Journal of Medical Virology 84:570–581 (2012)

A Clustering Phenomenon Among HCV-1a Strains Among Patients Coinfected With HIV From Buenos Aires, Argentina

Federico Bolcic,[1,3] Natalia Laufer,[1,3] Juan Trinchero,[1] Leandro R. Jones,[2,3] and Jorge Quarleri[1,3]*

[1]*Microbiology Department, National Reference Center for AIDS Research, School of Medicine, University of Buenos Aires, Buenos Aires, Argentina*
[2]*Molecular Biology Division, Photobiology Station, Playa Unión, Rawson, Chubut, Argentina*
[3]*National Council for Scientific and Technical Research (CONICET), Buenos Aires, Argentina*

The human immunodeficiency virus (HIV) and hepatitis C virus (HCV) share the same transmission routes which lead to high coinfection rates. Among HIV-infected individuals such rates reached 21% in Argentina, being HCV-1a the most predominant subtype. In this work, 25 HCV subtype 1a (HCV-1a) strains from Argentinean patients coinfected with HIV were studied based on E2 and NS5A sequences. Phylogenetic analyses indicated that 12 strains were highly related to each other, constituting a highly supported (posterior probability = 0.95) monophyletic group that we called ''M.'' The remaining HCV strains (group dispersed or ''D'') were interspersed along the phylogenetic trees. When comparing both groups of HCV-1a, 10 amino acid differences were located in functional domains of E2 and NS5A proteins that appeared to affect eventually the peptides binding to MHC-I molecules thus favoring immune escape and contributing to the divergence of HCV genotypes. Bayesian coalescent analyses for HCV-1a cluster M isolates indicated that the time to the most recent common ancestor (tMRCA) overlaps with the age estimated recently for the HIV-BF epidemic in Argentina. Furthermore, the genomic characterization based on *pol* gene analysis from HIV viremic patients showed that most HIV isolates from patients coinfected with HCV-1a cluster M were BF recombinants with identical recombination patterns. In conclusion, these results suggest the presence of an HCV-1a monophyletic cluster with a potential HIV co-transmission by phylogenetic analyses. ***J. Med. Virol. 84:570–581, 2012.*** © 2012 Wiley Periodicals, Inc.

KEY WORDS: hepatitis C virus; subtype 1a; coinfection; HIV; BF recombinant subtype

INTRODUCTION

Hepatitis C virus (HCV) is an enveloped, positive-stranded RNA virus and represents the Hepacivirus genus in the *Flaviviridae* family [Choo et al., 1989]. The human immunodeficiency virus-type 1 (HIV-1) is a member of the Lentivirus genus of the family *Retroviridae* [Ratner et al., 1985]. The World Health Organization (WHO) estimates that about 3% of the world's population has been infected with HCV and that more than 170 million chronic carriers are at risk of developing cirrhosis and/or liver cancer [NIH, 2002]. Seven million persons living with HIV/AIDS are co-infected with HCV [Soriano et al., 2010].

The genetic diversity and evolution of the HCV are important factors that shape the global epidemic, with implications for diagnosis, pathogenesis, and treatment duration and response. At least 7 major HCV genotypes and more than 67 subtypes have been identified [Simmonds et al., 2005; Kuiken and Simmonds, 2009]. Some genotypes are associated with a predominant transmission route [Ramalho et al., 2000] or

Grant sponsor: University of Buenos Aires (partial support); Grant number: SECYT-UBA 2010–2012; Grant sponsor: Agencia Nacional de Promoción Científica y Tecnológica (partial support); Grant number: PICT PRH-2008 120; Grant sponsor: Consejo Nacional de Investigaciones Científicas y Técnicas (partial support); Grant number: CONICET-PIP112 200801 01773; Grant sponsor: Fogarty International Center/NIH Grant (BF and LN; partial support); Grant sponsor: AIDS International, Training and Research Program—Mount Sinai School of Medicine (BF and LN; partial support); Grant number: D43TW001037.

*Correspondence to: Jorge Quarleri, PhD, Centro Nacional de Referencia para el SIDA, Departamento de Microbiología, Facultad de Medicina (UBA), Paraguay 2155—Piso 11, (C1121ABG) Ciudad de Buenos Aires, Argentina.
E-mail: quarleri@fmed.uba.ar

Accepted 11 January 2012
DOI 10.1002/jmv.23243
Published online in Wiley Online Library (wileyonlinelibrary.com).

with a particular interferon resistance profile [Maekawa and Enomoto, 2009]. In Argentina, the rate of HIV–HCV coinfection is approximately 21% [Laufer et al., 2010]. A change was reported previously in the prevalence of HCV subtypes [Quarleri et al., 2007], being HCV-1b the most prevalent in HCV monoinfected patients [Quarleri et al., 2000] and subtype 1a in patients coinfected with HIV and HCV [Bolcic et al., 2008b, 2011]. The rate of response to pegylated interferon and ribavirin (peg-IFN + RBV) therapy is lower in HCV subtype 1a than subtypes 1b and 4 [Legrand-Abravanel et al., 2009] and it appears to have substituted subtype 1b gradually over the past 20 years [Ross et al., 2000]. Furthermore, when comparing patients infected only with HCV-1 with patients coinfected with HIV and HCV-1, the sustained virological response is still lower among the latter [Chung et al., 2004]. Such differential therapy response rate could be influenced by host and viral-related factors. Among the last ones, the non-structural 5A-NS5A- and the envelope-2-E2-proteins are proposed but its clinical value as a predictive factor for IFN-RBV therapy still remains uncertain [Hofmann et al., 2005; Wohnsland et al., 2007].

Considering the increasing prevalence of HCV-1a among patients coinfected with HIV, the main goal of the present work was to empower the knowledge of HCV-1a phylodynamics among patients coinfected with HIV by analyzing the molecular diversity of the Argentinean HCV-1a strains in Buenos Aires, Argentina.

MATERIALS AND METHODS

Patients

Of the 84 patients coinfected with HCV and HIV attending the Fernandez Hospital in Buenos Aires city who were positive for HCV antibodies by enzyme immunoassay (HCV 3.0 Murex, Abbott Laboratories, Diagnostic Division, Chicago, IL) and plasma HCV-RNA (AMPLICOR HCV test kit, Roche Molecular Systems, Pleasanton, CA), 25 of these individuals were infected with HCV-1a subtype (Inno-LiPA HCV II, Innogenetics, N.V., Zwijnaarde, Belgium) and available peg-IFN and RBV therapy outcome, being eligible for recruitment. Their demographic and clinical data are depicted in Tables Ia and Ib. Samples from each individual were collected immediately prior to start peg-IFN and RBV therapy to rule out any possible clustering due to the adaptive convergence forced by mutations.

This study was approved by the Ethics Committee of the School of Medicine, University of Buenos Aires, and consents were obtained from all participating patients.

