Infection, Genetics and Evolution 26 (2014) 32–40

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Spreading of hepatitis C virus subtypes 1a and 1b through the central region of Argentina

Andrés Carlos Alberto Culasso [a], Adrián Farías [b], Federico Alejandro Di Lello [a], Marcelo Darío Golemba [a,c], Viviana Ré [b], Luciana Barbini [a], Rodolfo Campos [a,*]

[a] Cátedra de Virología, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Buenos Aires, Argentina
[b] Instituto de Virología Dr. José Maria Vanella, Universidad Nacional de Córdoba, Córdoba, Argentina
[c] Cellular Biology and Retroviruses Laboratory, Hospital de Pediatría S.A.M.I.C. "Prof. Dr. Juan P. Garrahan", Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 22 February 2014
Received in revised form 1 May 2014
Accepted 5 May 2014
Available online 10 May 2014

Keywords:
Hepatitis C virus
Phylogeography
Epidemiology
Argentina

ABSTRACT

The recent history of the hepatitis C virus (HCV) subtypes 1a and 1b in the central region of Argentina is hypothesized by phylogeographic reconstruction using coalescent based Bayesian analyses. Direct partial E2 sequences from HCV 1a and 1b infected patients attending different health-care centers of the country were analyzed.

The inferred date of the most recent common ancestor (t_{MRCA}) for HCV-1a was: 1962 (between 1943 and 1977) and for HCV-1b was earlier: 1929 (between 1895 and 1953). Diverse ancestral populations were inferred from both subtypes in Córdoba and in Buenos Aires cities and after that, HCV spread within and between larger cities and to other smaller cities. The analyses suggested that HCV-1b was dispersed first and it is currently in a stationary phase whereas HCV-1a was dispersed latter and it is still in a growth phase.

Finally, as it was observed in the developed countries, while the transmission of HCV-1b appears to have been somehow prevented, the HCV-1a may still represent a concern in the public health. Further work should be carried out to address their current transmission rate (and its main transmission route) in the Argentinean population.

1. Introduction

About 2% of world population is chronically infected with the hepatitis C virus (HCV) (Lavanchi, 2009). This blood-borne pathogen produces in most (about 80%) of the cases chronic infections that may cause severe liver diseases such as cirrhosis or hepatocellular carcinoma (Lauer and Walker, 2001). Taxonomically, HCV belongs to the *Flaviviridae* family as a member of *Hepacivirus* genus. By the phylogenetic analysis of its genomic sequences, HCV isolations can be classified into 7 genotypes (1–7) and at least 67 subtypes (named with small letters) (Smith et al., 2014). Some genotypes have an endemic distribution, such as genotypes 5 and 6 (prevalent in Southern Africa and Asia respectively), whereas others have a worldwide distribution like genotypes 1, 2 and 3 (particularly subtypes 1a, 1b, 2a and 3a) (Simmonds, 2004). In Argentina, the circulation of subtypes 1a, 1b, 2c, 3a and 4a was reported (Ré et al., 2007; Kershenobich et al., 2011; del Pino et al., 2013). Although each study had yielded a different estimation of the Argentinean HCV prevalence, between 0.32% and about 2%, all agree in that the genotype 1 account for more than the half of the cases.

Several researchers used Bayesian coalescence methodology to model the viral population dynamics and thus hypothesize about the possible causes for the global distribution of HCV (Njouom et al., 2007; Magiorkinis et al., 2009; Gray et al., 2013). Also, this methodology allowed the local scale study of the virus in many countries demonstrating different patterns of transmission associated in all cases with the regional demographic history of the inhabitants (Ciccozzi et al., 2011; Ré et al., 2011; Di Lello et al., 2013; Zehender et al., 2013; Golemba et al., 2013). The reconstruction of the events that shaped the actual epidemics allowed assessing the current status of the viral transmission in these communities which may contribute to the development of public health policies (Pybus et al., 2001).

The recent development of continuous phylogeographical models allows the reconstruction of the viral population dynamics in time and space and thus provides a richer picture of the virus

* Corresponding author. Address: Facultad de Farmacia y Bioquímica, Junín 956, 4th Floor, 1113 Buenos Aires, Argentina. Tel.: +54 11 4964 8264; fax: +54 11 4508 3645.
E-mail address: rcampos@ffyb.uba.ar (R. Campos).

http://dx.doi.org/10.1016/j.meegid.2014.05.008
1567-1348/

migrations at the inter-community level (Lemey et al., 2009). With this information, the viral epidemics could be better evaluated and potential public health risks may be anticipated (Kühnert et al., 2011).

In this work, sequences of the two most prevalent HCV subtypes in Argentina from different localizations of the central region are analyzed to reconstruct a possible scenario of the internal transmission of the virus in this region. This information is valuable since it allows the analyses of past and recent events that may have modulated the transmission paths and thus contribute to improve the implementation of public health policies.

2. Materials and methods

2.1. Ethical statement

Written informed consent to participate in this study was obtained from the patient. The study protocol was approved by the ethics committee of the "Facultad de Farmacia y Bioquímica de la Universidad de Buenos Aires": record number 732575/2010, in accordance with the 1975 Helsinki Declaration.

