FIG. 4. Evolution of the ATDALMTGY<sub>1436-1444</sub> epitope during chronic HCV infection. Sequences spanning the ATDALMTGY<sub>1436-1444</sub> epitope (highlighted in gray) from subjects with chronic HCV infection obtained at the first available time point (T1) and from a second time point (between 32 and 78 weeks later) are aligned to the consensus sequence. In the upper panel, bulk sequences from subjects 3a-Ch19, 3a-Ch22, 3a-Ch37, and 1b-Ch37 are shown. In the lower panel, clonal sequences from 3a-Ch-37 and 1b-Ch37 are shown, confirming a mutation back to the prototype in HCV genotype 3a. The viral load at each time point and the frequency of each clone is indicated.

they were assayed for their ability to unwind RNA (data not shown).

Since it has been established that the protease domain of NS3 plays a role in helicase action (13), we also examined the impact of the escape mutation on the helicase activity of the full-length NS3 protein. To this end, we subjected a plasmid encoding a full-length NS3 protein from the genotype 1b (con1) strain to mutagenesis. The resulting substitution (F1444Y) again reverted the escape mutation in the con1 background. The results with this protein were similar to those with the truncated NS3 protein (Fig. 5C). The full-length protein is about 2.5 times more active than truncated NS3 in this assay, and again the prototype sequence (NS3 [F1444Y]) is about 1.2 times more active than the escape mutant (NS3 [con1]). Again, a similar difference was observed when RNA was used as a substrate (not shown).

**Phylogenetic analysis reveals numerous unambiguous changes between Tyr and Phe at position 1444.** Recent reports suggest that the ability to detect HLA-associated sequence polymorphisms in HIV and HCV may be reduced by so-called “founder effects” in the phylogeny (5, 42). Sequences with a particular polymorphism may be closely related and form a distinct phylogenetic cluster. In this case, it is argued that it is less likely that the escape variant evolved by immune pressure independently within each individual; instead, it is considered more likely that a limited number of evolutionary events occurred in the founders of the lineage or cluster in question.

In order to address this in our data set a maximum-likelihood phylogeny was calculated for all genotype 1b and 3a sequences, upon which was mapped all unambiguous changes between Tyr and Phe in position 1444. Figure 6 shows the tree for all genotype 3a sequences; also included are all available

genotype 3a reference sequences from the public database. The branches of the tree are color-coded; blue branches represent the prototype sequence, and red branches represent the Y1444F substitution. Importantly, 10 different internal nodes in the tree split into a blue and a red branch (labeled with a black dot), indicating unambiguous and independent changes between Tyr and Phe in these cases. However, it is difficult to infer the direction of these changes from the tree, and not all such nodes are supported by high bootstrap values. Both changes from Tyr to Phe and from Phe to Tyr are possible. In a previous phylogenetic analysis approach (42), the ancestral state of the residue was reconstructed and used as a basis to determine the direction of sequence changes. However, the bootstrap values of the present maximum-likelihood tree were too low to do a similar analysis with any level of statistical confidence. Statistical variance arising from phylogenetic estimation also explains why a few sequences are associated with long terminal branches, since no alignment errors or frame-shifts were found in these sequences.

The same analysis for genotype 1b is more complex. Figure 7 depicts the maximum-likelihood tree of all genotype 1b sequences from our cohort and all available reference sequences from the public database. Only sequences for which the origin of the isolate was known were added. The branches in the tree are again color-coded to distinguish the prototype sequence (blue) and the Y1444F variant (red). A total of 16 unambiguous changes between Tyr and Phe at position 1444 were identified (nodes labeled with a black dot), including four cases with sequences from our cohort, again suggesting multiple mutational events at this site in genotype 1b. One of these sequence changes that is supported with a high bootstrap value (98%; indicated by an arrow in the figure) includes a prototype