RNA Isolation, Reverse Transcription, Amplification, and Sequencing

Viral RNA extraction, reverse transcription, and amplification were performed as follows: RNA was extracted from 200 μl plasma using a phenol/guanidine isothiocyanate solution (Trizol^{TM} LS Reagent; Gibco BRL, Carlsbad, CA), according to the manufacturer's instructions. The final product was dissolved in 20 μl of RNase-free water. Reverse transcription was performed with MMLV reverse transcriptase (Promega, Madison, WI) and random hexamers (Promega). The HCV-E2, HCV-NS5A and HIV-*pol* nested-PCR protocols were performed as described previously [Bolcic et al., 2008a; Dilernia et al., 2011] thus obtaining HCV-E2, HCV-NS5A, and HIV-*pol* amplicons with lengths of 321 bp (positions 2211–2531), 405 bp (positions 6918–7322), and 1,060 bp (positions 2253–3315) nucleotides, respectively. The products amplified were electrophoresed on 1.5% agarose gel and visualized by ethidium bromide staining. The DNA of each strain obtained from the purified nested-PCR products (Quick Spin; Qiagen, Venlo, Netherlands) was subjected directly to double-stranded sequencing with dye-labeled dideoxy terminators with those primers employed previously for each region in the nested-PCR (ABI PRISM 3100 automated sequencer, Applied Biosystems, Foster City, CA).

HCV and HIV Phylogenetic Analyses

DNA sequences were edited and assembled using the Sequencher software v.4.10.1 (Gene Codes). Reference sequences for the various HIV and HCV geno/subtypes were obtained from the GenBank [Kuiken et al., 2003, 2008]. Multiple sequence alignments were achieved by the Mafft program [Katoh et al., 2002, 2009].

Three HCV sequence datasets were incorporated into phylogenetic analyses, including 25 Argentinean and 260 reference sequences for E2, NS5A, or concatenated E2–NS5A.

For HIV phylogenetic analyses other three datasets were used. One corresponded to the whole *pol* amplicon and the other two to the unconnected B and F fractions. One hundred and ninety eight HIV reference sequences were included in the alignments. The unconnected fractions did not include the BF recombinant reference sequences to confirm and to determine the phylogenetic distribution of each fraction.

The most appropriate substitution model for the datasets was selected with hierarchical likelihood ratio testing implemented in ModelTest 3.7 [Posada and Crandall, 1998]. The best-fitting model for all datasets was the general time reversible model with a gamma distribution and a proportion of invariable sites (GTR + G + I). This model was used to estimate the phylogenetic relationships between the HCV-1a and HIV strains with a maximum likelihood method by heuristic searches implemented in the PhyML version 3.0 computer program [Guindon and Gascuel, 2003]. A BioNJ tree was used as an initial tree, which was reordered subsequently by Nearest Neighbor Interchange (NNI) and Subtree Pruning and Regrafting (SPR) branch permutations in order to maximize its

TABLE Ia. Baseline Characteristics of the 25 Patients

Patient	Gender	Age	Weight (kg)	Risk group	ALT (IU/ml)	CD4 (cells/ml)	HCV-VL (log_{10} IU/ml)	Fibrosis stage	Time from HIV infection[a]	Time from HCV infection[a]	Year sample collection
NR1	M	47	73	IDU	49	400	5.68	4	16	16	2007
NR2	M	37	77	IDU	66	522	6.01	1	2	16	2007
NR3	M	35	67	IDU	99	817	6.38	1	15	8	2007
NR4	M	39	76	IDU	48	414	5.61	3	8	12	2007
NR5	M	42	83	IDU	60	130	6.39	0	12	21	2007
NR6	F	51	ND	HTS	ND	547	5.25	ND	5	ND	2004
NR7	M	33	64	IDU	53	479	4.69	4	3	13	2007
NR8	M	33	72	IDU	124	376	6.08	2	5	5	2008
NR9	M	51	65	HTS	105	625	6.66	3	10	7	2008
NR10	M	40	75	IDU	97	303	5.86	2	21	10	2008
NR11	F	38	50	IDU	59	644	5.98	4	16	8	2009
NR12	M	45	70	HTS	50	483	5.40	1	1	5	2009
NR13	M	42	81	MSM	191	957	5.43	3	21	21	2009
NR14	M	41	58	MSM	92	262	6.53	1	ND	ND	2009
NR15	M	42	95	IDU	59	ND	6.38	2	5	12	2009
NR16	M	51	ND	IDU	ND	482	5.70	ND	7	ND	2004
NR17	M	42	ND	MSM	ND	468	5.70	ND	7	ND	2004
NR18	F	50	ND	IDU	ND	661	5.70	ND	5	ND	2003
NR19	M	55	ND	IDU	ND	422	5.52	ND	23	ND	2006
NR20	M	40	ND	MSM	ND	510	5.70	ND	9	ND	2005
SVR1	F	37	54	HTS	29	476	5.82	0	6	15	2007
SVR2	M	41	80	HTS	145	329	5.17	ND	12	12	2008
SVR3	M	34	67	IDU	41	1448	6.74	ND	25	25	2009
SVR4	M	45	65	MSM	115	989	4.76	ND	22	22	2009
SVR5	M	41	ND	IDU	ND	471	5.64	ND	17	ND	2004

IDU, injecting drug user; HTS, heterosexual; MSM, men who have sex with men; VL, viral load; ALT, alanine amino transferase; ND, not determined.
[a]Estimated since infection diagnosis.

likelihood score. Individual nodes' supports in the most likely tree topology were estimated with a non-parametric bootstrap test (1,000 replicates). The tree topology was midpoint rooted for illustration purposes and drawn using the Dendroscope v2.7.4 program [Huson et al., 2007].

Recombination Analysis

The 25 Argentinean HCV-1a sequences at both E2 and NS5A genes were screened for recombination by using the SimPlot program [Lole et al., 1999]. This software is based on a sliding window method and

TABLE Ib. Baseline Characteristics of the Patients: Comparison Between M and D Groups

Characteristics	M group (n = 12)	D group (n = 13)	*P*-Value
Gender			
Male	10	11	>0.05[a]
Female	2	2	
Risk group			
IDU	6	9	>0.05[a]
HTS	4	1	
MSM	2	3	
Non responders (NR)	9	11	>0.05[a]
Sustained virological responders (SVR)	3	2	
Fibrosis METAVIR 3–4 score	1.4	2.4	>0.05[b]
Age ± SD (years)	41 ± 5	42 ± 6	>0.05[b]
ALT ± SD (IU/ml)	82 ± 36	83 ± 56	>0.05[b]
CD4 ± SD (cells/ml)	456.1 ± 232.8	630.6 ± 293.1	>0.05[b]
HCV-VL ± SD (log copies/ml)	6.16 ± 6.16	6 ± 6.17	>0.05[b]
Estimated duration of HIV infection ± SD (years)[c]	11.5 ± 6.9	11.3 ± 7.9	>0.05[b]
Estimated duration of HCV infection ± SD (years)[c]	12.1 ± 5.9	15.8 ± 6.2	>0.05[b]

IDU, injecting drug user; HTS, heterosexual; MSM, men who have sex with men; VL, viral load; ALT, alanine-amino transferase; SD, standard deviation.
[a]Fisher's exact test.
[b]Mann–Whitney test.
[c]Time of serological diagnosis.

constitutes a way of displaying graphically the coherence of the sequence relationship over the entire length of a set of aligned homologous sequences. The window width and the step size were set to 100 and 20 bp, respectively [Lole et al., 1999]. The results obtained from the recombination analysis were confirmed by bootscanning [Salminen et al., 1995]. In this case, a sliding window of 100 nucleotides was used, moving 20 nucleotides at a time.