2.2. Population study

A retrospective study was carried out using sequences from patients infected with HCV of subtype 1a and 1b from different cities in the central region of Argentina.

Patients attending to health-care centers in Buenos Aires and Córdoba for HCV diagnosis were included in this study. The inclusion criterion was presence of viral RNA for HCV-1a or 1b.

Additional sequences were obtained from previous published epidemiological, phylogentical or coalescence studies carried out in the Argentina (Golemba et al., 2010, 2013; Culasso et al., 2012a,b) provided that the analyzed genomic region was the same and there were information about the collection place and date.

The studied population included male and female patients with different status for HIV infections. The basic data of each patient (gender, HIV status, reference) are depicted in the Supplementary data I.

2.3. Sequences

This study included both, new and downloaded sequences, that represent all the E2 sequences (nucleotides 1479–2036 in AF009606) for HCV-1a and HCV-1b that were collected in Argentina.

2.3.1. HCV-1a sequences

The new sequences from Buenos Aires and Córdoba were deposited in the Genbank under the accession numbers KJ467047–KJ467054 and KJ467064–KJ467072 respectively.

The downloaded ones were: JQ272915, JQ272916, JQ272919–JQ272921, JQ272923–JQ272925 and JQ272927–JQ272929 from Mar del Plata (Culasso et al., 2012b); AY876394 from Buenos Aires City (Alfonso et al., 2005) and KJ094593 and KJ362056 from patients with hemophilia in Buenos Aires City (Culasso et al., 2012a, 2014).

2.3.2. HCV-1b sequences

The new sequences from Córdoba were deposited in the Genbank under the accession numbers KJ467055–KJ467063.

The downloaded ones were: HM055629–HM055683 from Wheelwright (Golemba et al., 2010); KF733833–KF733863 from Buenos Aires (Golemba et al., 2013); KF733864–KF733900 from O'Brien (Golemba et al., 2013) and JQ272930 from Mar del Plata (Culasso et al., 2012b).

All the above-mentioned sequences were named according to the place of collection, then underscore, then the isolation (or sample) name, then underscore, and finally the year of collection in four digit format. The abbreviations used for the place of collection were: "BsAs" for Buenos Aires City, "Cba" for Córdoba City, "CdE" for Cruz del Eje, "MDQ" for Mar del Plata, "O" for General O'Brien, "ViA" for Villa Allende, "ViD" for Villa Dolores, "ViM" for Villa María and "W" for Wheelwright.

2.4. Phylogenetic analyses

For the phylogenetic analyses, additional sequences were downloaded from Genbank. After performing BLAST searches (nr database) the 10 most related sequences to each Argentinean ones were retrieved. Those sequences were renamed with their Genbank accession number, and with additional data such as the country of collection and the isolation name. Duplicated or highly similar cloned sequences were removed. No new sequences from Argentina were retrieved by this procedure.

The sequences were aligned with MUSCLE v 3.8.31 (Edgar, 2004) and were renamed and visually inspected with BioEdit v 7.2.3 (Hall, 1999). The best-fit nucleotide substitution models were selected by the Akaike information criterion implemented in jModelTest v 2.1.3 (Darriba et al., 2012) for each dataset. Then, for those datasets containing less than 200 sequences, the maximum likelihood tree was obtained by heuristic search with PhyML v 3.1 (Guindon and Gascuel, 2003), and the branch support was assessed by non-parametric bootstrapping (1000 pseudoreplica). For the large HCV-1b dataset (+200 sequences) the maximum likelihood tree and the bootstrapping was carried out with a faster program: RAxML v7.2.8 (Stamatakis, 2006). The trees were visualized and prepared for publication with FigTree v 1.4 (available at http://tree.bio.ed.ac.uk/software/figtree/).

2.5. Bayesian phylogeographical analyses

The t_{MRCA}, the population dynamic and the most probable pattern of migrations were jointly estimated in a Bayesian framework. This analysis was carried out with the program BEAST v1.7.5 (Drummond and Rambaut, 2007) setting up a continuous phylogeographical model (relaxed random walks with Cauchy distribution) with a flexible population dynamic (General Markov Random Fields Skyride) and a relaxed molecular clock (uncorrelated log-normal). The analyzed datasets contained the samples sequence, year of collection and geographical location (GPS data). For each dataset, the model of nucleotide substitution was selected in the same way that for the phylogenetic analysis with jModelTest, whereas the prior for the substitution rate was inferred with the help of an external dataset (Supplementary data II).