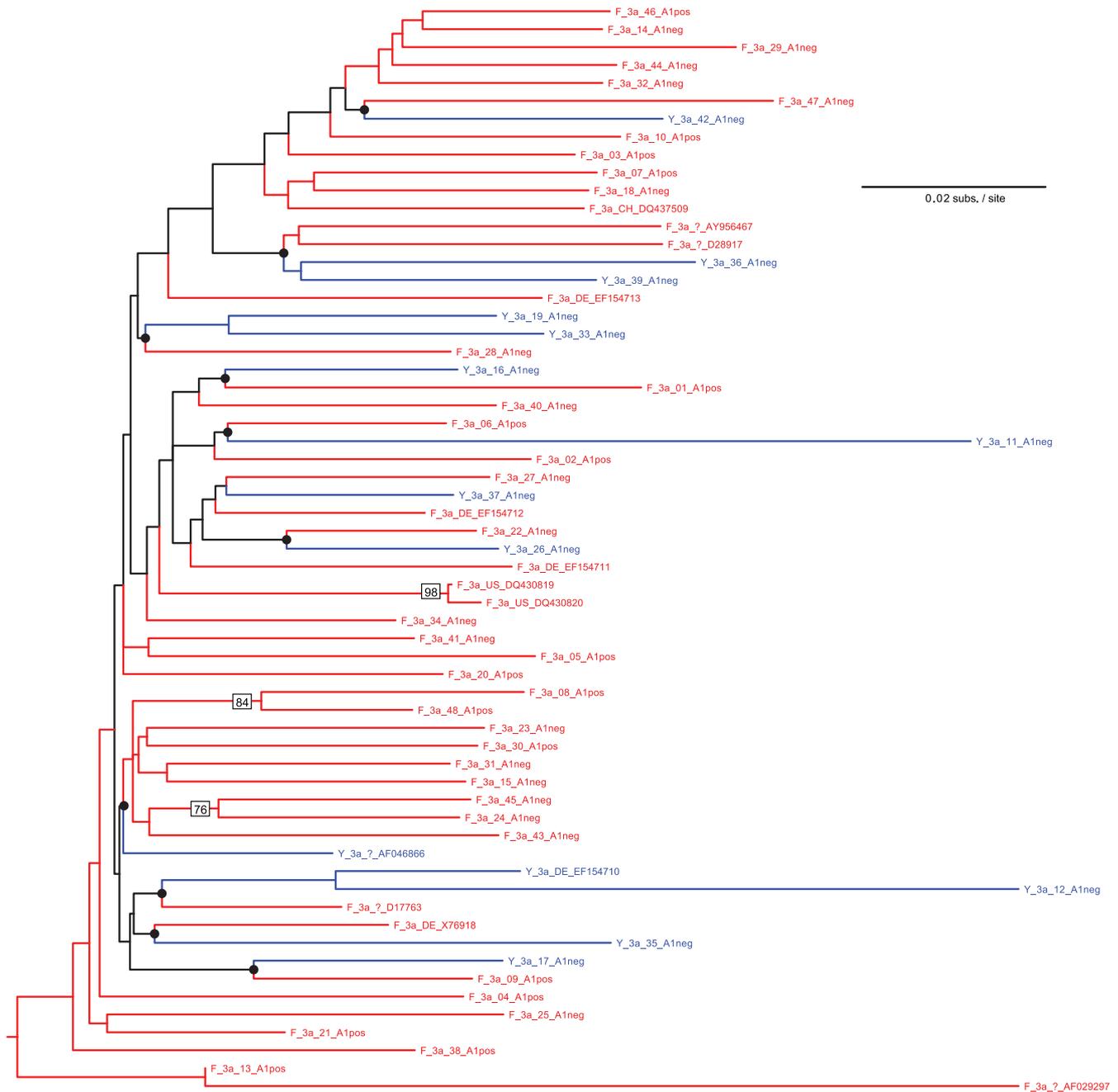
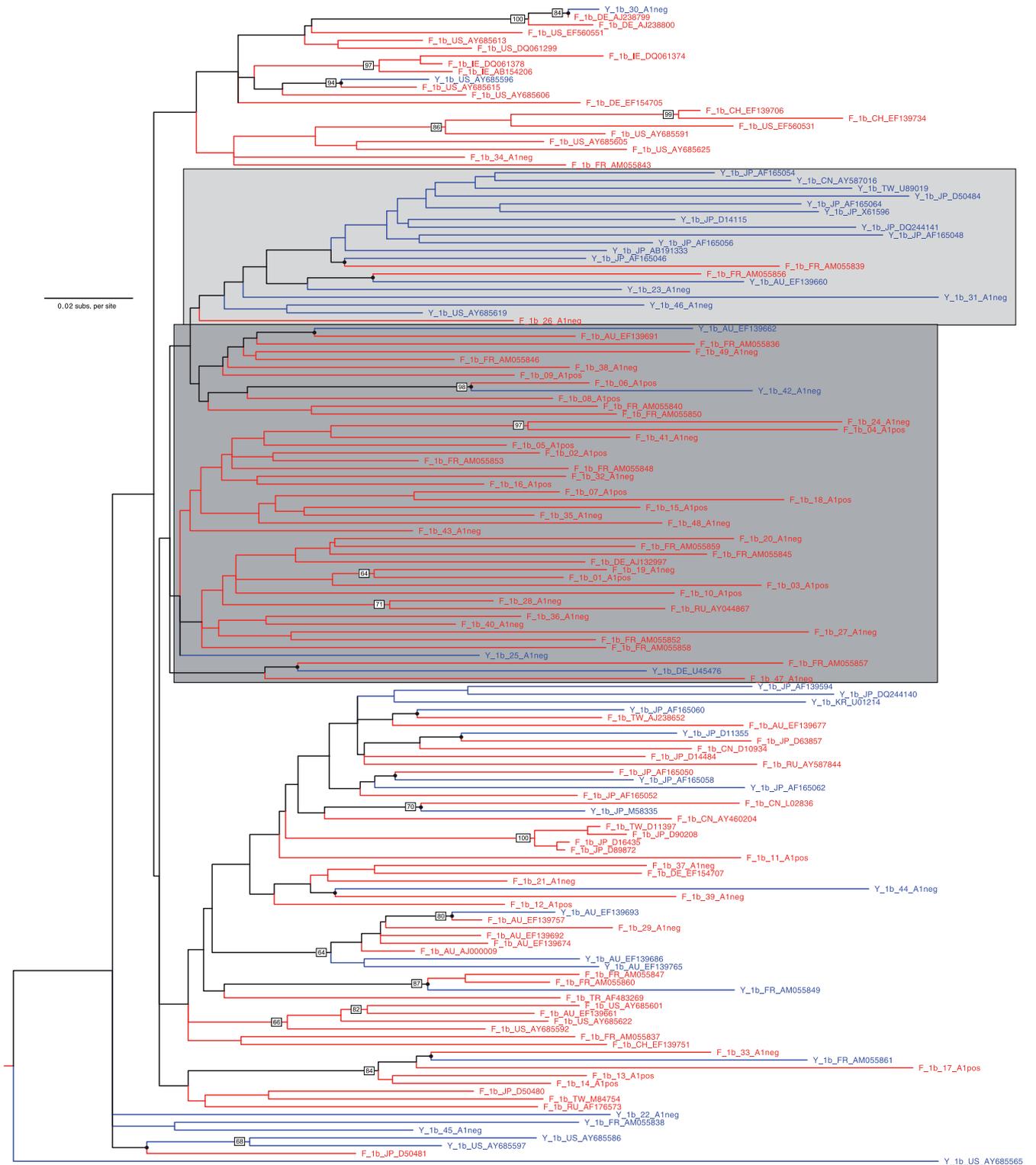