The HIV recombinant analyses were performed using the jumping profile Hidden Markov Model (jpHMM) probabilistic approach with program default parameters. They allowed comparing a sequence to a pre-calculated multiple alignment of the major HIV-1 subtypes. For a query sequence, phylogenetic recombination breakpoints are predicted and each region of the sequence is assigned to one HIV-1 subtype [Schultz et al., 2006, 2009; Zhang et al., 2006].

Amino Acid Analysis

The VESPA program [Korber and Myers, 1992] was used to characterize different amino acid signature patterns in a set of query sequences in relation to a set of reference sequences. VESPA calculates the frequency of each amino acid at each position in an alignment for the query and reference set, and selects the positions for which the most common character in the query set differs from that in the background set. The frequencies of characters at the signature sites were also calculated. The HCV H77 strain (GenBank accession number AF009606) was used as a reference sequence for amino acid numbering.

Comparative Analysis of the Major Histocompatibility Complex (MHC) Peptide Binding

Since the MHC–peptide binding is a necessary requirement for its recognition by a T cell, mutations affecting anchor positions (e.g., second position and the C-terminal position of the peptide) could affect their capability to bind MHC molecules. Considering the plausible different amino acid constitution in the E2 and NS5A derived-epitopes among Argentinean HCV-1a strains, the impact on the MHC–peptide binding was predicted by the algorithm based on matrix-based method NetMHC 3.2 server that exhibits the largest HLA allelic coverage [Lundegaard et al., 2011]. This tool uses artificial neural networks (ANNs) and weights matrices [Nielsen et al., 2003] trained for 57 different HLA alleles representing all 12 HLA A and B supertypes [Lund et al., 2004]. The HCV-1a cytotoxic T-lymphocyte (CTL) epitope sequences were defined according to those included in two databases: the HCV Immunology Database [www.hcv.lanl.gov] [Yusim et al., 2005] and the Immune Epitope Database [http://www.immuneepitope.com] [Vita et al., 2010] involving the epitopes that verified amino acid differences when comparing Argentinean HCV-1a groups.

Bayesian Coalescence Analyses

Rates of nucleotide substitutions per site and divergence times were estimated for the 25 HCV-1a sequences at both E2 and NS5A genes from Argentinean isolates. The sampling time was known for each sample and the time-span covered by the samples ranged from 2003 to 2009 (Table Ia). The Bayesian Markov chain Monte Carlo (MCMC) algorithm were used for the estimation of the posterior distribution of parameters, which is implemented in the BEAST v1.5.4 program [Drummond and Rambaut, 2007]. The posterior distribution of trees was estimated with Tamura–Nei variable base frequencies with a gamma distribution (TrN + G) model.

Three molecular clock models were compared to reconstruct trees under the most appropriate evolutionary model: a strict clock model and two models of relaxed clock, assuming either uncorrelated exponential or lognormal prior distribution of substitution rates among lineages [Drummond et al., 2006]. According to the Bayes Factor, the evolutionary model that best fit these datasets was the relaxed clock assuming an uncorrelated lognormal prior distribution and the demographic model used to determine the changes in effective population size over time was the Constant Size one. Analyses were run twice for 10^6 generations, sampling tree every 1,000 steps and discarding the first 10% as burn-in. MCMC convergence and effective sample size (ESS) were checked using the TRACER v1.5 program [Drummond and Rambaut, 2007]. Analyses were considered to have converged and reached stability after the burn-in period when the ESS value was higher than 200. Uncertainty in the estimates was indicated by the 95% highest probability density (HPD) values.

Nucleotide Sequences Accession Number

Argentinean sequences were submitted to GenBank with accession numbers JF749227–JF749251 for NS5A, JF749252–JF749276 for E2, and JN206619–JN206626 for HIV-*pol* sequences.

A full listing of the GenBank reference sequence accession numbers used in this study is available upon request.

Statistical Analysis

A descriptive analysis of baseline variables was conducted looking at the central tendency and dispersion. These values were compared aimed to ensure that the demographic, epidemiological, and clinical characteristics. Fisher's exact test was used to analyze qualitative variables and Mann–Whitney test to analyze quantitative variables. All statistical tests were two-sided, with a type I error of 5%. Statistical analyses were performed using the SPSS v.12.0 (SPSS Corporation, Chicago, IL).

RESULTS

According to the epidemiological data available, the HCV-1a subtype was transmitted mainly among injecting drug users. Prospectively, only five out of the 25 patients were sustained virological responders to peg-IFN and RBV treatment (no evidence of serum HCV-RNA at 6 months after the end-of-treatment) (Tables Ia and Ib). All patients were residents of Buenos Aires city and surrounding areas; no epidemiological link was found between them and they had no history of traveling abroad.

HCV Phylogenetic and Recombination Analyses

The phylogenetic relatedness among all characterized Argentinean HCV-1a strains were coincident irrespectively of the genomic region considered (E2, NS5A, and E2–NS5A concatenated sequences). These HCV-1a sequences formed two distinct groups among HCV-1a sequences from public databases. They include a monophyletic group of 12 HCV-1a closely related sequences defined as cluster "M" and supported statistically by a high bootstrap value. The remaining 13 sequences-named as group "D" were dispersed along the HCV-1a subtype (Fig. 1).

When comparing the characteristics of patients with HCV-1a from the groups M and D (Table Ib), no relationship was found among the phylogenetic distribution of Argentinean HCV-1a isolates, gender, route of infection, or the end-of-treatment response ($P > 0.05$, Fisher's exact test).

The recombination analysis showed that no HCV-1a Argentinean sequence presented recombination events in any of the genomic regions analyzed.

Amino Acid Analysis and Comparative HCV-1a Derived-Peptides MHC Class I Binding Predictions

The VESPA program was used to characterize different amino acid signature patterns for groups M and D through E2 and the NS5A amino acid sequences (Table II).

Group M sequences differed from those belonging to group D in 10 amino acid positions (N653D, E655D, S686T, and A710V for E2, and L2198V, Q2228E,

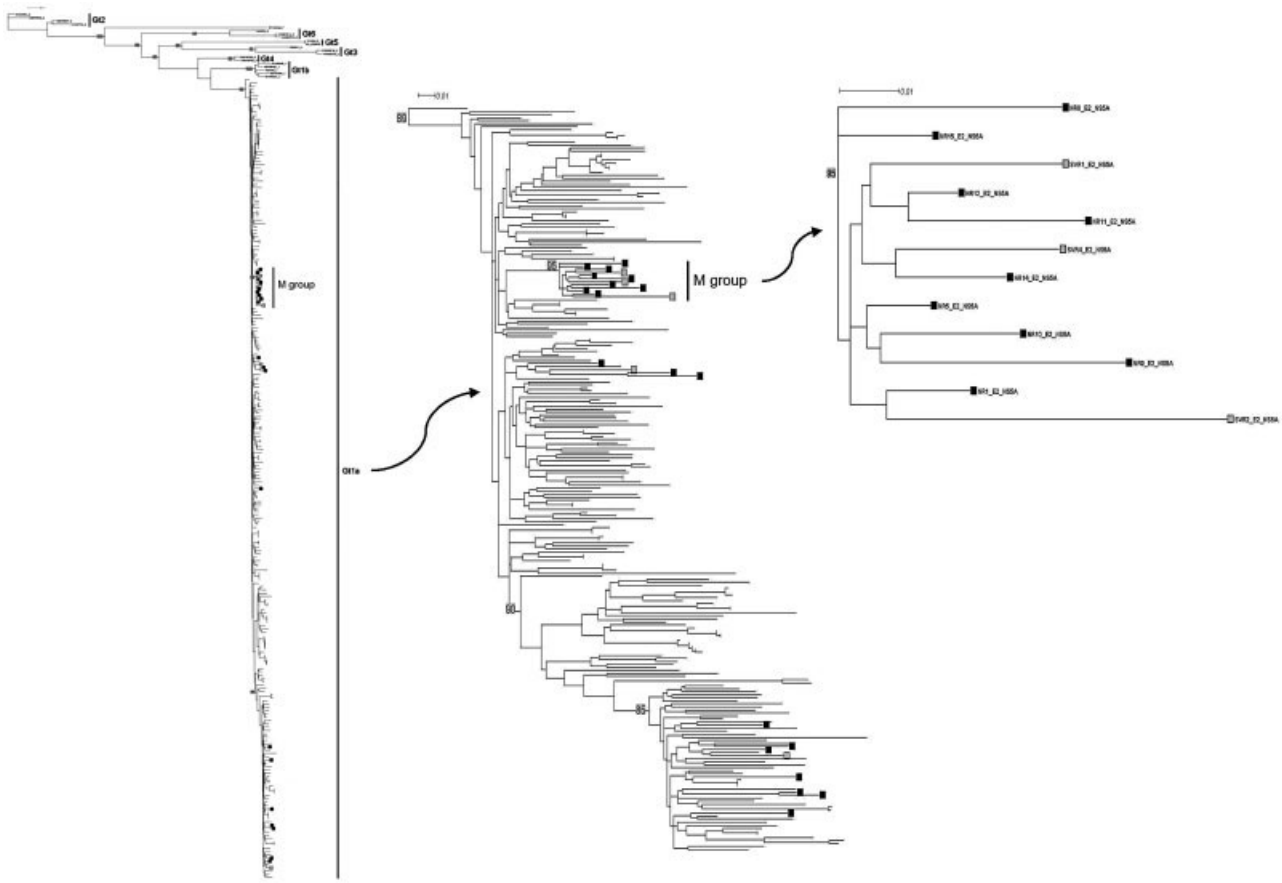

Fig. 1. Maximum likelihood tree of 285 HCV sequences including 25 HCV-1a Argentinean sequences. The HCV Argentinean monophyletic (M) group is denoted. Black and gray boxes indicate HCV isolates obtained for non responder (NR) and sustained virological responder (SVR) patients, respectively. Bootstrap values are shown on branches. Branch lengths are proportional to the number of nucleotide substitutions per aligned site (bar = 0.01 substitutions).

TABLE II. Summary of VESPA Program Results Comparing the HCV-la Group M and D Sequences

	E2			
Position	653	655	686	710
M group	**N 83%**; D 8%	**E 92%**; D 8%	**S 100%**; T 0%	**A 92%**; V 8%
D group	**D 77%**; N 8%	D 54%; E 38%	T 54%; S 46%	V 62%; A 38%

	NS5A					
Position	2198	2228	2268	2283	2287	2320
M group	**L 100%**; V 0%	**Q 75%**; E25	**I 83%**; V 8%	Q 58%; P 41%	V 67%; I25%	**Q 83%**; Q 17%
D group	V 46%; L 23%	E 62%; Q15%	**V 70%**; I23%	**P 92%**; Q 7%	**I 70%**; V23%	R 54%; Q 23%

Marker amino acids (≥70%) are in bold ($P \leq 0.05$).

I2268V, Q2283P, V2287I, and Q2320R for NS5A) (Table II), including twelve CTL epitopes. Six of them showed altered predicted binding affinity (four E2-derived and two NS5A-derived) to eleven MHC-I alleles by in silico analysis (Table III). The HCV-1a peptides are E2654–662 LEDRDRSEL, E2686–700 TTGLIHLHQNIVDV, E2701–715 YLYGVGSSIVSWAI, E2708–716 SIVSWAIKW, NS5A2197–2205 SVASSSASQ, and NS5A2266–2275 REISVPAEIL.

However, it is important to recognize the limitations inherent in the use of epitope prediction programs. No epitope prediction program is perfectly accurate, and there is substantial disagreement among epitopes predicted by each program [Gowthaman and Agrewala, 2008; Wang et al., 2008].

Rates and Dates of HCV Evolution

Demographic reconstruction analyses were performed in order to investigate the origin and spread of the 25 Argentinean HCV-1a strains. A relaxed clock was used in the datasets as the best model to fit the samples.

The time of the most recent common ancestor (tMRCA) for the sequences from the group M was estimated to be around 1991 for E2 and 1990 for NS5A while for the group D it was around 1973 for E2 and 1971 for NS5A. Irrespectively of the HCV genomic region considered, no differences were found when comparing the substitution rates between groups M and D (Table IV).

HIV Phylogenetic and Recombination Analyses

Considering the HIV–HCV coexistence in a given host and plausible reciprocal influences, HIV sequences from patients included in the study were also analyzed based on partial *pol* gene sequences. Given that all the individuals from the study population were under antiretroviral therapy, the HIV RNA was only detectable (>50 copies/ml) in 8 out of 25 patients. Among them, 6 corresponded to patients infected with HCV-1a strains who were included into the cluster M, and infected with HIV subtype BF, whereas the other two were infected with HIV subtype B (Fig. 2). The former were isolated from intravenous drug users (NR1, NR8; NR14), heterosexual (NR5), and men having sex with men (NR4, NR15) patients while for the latter were both IDU individuals. In order to determine the recombinant pattern of HIV sequences, the jpHMM algorithm was used to distinguish the different mosaics. Recombination analyses showed that five out of six HIV-BF recombinant strains (NR1, NR5, NR8, NR14, and NR15) exhibited the same *pol* gene mosaic pattern and confirmed that the B sequences (NR2 and NR17) were HIV-B (Fig. 3).