FIG. 6. Maximum-likelihood phylogeny for HCV genotype 3a sequences. A maximum-likelihood phylogeny was calculated for the genotype 3a sequences from our cohort and all available reference sequences from the HCV sequence database >300 bp in length that span the sequenced region. The accession numbers and countries of origin of the reference sequences are indicated in the branch labels. The branches of the tree are color-coded: blue branches represent the prototype sequence, and red branches represent the Y1444F substitution. Ten different internal nodes in the tree split into a blue and a red branch (labeled with a black dot), indicating unambiguous and independent changes between Tyr and Phe. Bootstrap scores of >60% are shown in boxes on the appropriate branches.

sequences from our cohort (boxed in dark gray). This raised the possibility that there are region-specific differences in the frequency of the prototype and the Y1444F substitution. A total of 149 sequences are included in the tree: 49 sequences are from our cohort recruited in Germany (Table 2), an additional 35 previously published sequences are from Europe (France, Germany, Russia, Ireland, and Switzerland), 35 sequences are of Asian origin (Japan, China, Taiwan, and Ko-

rea), 17 sequences are from the United States, 12 sequences are from Australia, and 1 more is from Turkey. The frequency of the Y1444F substitution is regionally different, with the highest frequency in published sequences from Europe (88.6%) and the lowest frequency in isolates from Asian countries (42.9%). These region-specific differences in the distribution of the prototype and the Y1444F variant correlate with the frequency of the HLA-A\*01-allele in different populations.



Downloaded from jvi.asm.org at Yale University on March 18, 2008

FIG. 7. Maximum-likelihood phylogeny for HCV genotype 1b sequences. A maximum-likelihood phylogeny was calculated for the genotype 1b sequences from our cohort and all available reference sequences from the HCV sequence database >300 bp in length that span the sequenced region. The accession numbers and countries of origin of the reference sequences are indicated in the branch labels. The branches of the tree are color-coded: blue branches represent the prototype sequence, and red branches represent the Y1444F substitution. Sixteen different internal nodes in the tree split into a blue and a red branch (labeled with a black dot), indicating unambiguous and independent changes between Tyr and Phe. A case of an unambiguous change that includes two sequences from our cohort is indicated with an arrow. The tree also contains unique clusters with predominantly prototype sequences (boxed in light gray) or predominantly the Y1444F variant (boxed in dark gray). Bootstrap scores of >60% are shown in boxes on the appropriate branches.

TABLE 2. Regional differences in the frequency of the Y1444F substitution

Region	No. of sequences	Y/F	% F	Consensus
Germany	49	9/40	81.6	F
Europe	35	4/31	88.6	F
Asia	35	20/15	42.9	Y
United States	17	5/12	70.6	F
Australia	12	5/7	58.3	F
Remaining regions	1	0/1		

HLA-A\*01 is a common allele in Europe, with phenotype frequencies between 20 and 35% (Fig. 8). In contrast, HLA-A\*01 is rare in Asians, with reported frequencies of <6%. Interestingly, only in isolates of Asian origin the consensus residue in position 1444 is the prototype residue tyrosine (Table 2), suggesting that isolates from this region are potentially less well adapted to HLA-A\*01-mediated selection pressure.

### DISCUSSION

The propensity of HCV to evolve with a high mutation rate, enabling selection of escape variants in targeted regions, makes the design of effective immune therapies difficult. Here we describe the impact of MHC class I-associated selection pressure on the frequency of an escape mutation in an HLA-A\*01-restricted CD8 epitope in NS3. Continuous HLA-A\*01-mediated selection pressures in two cohorts infected with HCV genotype 1b or 3a replace the prototype epitope sequence (ATDALMTGY) with an escape variant (ATDALMTGF) that seems to impair MHC class I binding.

Recent studies in HIV-1 suggest that escape variants can become predominant in circulating strains in a population (3, 26, 31). In an analysis of HLA class I-associated sequence polymorphisms in viral genomes, these escape mutations represent “negative associations” (16, 31, 42), which means that a particular HLA allele is associated with presence of the consensus sequence. As a driving mechanism for these negative associations, reproducible selection in the presence of CD8 immune pressure, stability of the mutation upon transmission, and the lack of reversion in the absence of immune pressure have been proposed (26). Under these conditions continuous accumulation of the escape variant ultimately replaces the “wild-type” sequence as the most frequent strain in a population, and the escape residue becomes the predominant consensus residue. It has been highlighted that this mechanism may potentially lead to complete deletion of epitopes from the population after sufficient time has elapsed and all circulating “wild type” sequences are replaced (26). Here we similarly demonstrate for HCV that an escape mutation in an HLA-A\*01-restricted epitope in HCV NS3 is predominant in circulating isolates. Mutational escape during acute infection has been described for this epitope in a subject infected with genotype 1a (8). Moreover, significant associations between the Y1444F substitution and expression of the HLA-A\*01 allele in larger cohorts with chronic infection have been previously reported (16, 42). However, different frequencies of this escape mutation have been observed. The predominant residue in a cohort infected with genotype 1a in North America was the

prototype, whereas in a mixed cohort from Switzerland and Western Australia infected with genotype 1a or 1b the escape mutation Y1444F was predominant, indicating regional differences and potentially differences between genotypes or subtypes.

Even though the Y1444F substitution in the HLA-A\*01 epitope is located in the anchor position for MHC class I binding, Cox et al. (8) observed in a previous escape study no differential impact on stimulation of IFN- $\gamma$  secretion both *ex vivo* and using expanded specific T cells from a subject with acute HCV infection. The hydrophilic tyrosine is the common residue for the anchor in position 9 of the HLA-A\*01 motif (9). Substitution with the similar but hydrophobic phenylalanine at this crucial site likely preserves the peptide’s structure but impairs binding to the MHC molecule. This is supported by another study suggesting that the tyrosine-to-phenylalanine substitution in this anchor position reduces HLA-A\*01 binding (20) and is in line with a previously reported >200-fold-reduced binding affinity of the Y1444F variant to the HLA-A\*01 molecule compared to the prototype (8). We observed substantial cross-recognition between the two variants when added in an ICS assay, where the peptide is present throughout the whole incubation period. However, when loaded on B cells that are washed after the incubation period, the presentation of the variant peptide is significantly impaired. This suggests that MHC class I binding is likely affected, resulting in a higher drop-off rate of the variant peptide. A similar subtle effect of an escape mutation that only unmasked when the off-rate from pulsed B cells was determined has been previously described for HIV (17, 26).

Accumulation of escape mutations in circulating strains may have important implications for the ability to prime a functional antiviral CD8 immune response upon acute infection. Even though the plasticity of the T-cell response allows priming of CD8 T cells against various sequences including escape variants (2, 7), mutations of the anchor residue for MHC class

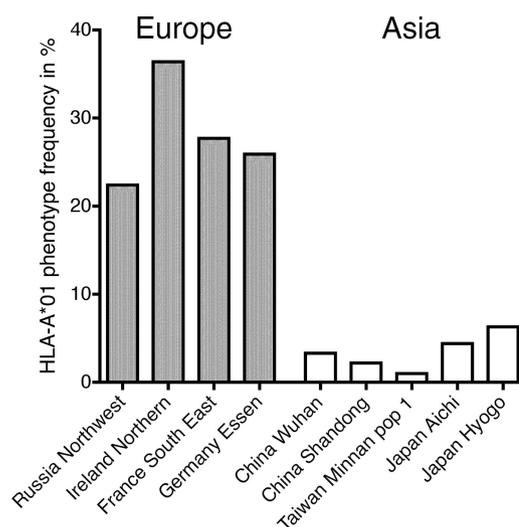


FIG. 8. Frequency of the HLA-A\*01 allele in populations in Europe and Asia. The frequencies of the HLA-A\*01 allele in different populations were retrieved from a public database (<http://www.allelefrequencies.net>). The observed phenotype frequencies in different cohorts are shown.

I binding as observed here make this scenario unlikely. Lack of efficient processing and presentation of the antigen completely prevents effective priming. In line with this we were reproducibly not able to expand specific T cells targeting the ATDALMTGY<sub>1436-1444</sub> epitope upon stimulation with the escape variant. Transmission of a virus harboring an antigen-processing mutation or an escape mutation in the anchor residue resulted in the lack of priming of a normally immunodominant CD8 response during acute HIV-1 infection (1, 3). The ATDALMTGY<sub>1436-1444</sub> epitope analyzed here represents an immunodominant CD8 response during acute and resolved HCV infection (23, 24). However, it is unclear how lack of such an immunodominant CD8 response may affect the functionality of the overall T-cell response and the ability to control viral replication. In animal models, conflicting results on the consequences of the lack of immunodominant CD8 epitopes on disease control have been reported. Infection of mice with HSV-1 or influenza virus lacking one immunodominant CD8 epitope resulted in a shifted epitope hierarchy, but enhancement of otherwise-subdominant CD8 responses compensated for the missing dominant response (4, 37). In contrast, mice infected with a lymphocytic choriomeningitis virus escape variant in one or two immunodominant CD8 epitopes presented with prolonged viremia and disease (32, 45). In HCV, expression of the HLA A\*01-B\*08-Cw07-DRB1\*03011-DQB1\*0201 haplotype was found in one study to be associated with viral persistence (30). That study was based on an analysis of a cohort infected by contaminated anti-D immune globulin from a single source in Ireland. Interestingly, the infecting virus harbored two escape mutations: one was the Y1444F mutation, and the other was located in an immunodominant HLA-B\*08-restricted epitope (HSKKKCEDL<sub>1395-1403</sub>) in NS3 (35). It is tempting to speculate that adaptation in the ATDALMTGY<sub>1436-1444</sub> epitope and that the B\*08-epitope contributed to the failure of this HLA-haplotype to control viral replication.

We detected the CD8 response against the ATDALMTGY<sub>1436-1444</sub> epitope in ca. 50% of all subjects. However, only 25% of circulating isolates contain the prototype sequence. Assuming that both variants can be transmitted upon infection, this high frequency of the CD8 response may be unexpected. As a possible explanation, it was hypothesized for a similar observation in the simian immunodeficiency virus-infected rhesus macaque model that viral variants may have reverted back to the prototype sequence and primed the response but were rapidly eliminated by specific T cells and therefore not detectable in serum (15). Transient reversion to the prototype as the basis for priming of a CD8 response after infection with an escape variant has been documented (1, 11), even though this may result in delayed priming of the response (1). In our study we were not able to find the prototype epitope sequence even in minor frequencies in HLA-A\*01-positive subjects. However, we observed one case of evolution back to the prototype residue in the absence of the HLA-A\*01 allele, indicating that such mutations occur at least in some isolates. It therefore seems possible that a low frequency of the prototype sequence was also present in some HLA-A\*01-positive subjects but was missed in our quasispecies analysis. This may have driven the corresponding CD8 response at a very low detection level, and

the immune response at the same time may have prevented full mutation back to the prototype.