As the available phylogenetic tools are incapable of considering recombinant events, phylogenetic inferences of the unconnected B and F fractions of the sequences studied here were carried out. The ML tree corresponding to the B fraction showed no relationship among Argentinean HIV sequences. Conversely, the ML tree corresponding to the F portion suggests a phylogenetic association among five Argentinean HIV sequences (NR1, NR5, NR8, NR14, and NR15) (Fig. 4). Interestingly, those isolates were obtained from patients infected with HCV-1a subtype included in the cluster M.

DISCUSSION

In Argentina, HCV-1a is the most prevalent HCV subtype in patients coinfected with HIV [Bolcic et al., 2008b, 2011], whereas in HCV monoinfected patients, the most prevalent subtype was characterized as 1b [Quarleri et al., 2000]. It has been reported that HCV-1b in HCV monoinfected patients cluster in two distinct groups, and a statistical association was observed between HCV-1b sequence distribution and the final HCV treatment response [Di Lello et al., 2008].

This work describes the HCV-1a phylogenetic distribution of strains in patients coinfected with HIV from Buenos Aires, Argentina, and their relationship with HCV sequences around the world. A group of highly related sequences (cluster M) was found by analyzing two HCV genomics regions (E2 and NS5A) separately or concatenated. No association was observed between the HCV-1a sequence distribution and the route of infection, time of isolation or end-of-treatment response.

TABLE III. Amino Acids Variation Between Argentinean Sequences and the Relationship With Predicted HLA Binding

			HLA Alleles									
	Position	Peptide	A0202	A0203	A0250	A6801	A6802	B150	B1501	B1517	B4001	B4002	B4403
H77	654–662	L**E**DRDRSEL											
M group		*********	NB	NB	NB	NB	NB	NB	NB	NB	**B**	NB	NB
D group		*D********	NB	NB	NB	NB	NB	NB	NB	NB	NB	NB	NB
H77	686–701	**S**TGLIHLHQNIVDVQ											
M group		**********	NB	NB	NB	NB	NB	NB	NB	NB	NB	NB	NB
D group		T**********	NB	NB	NB	NB	**B**	NB	NB	NB	NB	NB	NB
H77	701–716	YLYGVGSSI**A**SWAIK											
M group		**********	NB	NB	NB	NB	NB	NB	NB	NB	NB	NB	NB
D group		**********V*******	NB	NB	NB	NB	**B**	**B**	**B**	NB	NB	NB	NB
H77	708–716	SI**A**SWAIKW											
M group		*********	NB	NB	NB	NB	NB	NB	NB	**B**	NB	NB	NB
D group		**V********	NB	NB	NB	NB	NB	NB	NB	NB	NB	NB	NB
H77	2197–2205	S**V**ASSSASQ											
M group		*L********	**B**	**B**	**B**	NB	NB	NB	**B**	NB	NB	NB	NB
D group		**********	NB	NB	NB	**B**	NB	NB	NB	NB	NB	NB	NB
H77	2266–2275	RE**I**SVPAEIL											
M group		**********	NB	NB	NB	NB	NB	NB	NB	NB	B	**B**	**B**
D group		**V*******	NB	NB	NB	NB	NB	NB	NB	NB	B	NB	NB

Gray boxes denote potential change in peptide binding affinity. NB, non-binding to HLA alleles; B, binding to HLA alleles.
Bolded and underlined amino acids in the prototype HCV-1a sequence (H77) are those that exhibit variations in comparison with in M or D groups of HCV sequences.
*Identical amino acid sequence.

The HCV-1a stains included in the cluster M presented a characteristic amino acid signature pattern that differed in 10 positions (4 in E2; 6 in NS5A) in comparison to group D. By in silico analysis, these amino acid variations seemed to have a potential impact on the binding to MHC-I molecules and could eventually favor the escape to immune surveillance which is a significant driving force in the divergence of the HCV genotypes [Duan et al., 2010; Shehzadi et al., 2011; de Queiroz et al., 2011]. However, the potential changes in binding affinities deserve further confirmation by biological assays taking into account that they were drawn from bioinformatics analyses and that the HLA haplotypes in these patients were not determined.

The tMRCA of the HCV-1a cluster M sequences was calculated to be around 20 years ago with a high posterior density going from the 1970s to the beginning of 2000. Interestingly, several studies [Aulicino et al., 2007; Bello et al., 2010; Dilernia et al., 2011] have reported recently that the origin of the HIV-BF subtype in South America dates from the beginning of the 1970s, overlapping with the HPD 95% found for the HCV-1a Argentinean sequences included into the cluster M. Nevertheless, the 95% HPD intervals found for the tMRCA of the HCV-1a cluster, final conclusions remain elusive. This limitation deserves further investigation with a longer time span of sampling that allows reaching short confidence intervals for dates, and higher precision. Another possibility would be the use of an alternative genomic region with available external calibration rates, like the NS5B gene. Considering that patients enrolled in this study come from Buenos Aires and surrounding areas, further research is needed to elucidate whether additional HCV-1a sequences from other Argentinean provinces might be incorporated into the cluster M and if these HCV-1a isolates are combined with HIV BF isolates when they are identified in patients co-infected with HIV.

HIV strains in these co-infected patients were B or BF recombinants. In Argentina, the CRF12_BF and related recombinants are predominant among injecting drug users and heterosexually infected individuals, while subtype B is found preferentially among MSM [Avila et al., 2002] but here the distribution did

TABLE IV. Estimated Evolutionary Parameters of Argentinean Sequences Included in the Different Groups for Both E2 and NS5A Analyzed Regions

Region	Description	Substitution rate[a]	tMRCA (95% HPD)
E2	M group	9.2E-3 (1.68E-5/2E-2)	1991 (2003–1973)
	D group	5.2E-3 (1.3E-3/9.5E-3)	1977 (1999–1941)
NS5A	M group	1.04E-2 (1.2E-5/2E-2)	1990 (2006–1971)
	D group	7.01E-3 (4.8E-3/9.2E-3)	1976 (2000–1931)

All values have ESS > 200.
[a]Substitutions per site per year.

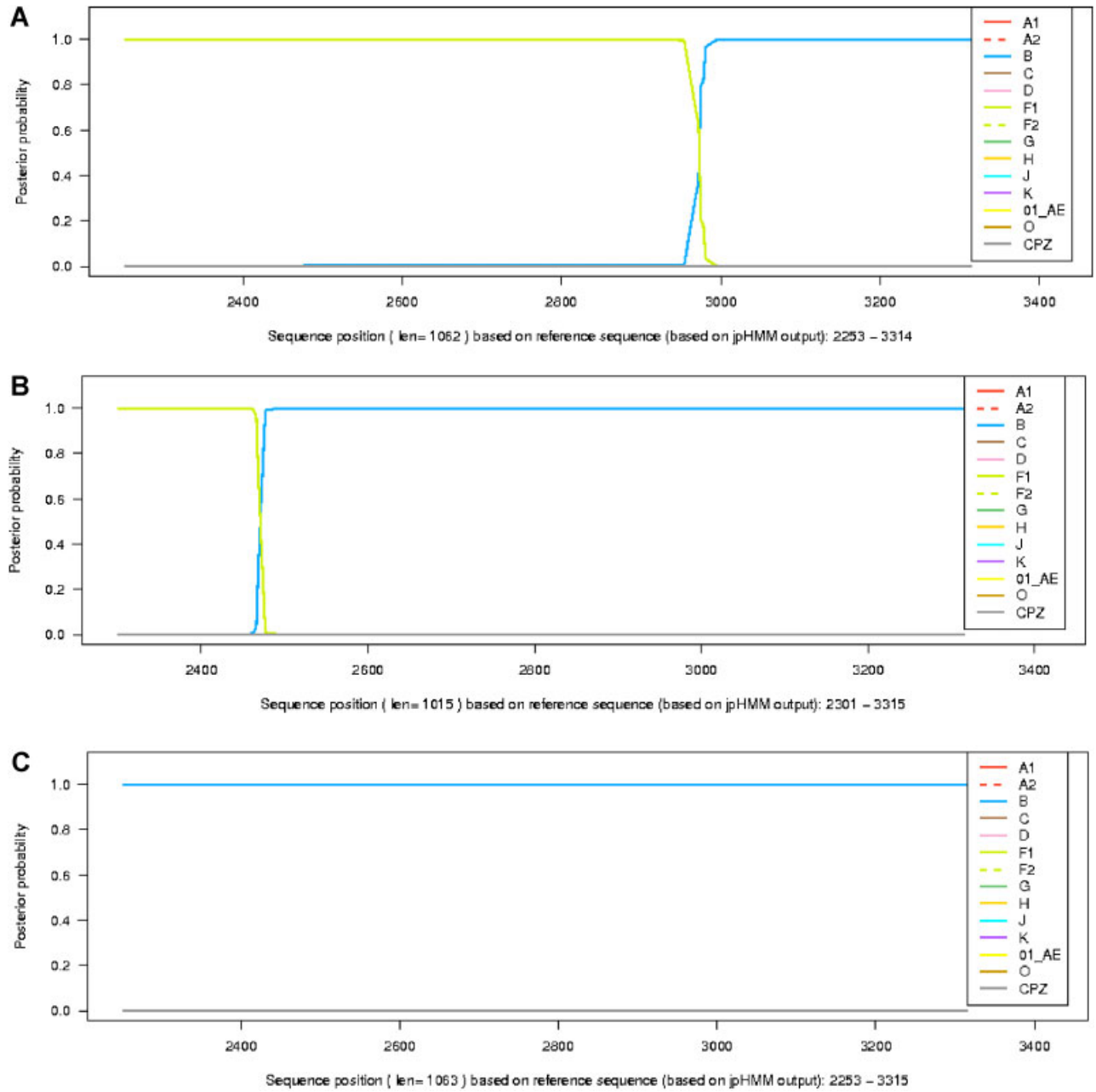

Fig. 2. HIV-*pol* representative plots for recombination analysis. Three different patterns were obtained: (**A**) BF recombinant pattern identified for HIV-*pol* sequences in non-responder patient 1 (NR1), NR8, NR12, NR14, and NR15; (**B**) BF recombinant pattern identified for HIV-*pol* sequences in a sustained virological responder patient 4 (SVR4); and (**C**) HIV-*pol* sequences subtype B identified in NR2 and NR17.

not appear to be related with the associated risk of infection. Nevertheless the relatively small sample size precludes any final conclusion.

The plausible coincidence observed between the tMRCA of the HCV-1a cluster M and the Argentinean HIV-BF epidemic allows hypothesizing that a lineage of HCV-1a could have been introduced into Buenos Aires, Argentina, together with the HIV-BF epidemic. In this regard, three additional observations give additional support. First, the HIV-BF recombinants were isolated coincidentally from coinfected patients whose HCV-1a strains belonged to the cluster M. Second, most of these HIV-BF recombinant strains exhibited an identical *pol* gene mosaic pattern. Finally, they were related phylogenetically to each other after the analysis of the F unconnected fraction. Phylogenetic analyses of the corresponding B fraction of the genome indicated no association, but considering that

Fig. 3. Maximum likelihood tree of 200 HIV-*pol* sequences, including 8 Argentinean sequences. Gray and black boxes indicate the sequences linked to HCV strains from the M and D, respectively. Bootstrap values are given on branches. Branch lengths are proportional to the number of nucleotide substitutions per aligned site (bar = 0.1 substitutions).