We observed one case of evolution toward the prototype residue in a subject with chronic HCV genotype 3a infection. Similarly, reversion in this epitope was observed in 4 of 11 subjects in the Irish cohort infected by a single source of HCV genotype 1b (32). In contrast, in a similar single source outbreak with HCV genotype 1b in East Germany (46), only 1 of 19 sequences from HLA-A\*01-negative subjects has reverted 20 years after infection (data not shown). This raises the possibility that the Y1444F substitution is at least in some but not all isolates associated with fitness costs to the virus. To address the impact of the Y1444F substitution on the function of the encoded HCV helicase, the protein from a genotype 3a isolate and from genotype 1b (con1) was expressed, and the helicase activity was determined *in vitro*. The helicase activity of the Y1444F variant was reproducibly diminished compared to the prototype in the genotype 3a isolate. Similarly, in the con1 context the helicase containing the prototype epitope sequence (ATDALMTGY) was more active compared to the variant; however, the effect was in this case only minor. The recently developed HCV culture systems that produce infectious viral particles are based on HCV genotype 2a (28, 44, 47). Therefore, direct evaluation of the impact of this escape mutation observed in HCV genotype 3a and 1b on viral replication capacity is at the moment difficult. It is unclear how this difference in helicase function translates into differences in overall replication capacity. In a transient-replication model utilizing a subgenomic replicon based on con1, both the prototype and the variant were replication competent on a similar level, even though subtle differences in the replication level were difficult to dissect (data not shown). If optimal performance of the helicase supports viral replication, it seems possible that the prototype sequence is advantageous overall, even though clear evidence of a fitness benefit is lacking from the *in vitro* assays. Potentially, even subtle differences in replication capacity result in a selection advantage of a variant *in vivo*. However, the observation in our cohort that the majority of HLA-A\*01-negative subjects are infected with the escape variant indicates that the Y1444F substitution is relatively stable in many isolates and that evolution toward the prototype in the absence of immune pressure is not universal. Of note, there was no significant difference in viral load between subjects with the prototype and the Y1444F variant (data not shown).

The frequency of escape mutations in circulating strains in a population is not only determined by the associated fitness costs and consequently the rate of reversion. CD8 immune pressure mediated by frequent HLA class I alleles may contribute proportionally more to viral evolution at the population level compared to rare alleles. Hence, circulating strains in a population may be more adapted to immune pressure mediated by frequent HLA alleles and less adapted to immune pressure by rare alleles. In HIV it has been suggested that the expression of rare HLA alleles to which the virus is not yet adapted is even advantageous for containing viremia (43). HLA-A\*01 is a frequent allele in a Caucasian population, with reported phenotype frequencies between 20 and 35%. Reproducible immune pressure mediated by this common allele may support a high prevalence of the Y1444F escape variant in circulating strains. Inter-

estingly, the frequency of the Y1444F escape variant in isolates from different regions correlates well with reported HLA-A\*01 frequencies. Sequences from Asia, where HLA-A\*01 is rare, predominantly harbor the prototype sequence, indicating that circulating viruses in Asia are less frequently exposed to immune pressure mediated by this allele. It will therefore be interesting to investigate whether HLA-A\*01 is beneficial in such populations, i.e., in which this allele is rare and the circulating strains are predominantly prototype. However, the phylogenetic analysis of genotype 1b sequences suggests that in addition to numerous unambiguous sequence changes there are region-specific clusters that predominantly contain either prototype or the Y1444F variant. The interpretation of this observation is difficult. A recent phylogenetic reevaluation of HLA-associated sequence polymorphisms in HIV suggested that these clusters are unlikely to be selected by immune pressure; this phenomenon was termed "founder effect" (5). Many of the reported viral sequence differences in HIV corresponded to different subtypes, resulting in a detection bias toward associations between subtype-specific sequence polymorphisms and the predominant HLA allele of the population in which this subtype is common. However, in the present study, in which subtypes were analyzed separately, multiple sequence changes in position 1444 in HCV genotype 1b indicate that this site is continuously evolving. Even though we cannot completely rule out other mechanisms, we hypothesize that the Asian cluster predominantly contains prototype sequences because it was less exposed to HLA-A\*01-associated selection pressure. In turn, sequences in clusters predominantly containing Y1444F sequences may have additional sequence changes that compensate for fitness costs, which is the subject of additional ongoing studies. This may prevent frequent reversion, giving the phylogeny a less dynamic picture. Additional studies are clearly needed to address the impact of regional differences in HLA frequencies on the global evolution of HCV.

In summary, we demonstrate here that several factors influence the evolution in an HLA-A\*01-restricted epitope in NS3 in the individual but also at the population level. Importantly, this epitope is a frequent target. Specific CD8<sup>+</sup> T cells are detectable in approximately half of all HLA-A\*01-positive patients, even during chronic HCV infection, and reproducible immune pressure results in the selection of a CD8 escape variant that seemingly impairs HLA binding in all HLA-A\*01-positive patients. However, mutational escape in this epitope is potentially associated with viral fitness costs at least in some isolates, and evolution toward the prototype residue in the absence of the restricting HLA allele occurs, albeit seemingly at a low rate. Interestingly, in this interplay of positive and negative selection pressures the frequency of the HLA\*01-allele in a population seems to be an additional important factor that influences the overall frequency of the escape mutation in circulating isolates. In HLA-A\*01-rich populations the majority of HCV isolates is adapted to immune pressure in this epitope, and further compensatory evolution may have tipped the balance for the predominant consensus residue from prototype to the escape variant. Adaptation of circulating viruses to HLA class I-associated selection pressure in a pop-

ulation may further complicate the design of T-cell-based immune therapies.

#### ACKNOWLEDGMENTS

C.N.-H. and J.T. are supported by the German Hepatitis Network (Hep-Net). J.T. is supported by the DFG (TI 323/3-1). D.N.F. is supported by the National Institutes of Health (AI052395). P.K. is supported by the Wellcome Trust and the James Martin School of the 21st Century, Oxford, England.

We thank the German Hepatitis Network (Kompetenznetz Hepatitis) for the contribution of samples.

#### REFERENCES

- Allen, T. M., M. Altfeld, X. G. Yu, K. M. O'Sullivan, M. Lichterfeld, S. Le Gall, M. John, B. R. Mothe, P. K. Lee, E. T. Kalife, D. E. Cohen, K. A. Freedberg, D. A. Strick, M. N. Johnston, A. Sette, E. S. Rosenberg, S. A. Mallal, P. J. Goulder, C. Brander, and B. D. Walker. 2004. Selection, transmission, and reversion of an antigen-processing cytotoxic T-lymphocyte escape mutation in human immunodeficiency virus type 1 infection. *J. Virol.* **78**:7069–7078.
- Allen, T. M., X. G. Yu, E. T. Kalife, L. L. Reyor, M. Lichterfeld, M. John, M. Cheng, R. L. Allgaier, S. Mui, N. Frahm, G. Alter, N. V. Brown, M. N. Johnston, E. S. Rosenberg, S. A. Mallal, C. Brander, B. D. Walker, and M. Altfeld. 2005. De novo generation of escape variant-specific CD8<sup>+</sup> T-cell responses following cytotoxic T-lymphocyte escape in chronic human immunodeficiency virus type 1 infection. *J. Virol.* **79**:12952–12960.
- Altfeld, M., T. M. Allen, E. T. Kalife, N. Frahm, M. M. Addo, B. R. Mothe, A. Rathod, L. L. Reyor, J. Harlow, X. G. Yu, B. Perkins, L. K. Robinson, J. Sidney, G. Alter, M. Lichterfeld, A. Sette, E. S. Rosenberg, P. J. Goulder, C. Brander, and B. D. Walker. 2005. The majority of currently circulating human immunodeficiency virus type 1 clade B viruses fail to prime cytotoxic T-lymphocyte responses against an otherwise immunodominant HLA-A2-restricted epitope: implications for vaccine design. *J. Virol.* **79**:5000–5005.
- Andreansky, S. S., J. Stambas, P. G. Thomas, W. Xie, R. J. Webby, and P. C. Doherty. 2005. Consequences of immunodominant epitope deletion for minor influenza virus-specific CD8<sup>+</sup>-T-cell responses. *J. Virol.* **79**:4329–4339.
- Bhattacharya, T., M. Daniels, D. Heckerman, B. Foley, N. Frahm, C. Kadie, J. Carlson, K. Yusim, B. McMahon, B. Gaschen, S. Mallal, J. I. Mullins, D. C. Nickle, J. Herbeck, C. Rousseau, G. H. Learn, T. Miura, C. Brander, B. Walker, and B. Korber. 2007. Founder effects in the assessment of HIV polymorphisms and HLA allele associations. *Science* **315**:1583–1586.
- Bowen, D. G., and C. M. Walker. 2005. Adaptive immune responses in acute and chronic hepatitis C virus infection. *Nature* **436**:946–952.
- Brander, C., K. E. Hartman, A. K. Trocha, N. G. Jones, R. P. Johnson, B. Korber, P. Wentworth, S. P. Buchbinder, S. Wolinsky, B. D. Walker, and S. A. Kalams. 1998. Lack of strong immune selection pressure by the immunodominant, HLA-A\*0201-restricted cytotoxic T lymphocyte response in chronic human immunodeficiency virus-1 infection. *J. Clin. Investig.* **101**:2559–2566.
- Cox, A. L., T. Mosbrugger, Q. Mao, Z. Liu, X. H. Wang, H. C. Yang, J. Sidney, A. Sette, D. Pardoll, D. L. Thomas, and S. C. Ray. 2005. Cellular immune selection with hepatitis C virus persistence in humans. *J. Exp. Med.* **201**:1741–1752.
- DiBrino, M., K. C. Parker, J. Shiloach, R. V. Turner, T. Tsuchida, M. Garfield, W. E. Biddison, and J. E. Coligan. 1994. Endogenous peptides with distinct amino acid anchor residue motifs bind to HLA-A1 and HLA-B8. *J. Immunol.* **152**:620–631.
- Erickson, A. L., Y. Kimura, S. Igarashi, J. Eichelberger, M. Houghton, J. Sidney, D. McKinney, A. Sette, A. L. Hughes, and C. M. Walker. 2001. The outcome of hepatitis C virus infection is predicted by escape mutations in epitopes targeted by cytotoxic T lymphocytes. *Immunity* **15**:883–895.
- Fernandez, C. S., M. Z. Smith, C. J. Batten, R. De Rose, J. C. Reece, E. Rollman, V. Venturi, M. P. Davenport, and S. J. Kent. 2007. Vaccine-induced T cells control reversion of AIDS virus immune escape mutants. *J. Virol.* **81**:4137–4144.
- Frick, D. N. 2007. The hepatitis C virus NS3 protein: a model RNA helicase and potential drug target. *Curr. Issues Mol. Biol.* **9**:1–20.
- Frick, D. N., R. S. Rypma, A. M. Lam, and B. Gu. 2004. The nonstructural protein 3 protease/helicase requires an intact protease domain to unwind duplex RNA efficiently. *J. Biol. Chem.* **279**:1269–1280.
- Friedrich, T. C., E. J. Dodds, L. J. Yant, L. J. Vojnov, R. Rudersdorf, C. Cullen, D. T. Evans, R. C. Desrosiers, B. R. Mothe, J. Sidney, A. Sette, K. Kunstman, S. Wolinsky, M. Piatak, J. Lifson, A. L. Hughes, N. Wilson, D. H. O'Connor, and D. I. Watkins. 2004. Reversion of CTL escape-variant immunodeficiency viruses in vivo. *Nat. Med.* **10**:275–281.
- Friedrich, T. C., A. B. McDermott, M. R. Reynolds, S. Piaskowski, S. Fuenger, I. P. De Souza, R. Rudersdorf, C. Cullen, L. J. Yant, L. Vojnov, J. Stephany, S. Martin, D. H. O'Connor, N. Wilson, and D. I. Watkins. 2004. Consequences of cytotoxic T-lymphocyte escape: common escape mutations in simian immunodeficiency virus are poorly recognized in naive hosts. *J. Virol.* **78**:10064–10073.