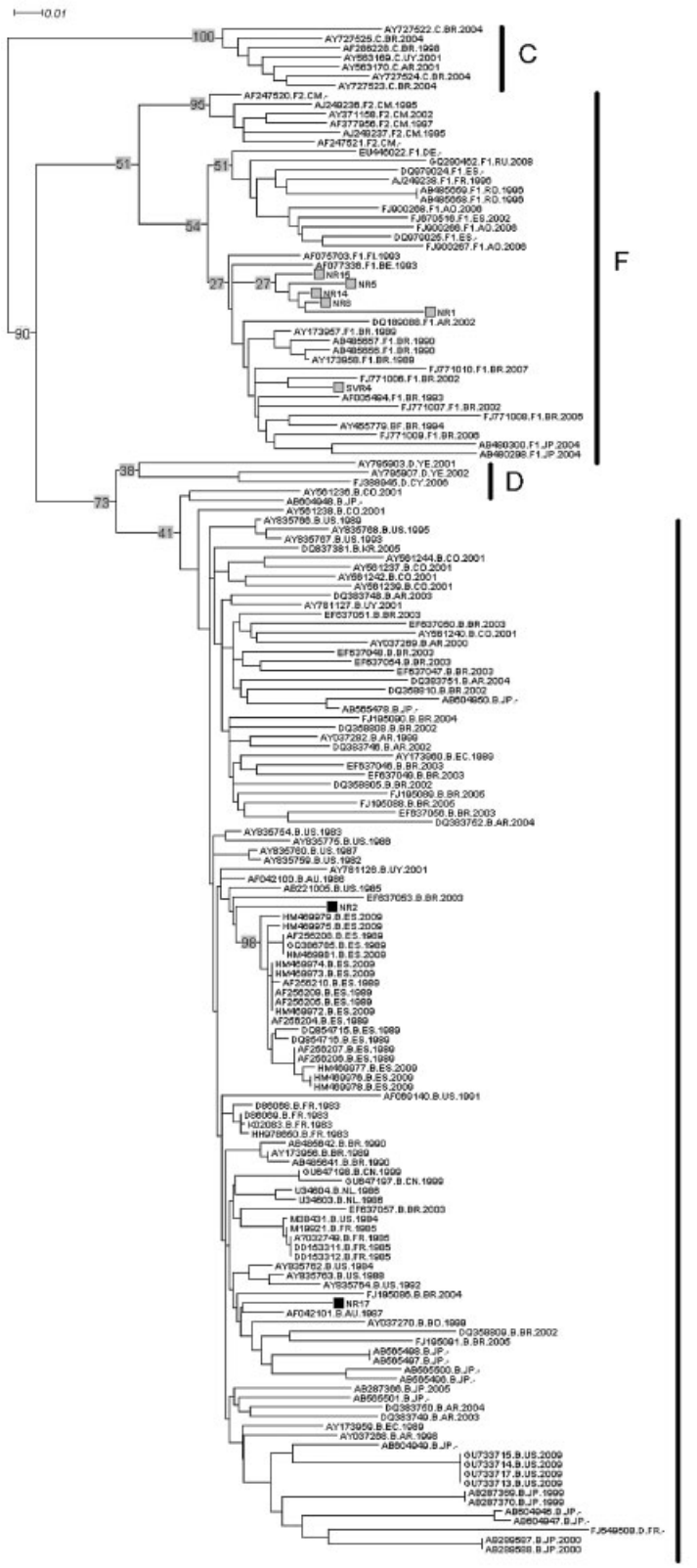

Fig. 4. Maximum likelihood tree of the *pol* HIV-F fractions of the HIV-BF recombinant sequences. Gray and black boxes indicate the sequences linked to HCV strains from the M and D groups, respectively. Bootstrap values are given on branches. Branch lengths are proportional to the number of nucleotide substitutions per aligned site (bar = 0.1 substitutions).

the B fragment is considerably smaller than the F portion, a relatively weak phylogenetic signal in these sequences was conceivable.

In conclusion, the analyses of HCV subtype 1a strains isolated from patients coinfected with HCV and HIV attended in a single Hospital in Buenos Aires, Argentina, allowed identification of a monophyletic cluster with a potential HCV-1a/HIV cotransmission by phylogenetic analyses.

ACKNOWLEDGMENTS

We thank Sergio Mazzini (Centro Nacional de Referencia para el SIDA, Departamento de Microbiología, Facultad de Medicina, Universidad de Buenos Aires, Argentina) for his assistance during the manuscript preparation.

REFERENCES

Aulicino PC, Holmes EC, Rocco C, Mangano A, Sen L. 2007. Extremely rapid spread of human immunodeficiency virus type 1 BF recombinants in Argentina. J Virol 81:427–429.

Avila MM, Pando MA, Carrion G, Peralta LM, Salomon H, Carrillo MG, Sanchez J, Maulen S, Hierholzer J, Marinello M, Negrete M, Russell KL, Carr JK. 2002. Two HIV-1 epidemics in Argentina: Different genetic subtypes associated with different risk groups. J Acquir Immune Defic Syndr 29:422–426.

Bello G, Aulicino PC, Ruchansky D, Guimaraes ML, Lopez-Galindez C, Casado C, Chiparelli H, Rocco C, Mangano A, Sen L, Morgado MG. 2010. Phylodynamics of HIV-1 circulating recombinant forms 12_BF and 38_BF in Argentina and Uruguay. Retrovirology 7:22–30.

Bolcic F, Bull L, Martinez L, Reynoso R, Salomon H, Arduino R, Barnett B, Quarleri J. 2008a. Analysis of sequence configurations of the PKR-interacting HCV proteins from plasma and PBMC as predictors of response to interferon-alpha and ribavirin therapy in HIV-coinfected patients. Intervirology 51:261–264.

Bolcic F, Laufer N, Quarleri J. 2008b. Hepatitis C virus genotypes distribution and viral load level among HIV-coinfected patients. Acta Gastroenterol Latinoam 38:307.

Bolcic F, Jones LR, Laufer N, Quarleri J. 2011. Molecular characterization of hepatitis C virus genotype 4 sequences in HIV-coinfected patients from Argentina. J Med Virol 83:935–940.

Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. 1989. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. Science 244:359–362.

Chung RT, Andersen J, Volberding P, Robbins GK, Liu T, Sherman KE, Peters MG, Koziel MJ, Bhan AK, Alston B, Colquhoun D, Nevin T, Harb G, van der Horst C. 2004. Peginterferon Alfa-2a plus ribavirin versus interferon alfa-2a plus ribavirin for chronic hepatitis C in HIV-coinfected persons. N Engl J Med 351:451–459.

de Queiroz AT, Maracaja-Coutinho V, Jardim AC, Rahal P, de Carvalho-Mello IM, Matioli SR. 2011. Relation of pretreatment sequence diversity in NS5A region of HCV genotype 1 with immune response between pegylated-INF/ribavirin therapy outcomes. J Viral Hepat 18:142–148.

Di Lello F, Garcia G, Kott V, Sookoian S, Campos R. 2008. Diversity of hepatitis C virus genotype 1b in Buenos Aires, Argentina: Description of a new cluster associated with response to treatment. J Med Virol 80:619–627.

Dilernia DA, Jones LR, Pando MA, Rabinovich RD, Damilano GD, Turk G, Rubio AE, Pampuro S, Gomez-Carrillo M, Salomon H. 2011. Analysis of HIV type 1 BF recombinant sequences from South America dates the origin of CRF12_BF to a recombination event in the 1970s. AIDS Res Hum Retroviruses 27:569–578.

Drummond AJ, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214–221.

Drummond AJ, Ho SY, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol 4:e88–e99.

Duan H, Struble E, Zhong L, Mihalik K, Major M, Zhang P, Feinstone S, Feigelstock D. 2010. Hepatitis C virus with a naturally occurring single amino-acid substitution in the E2 envelope protein escapes neutralization by naturally-induced and vaccine-induced antibodies. Vaccine 28:4138–4144.

Gowthaman U, Agrewala JN. 2008. In silico tools for predicting peptides binding to HLA-class II molecules: More confusion than conclusion. J Proteome Res 7:154–163.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704.

Hofmann WP, Zeuzem S, Sarrazin C. 2005. Hepatitis C virus-related resistance mechanisms to interferon alpha-based antiviral therapy. J Clin Virol 32:86–91.

Huson DH, Richter DC, Rausch C, Dezulian T, Franz M, Rupp R. 2007. Dendroscope: An interactive viewer for large phylogenetic trees. BMC Bioinformatics 8:460–465.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066.

Katoh K, Asimenos G, Toh H. 2009. Multiple alignment of DNA sequences with MAFFT. Methods Mol Biol 537:39–364.