16. Gaudieri, S., A. Rauch, L. P. Park, E. Freitas, S. Herrmann, G. Jeffrey, W. Cheng, K. Pfafferott, K. Naidoo, R. Chapman, M. Battegay, R. Weber, A. Telenti, H. Furrer, I. James, M. Lucas, and S. A. Mallal. 2006. Evidence of viral adaptation to HLA class I-restricted immune pressure in chronic hepatitis C virus infection. *J. Virol.* **80**:11094–11104.
17. Goulder, P. J., R. E. Phillips, R. A. Colbert, S. McAdam, G. Ogg, M. A. Nowak, P. Giangrande, G. Luzzi, B. Morgan, A. Edwards, A. J. McMichael, and S. Rowland-Jones. 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nat. Med.* **3**:212–217.
18. Guglietta, S., A. R. Garbuglia, V. Pacciani, C. Scotta, M. P. Perrone, L. Laurenti, E. Spada, A. Mele, M. R. Capobianchi, G. Taliani, A. Folgori, A. Vitelli, L. Ruggeri, A. Nicosia, E. Piccolella, and P. Del Porto. 2005. Positive selection of cytotoxic T lymphocyte escape variants during acute hepatitis C virus infection. *Eur. J. Immunol.* **35**:2627–2637.
19. Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**:160–174.
20. Kubo, R. T., A. Sette, H. M. Grey, E. Appella, K. Sakaguchi, N. Z. Zhu, D. Arnott, N. Sherman, J. Shabanowitz, H. Michel, et al. 1994. Definition of specific peptide motifs for four major HLA-A alleles. *J. Immunol.* **152**:3913–3924.
21. Kuiken, C., K. Yusim, L. Boykin, and R. Richardson. 2005. The Los Alamos hepatitis C sequence database. *Bioinformatics* **21**:379–384.
22. Lam, A. M., D. Keeney, P. Q. Eckert, and D. N. Frick. 2003. Hepatitis C virus NS3 ATPases/helicases from different genotypes exhibit variations in enzymatic properties. *J. Virol.* **77**:3950–3961.
23. Lauer, G. M., E. Barnes, M. Lucas, J. Timm, K. Ouchi, A. Y. Kim, C. L. Day, G. K. Robbins, D. R. Casson, M. Reiser, G. Dusheiko, T. M. Allen, R. T. Chung, B. D. Walker, and P. Klenerman. 2004. High resolution analysis of cellular immune responses in resolved and persistent hepatitis C virus infection. *Gastroenterology* **127**:924–936.
24. Lauer, G. M., M. Lucas, J. Timm, K. Ouchi, A. Y. Kim, C. L. Day, J. Schulze Zur Wiesch, G. Paranhos-Baccala, I. Sheridan, D. R. Casson, M. Reiser, R. T. Gandhi, B. Li, T. M. Allen, R. T. Chung, P. Klenerman, and B. D. Walker. 2005. Full-breadth analysis of CD8<sup>+</sup> T-cell responses in acute hepatitis C virus infection and early therapy. *J. Virol.* **79**:12979–12988.
25. Lauer, G. M., K. Ouchi, R. T. Chung, T. N. Nguyen, C. L. Day, D. R. Purkis, M. Reiser, A. Y. Kim, M. Lucas, P. Klenerman, and B. D. Walker. 2002. Comprehensive analysis of CD8<sup>+</sup> T-cell responses against hepatitis C virus reveals multiple unpredicted specificities. *J. Virol.* **76**:6104–6113.
26. Leslie, A., D. Kavanagh, I. Honeyborne, K. Pfafferott, C. Edwards, T. Pillay, L. Hilton, C. Thobakgale, D. Ramduth, R. Draenert, S. Le Gall, G. Luzzi, A. Edwards, C. Brander, A. K. Sewell, S. Moore, J. Mullins, C. Moore, S. Mallal, N. Bhardwaj, K. Yusim, R. Phillips, P. Klenerman, B. Korber, P. Kiepiela, B. Walker, and P. Goulder. 2005. Transmission and accumulation of CTL escape variants drive negative associations between HIV polymorphisms and HLA. *J. Exp. Med.* **201**:891–902.
27. Leslie, A. J., K. J. Pfafferott, P. Chetty, R. Draenert, M. M. Addo, M. Feeney, Y. Tang, E. C. Holmes, T. Allen, J. G. Prado, M. Altfeld, C. Brander, C. Dixon, D. Ramduth, P. Jeena, S. A. Thomas, A. S. John, T. A. Roach, B. Kupfer, G. Luzzi, A. Edwards, G. Taylor, H. Lyall, G. Tudor-Williams, V. Novelli, J. Martinez-Picado, P. Kiepiela, B. D. Walker, and P. J. Goulder. 2004. HIV evolution: CTL escape mutation and reversion after transmission. *Nat. Med.* **10**:282–289.
28. Lindenbach, B. D., M. J. Evans, A. J. Syder, B. Wolk, T. L. Tellinghuisen, C. C. Liu, T. Maruyama, R. O. Hynes, D. R. Burton, J. A. McKeating, and C. M. Rice. 2005. Complete replication of hepatitis C virus in cell culture. *Science* **309**:623–626.
29. Lindenbach, B. D., B. M. Pragai, R. Montserret, R. K. Beran, A. M. Pyle, F. Penin, and C. M. Rice. 2007. The C terminus of hepatitis C virus NS4A encodes an electrostatic switch that regulates NS5A hyperphosphorylation and viral replication. *J. Virol.* **81**:8905–8918.
30. McKiernan, S. M., R. Hagan, M. Curry, G. S. McDonald, A. Kelly, N. Nolan, A. Walsh, J. Hegarty, E. Lawlor, and D. Kelleher. 2004. Distinct MHC class I and II alleles are associated with hepatitis C viral clearance, originating from a single source. *Hepatology* **40**:108–114.