Korber B, Myers G. 1992. Signature pattern analysis: A method for assessing viral sequence relatedness. AIDS Res Hum Retroviruses 8:1549–1560.

Kuiken C, Simmonds P. 2009. Nomenclature and numbering of the hepatitis C virus. Methods Mol Biol 510:33–53.

Kuiken C, Korber B, Shafer RW. 2003. HIV sequence databases. AIDS Rev 5:52–61.

Kuiken C, Hraber P, Thurmond J, Yusim K. 2008. The hepatitis C sequence database in Los Alamos. Nucleic Acids Res 36:D512–D516.

Laufer N, Quarleri J, Bouzas MB, Juncos G, Cabrini M, Moretti F, Bolcic F, Fernandez-Giuliano S, Mammana L, Salomon H, Cahn P. 2010. Hepatitis B virus, hepatitis C virus and HIV coinfection among people living with HIV/AIDS in Buenos Aires, Argentina. Sex Transm Dis 37:342–343.

Legrand-Abravanel F, Colson P, Leguillou-Guillemette H, Alric L, Ravaux I, Lunel-Fabiani F, Bouviers-Alias M, Trimoulet P, Chaix ML, Hezode C, Foucher J, Fontaine H, Roque-Afonso AM, Gassin M, Schvoerer E, Gaudy C, Roche B, Doffoel M, D'Alteroche L, Vallet S, Baazia Y, Pozzetto B, Thibault V, Nousbaum JB, Roulot D, Coppere H, Poynard T, Payan C, Izopet J. 2009. Influence of the HCV subtype on the virological response to pegylated interferon and ribavirin therapy. J Med Virol 81:2029–2035.

Lole KS, Bollinger RC, Paranjape RS, Gadkari D, Kulkarni SS, Novak NG, Ingersoll R, Sheppard HW, Ray SC. 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J Virol 73:152–160.

Lund O, Nielsen M, Kesmir C, Petersen AG, Lundegaard C, Worning P, Sylvester-Hvid C, Lamberth K, Roder G, Justesen S, Buus S, Brunak S. 2004. Definition of supertypes for HLA molecules using clustering of specificity matrices. Immunogenetics 55:797–810.

Lundegaard C, Lund O, Nielsen M. 2011. Prediction of epitopes using neural network based methods. J Immunol Methods 374:26–34.

Maekawa S, Enomoto N. 2009. Viral factors influencing the response to the combination therapy of peginterferon plus ribavirin in chronic hepatitis C. J Gastroenterol 44:1009–1015.

Nielsen M, Lundegaard C, Worning P, Lauemoller SL, Lamberth K, Buus S, Brunak S, Lund O. 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci 12:1007–1017.

NIH. 2002. NIH consensus statement on management of hepatitis C: 2002. NIH Consens State Sci Statements 19:1–46.

Posada D, Crandall KA. 1998. MODELTEST: Testing the model of DNA substitution. Bioinformatics 14:817–818.

Quarleri JF, Robertson BH, Mathet VL, Feld M, Espinola L, Requeijo MP, Mando O, Carballal G, Oubina JR. 2000. Genomic and phylogenetic analysis of hepatitis C virus isolates from argentine patients: A six-year retrospective study. J Clin Microbiol 38:4560–4568.

Quarleri JF, Bolcic FM, Bouzas MB, Laufer N, Gomez Carrillo M, Mammana L, Kaufman S, Perez H, Cahn P, Salomon H. 2007.

HCV genotype distribution among HIV co-infected individuals in Argentina: Relationship with host and viral factors. Acta Gastroenterol Latinoam 37:76–83.

Ramalho F, Costa A, Pires A, Cabrita P, Serejo F, Correia AP, Fatela N, Cloria H, Lopes J, Pinto HC, Marinho R, Raimundo M, Velosa J, Batista A, de Moura MC. 2000. Correlation of genotypes and route of transmission with histologic activity and disease stage in chronic hepatitis C. Dig Dis Sci 45:182–187.

Ratner L, Gallo RC, Wong-Staal F. 1985. HTLV-III, LAV, ARV are variants of same AIDS virus. Nature 313:636–637.

Ross RS, Viazov S, Renzing-Kohler K, Roggendorf M. 2000. Changes in the epidemiology of hepatitis C infection in Germany: Shift in the predominance of hepatitis C subtypes. J Med Virol 60: 122–125.

Salminen MO, Carr JK, Burke DS, McCutchan FE. 1995. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. AIDS Res Hum Retroviruses 11:1423–1425.

Schultz AK, Zhang M, Leitner T, Kuiken C, Korber B, Morgenstern B, Stanke M. 2006. A jumping profile Hidden Markov Model and applications to recombination sites in HIV and HCV genomes. BMC Bioinformatics 7:265–279.

Schultz AK, Zhang M, Bulla I, Leitner T, Korber B, Morgenstern B, Stanke M. 2009. jpHMM: Improving the reliability of recombination prediction in HIV-1. Nucleic Acids Res 37:W647–W651.

Shehzadi A, Ur Rehman S, Idrees M. 2011. Promiscuous prediction and conservancy analysis of CTL binding epitopes of HCV 3a viral proteome from Punjab Pakistan: An in silico approach. Virol J 8:55–67.

Simmonds P, Bukh J, Combet C, Deleage G, Enomoto N, Feinstone S, Halfon P, Inchauspe G, Kuiken C, Maertens G, Mizokami M, Murphy DG, Okamoto H, Pawlotsky JM, Penin F, Sablon E, Shin IT, Stuyver LJ, Thiel HJ, Viazov S, Weiner AJ, Widell A. 2005. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. Hepatology 42:962–973.

Soriano V, Vispo E, Labarga P, Medrano J, Barreiro P. 2010. Viral hepatitis and HIV co-infection. Antiviral Res 85:303–315.

Vita R, Zarebski L, Greenbaum JA, Emami H, Hoof I, Salimi N, Damle R, Sette A, Peters B. 2010. The immune epitope database 2.0. Nucleic Acids Res 38:D854–D862.

Wang P, Sidney J, Dow C, Mothe B, Sette A, Peters B. 2008. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. PLoS Comput Biol 4:e1000048–e1000057.

Wohnsland A, Hofmann WP, Sarrazin C. 2007. Viral determinants of resistance to treatment in patients with hepatitis C. Clin Microbiol Rev 20:23–38.

Yusim K, Richardson R, Tao N, Dalwani A, Agrawal A, Szinger J, Funkhouser R, Korber B, Kuiken C. 2005. Los alamos hepatitis C immunology database. Appl Bioinformatics 4:217–225.

Zhang M, Schultz AK, Calef C, Kuiken C, Leitner T, Korber B, Morgenstern B, Stanke M. 2006. jpHMM at GOBICS: A web server to detect genomic recombinations in HIV-1. Nucleic Acids Res 34:W463–W465.