31. Moore, C. B., M. John, I. R. James, F. T. Christiansen, C. S. Witt, and S. A. Mallal. 2002. Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level. *Science* **296**:1439–1443.
32. Moskopidhis, D., and R. M. Zinkernagel. 1995. Immunobiology of cytotoxic T-cell escape mutants of lymphocytic choriomeningitis virus. *J. Virol.* **69**:2187–2193.
33. Neumann-Haefelin, C., T. Killinger, J. Timm, S. Southwood, D. McKinney, H. E. Blum, and R. Thimme. 2007. Absence of viral escape within a frequently recognized HLA-A26-restricted CD8<sup>+</sup> T-cell epitope targeting the functionally constrained hepatitis C virus NS5A/5B cleavage site. *J. Gen. Virol.* **88**:1986–1991.
34. Neumann-Haefelin, C., S. McKiernan, S. Ward, S. Viazov, H. C. Spangenberg, T. Killinger, T. F. Baumert, N. Nazarova, I. Sheridan, O. Pybus, F. von Weizsacker, M. Roggendorf, D. Kelleher, P. Klenerman, H. E. Blum, and R. Thimme. 2006. Dominant influence of an HLA-B27 restricted CD8<sup>+</sup> T-cell response in mediating HCV clearance and evolution. *Hepatology* **43**:563–572.
35. Ray, S. C., L. Fanning, X. H. Wang, D. M. Netski, E. Kenny-Walsh, and D. L. Thomas. 2005. Divergent and convergent evolution after a common-source outbreak of hepatitis C virus. *J. Exp. Med.* **201**:1753–1759.
36. Soderholm, J., G. Ahlen, A. Kaul, L. Frelin, M. Alheim, C. Barnfield, P. Liljestrom, O. Weiland, D. R. Milich, R. Bartenschlager, and M. Sallberg. 2006. Relation between viral fitness and immune escape within the hepatitis C virus protease. *Gut* **55**:266–274.
37. Stock, A. T., C. M. Jones, W. R. Heath, and F. R. Carbone. 2006. CTL response compensation for the loss of an immunodominant class I-restricted HSV-1 determinant. *Immunol. Cell Biol.* **84**:543–550.
38. Swofford, D. L. 2002. PAUP\*: phylogenetic analysis using parsimony (\*and other methods). Sinauer Associates, Sunderland, MA.
39. Tester, I., S. Smyk-Pearson, P. Wang, A. Wertheimer, E. Yao, D. M. Lewinsohn, J. E. Tavis, and H. R. Rosen. 2005. Immune evasion versus recovery after acute hepatitis C virus infection from a shared source. *J. Exp. Med.* **201**:1725–1731.
40. Thimme, R., D. Oldach, K. M. Chang, C. Steiger, S. C. Ray, and F. V. Chisari. 2001. Determinants of viral clearance and persistence during acute hepatitis C virus infection. *J. Exp. Med.* **194**:1395–1406.
41. Timm, J., G. M. Lauer, D. G. Kavanagh, I. Sheridan, A. Y. Kim, M. Lucas, T. Pillay, K. Ouchi, L. L. Reyor, J. Schulze zur Wiesch, R. T. Gandhi, R. T. Chung, N. Bhardwaj, P. Klenerman, B. D. Walker, and T. M. Allen. 2004. CD8 epitope escape and reversion in acute HCV infection. *J. Exp. Med.* **200**:1593–1604.
42. Timm, J., B. Li, M. G. Daniels, T. Bhattacharya, L. L. Reyor, R. Allgaier, T. Kuntzen, W. Fischer, B. E. Nolan, J. Duncan, J. Schulze Zur Wiesch, A. Y. Kim, N. Frahm, C. Brander, R. T. Chung, G. M. Lauer, B. T. Korber, and T. M. Allen. 2007. Human leukocyte antigen-associated sequence polymorphisms in hepatitis C virus reveal reproducible immune responses and constraints on viral evolution. *Hepatology* **46**:339–349.
43. Trachtenberg, E., B. Korber, C. Sollars, T. B. Kepler, P. T. Hraber, E. Hayes, R. Funkhouser, M. Fugate, J. Theiler, Y. S. Hsu, K. Kunstman, S. Wu, J. Phair, H. Erlich, and S. Wolinsky. 2003. Advantage of rare HLA supertype in HIV disease progression. *Nat. Med.* **9**:928–935.
44. Wakita, T., T. Pietschmann, T. Kato, T. Date, M. Miyamoto, Z. Zhao, K. Murthy, A. Habermann, H. G. Krausslich, M. Mizokami, R. Bartenschlager, and T. J. Liang. 2005. Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat. Med.* **11**:791–796.
45. Weidt, G., O. Utermohlen, J. Heukeshoven, F. Lehmann-Grube, and W. Deppert. 1998. Relationship among immunodominance of single CD8<sup>+</sup> T-cell epitopes, virus load, and kinetics of primary antiviral CTL response. *J. Immunol.* **160**:2923–2931.
46. Wiese, M., F. Berr, M. Lafrenz, H. Porst, and U. Oesen. 2000. Low frequency of cirrhosis in a hepatitis C (genotype 1b) single-source outbreak in Germany: a 20-year multicenter study. *Hepatology* **32**:91–96.
47. Zhong, J., P. Gastaminza, G. Cheng, S. Kapadia, T. Kato, D. R. Burton, S. F. Wieland, S. L. Uprichard, T. Wakita, and F. V. Chisari. 2005. Robust hepatitis C virus infection in vitro. *Proc. Natl. Acad. Sci. USA* **102**:9294–9299.