For the Bayesian analysis, Monte Carlo Markov Chains (MCMC) were, initially, run for 5×10^7 generations sampling each 5000 generations. Then the samples were examined with Tracer 1.5 (available at http://tree.bio.ed.ac.uk/software/tracer/) to assess whether they have achieved convergence by observing if the Effective Sample Size (ESS) values where greater than 200 and by the visual inspection of lack of tendencies in the traces of each parameter. For HCV-1a (31 sequences) the analysis converged with the initial parameters, but in the case of the HCV-1b dataset (113 sequences), the chain failed to achieve convergence even after being run for 1×10^9 generations and longer runs were computationally prohibitive. To overcome this issue, the analysis was partitioned: first, the geographical data was removed to run a GMRF Skyride analysis of the full HCV-1b dataset. This simplified analysis achieved convergence after 1×10^8 generations. The t_{MRCA}, the

maximum clade credibility tree (MCCT) and the population dynamic plot presented in the results section corresponds to this analysis. Then, by taking into account the topology of the MCCT, the dataset was reduced by leaving only 6 sequences to represent Wheelwright and 6 for O'Brien creating a reduced HCV-1b dataset (53 sequences). The reduced dataset was analyzed under the GTR + Γ + I model (selected by jModelTest analysis). It achieved convergence after 5×10^7 generations and showed similar results for t_{MRCA}, MCCT and population dynamics that the full dataset but also allowed to perform the phylogeographic analysis.

The MCCT and the trees sampled in the MCMC were analyzed with SPREAD v 1.06 (available at http://www.kuleuven.be/aidslab/phylogeography/SPREAD.html) to represent the trees space-temporally in a map, making possible to infer the pattern of spreading of the viral lineages in the central region of Argentina. The resulting time and spatial annotated trees (Supplementary data III) where visualized with GoogleEarth.

3. Results and discussion

3.1. The HCV-1a and 1b isolations from Argentina are not monophyletic

To assess the diversity of the HCV circulating in Argentina a phylogenetic analysis of the Argentinean sequences and its closest related ones deposited in Genbank was carried out.

The BLAST searches with the 31 HCV-1a sequences from Argentina retrieved 81 sequences from Genbank showing identity values ranging from 96.44% to 89.07%.

The phylogeny inferred by maximum likelihood (GTR + Γ + I model) for HCV-1a could be described as the joining of two subtrees (clades ST-1 y ST-2 in Fig. 1A), forming the two commonly observed main clusters for this subtype (Pickett et al., 2011). Five out of 31 sequences from Argentina were included into ST-1 and twenty-six into ST-2.

Twenty-two out of 26 sequences from Argentina included in ST-2 grouped as an unsupported cluster (Argentinean clade). The Argentinean clade contained sequences from different cities. Despite the low bootstrap support value (11%), this cluster may represent the circulation of a country specific HCV-1a variant along with the worldwide distributed ones and may reflect the sharing of some feature, such as the route of transmission rather than geographical (local scale) isolation as it was recently suggested by the analysis of NS5B sequences (del Pino et al., 2013). Within the remaining four Argentinean sequences in ST-2, two from Córdoba City formed a highly supported cluster while two from Buenos Aires where intermingled. According to this phylogenetic topology, the Argentinean sequences do not form a monophyletic cluster, supporting the epidemiologic hypothesis of multiple introductions of HCV-1a to the country.

For HCV-1b, the BLAST searches with the 133 sequences from Argentina retrieved 143 additional sequences from Genbank showing identity values ranging from 93.61% to 83.90%.

The phylogeny (Fig. 1B) inferred by maximum likelihood (GTR + Γ + I model) for this dataset was represented as a tree with long branches, which showed several unrelated large clusters (+10 sequences). Three of these clusters were supported by bootstrap analysis (supported groups: SG-1, 2 and 3 with 100%, 93% and 96% of bootstrap values respectively). While SG-2 was composed by related sequences from Genbank, the other two clusters involved sequences from Argentina: SG-3 included 36 out of the 37 sequences from O'Brien and SG-1 included 14 sequences from different locations. Another large, although with low bootstrap support (19%), group (unsupported group 1, UG-1) contained all (55) sequences from Wheelwright and one sequence from Córdoba City. UG-1 was sister to a four-sequence cluster from Buenos Aires.

In addition, many small supported clusters, represented by two or three sequences, were observed. All of this grouplets were part of some of the already described larger groups. Finally, 27 sequences were found intermingled throughout the tree. As in the case of HCV-1a, the 1b sequences from Argentina do not form a monophyletic group and, once again, this finding supports the hypothesis of multiple introductions of HCV-1b to Argentina. As in the case of HCV-1a, the clusters of Argentinean HCV-1b sequences may reflect the circulation of country specific strains. However, in this case there are two types of clusters: the geographically delimited ones (SG-3 and UG-1) that supports the epidemiologic hypothesis of one event of introduction and local-level circulation of strains in O'Brien and in Wheelwright (Picchio et al., 2006; Golemba et al., 2010), and the country-wide SG-1 that may be related with some unknown factor such as the route of transmission. In fact, the seven sequences from Buenos Aires found in the SG-1 cluster were already described as "Buenos Aires 7+" Cluster (BA7+) in a previous work (Golemba et al., 2013) which in turn may represent a treatment susceptible strain of HCV-1b circulating in Argentina (Di Lello et al., 2008).

3.2. HCV-1b was dispersed before HCV-1a

To describe the population dynamics of the HCV in the central region of Argentina, coalescence analyses of the HCV-1a and 1b were carried out using the GMRF Skyride (GMRFS) as population model. The substitutions rates priors were set as normal distributions with a mean (and standard deviations) of 3.41×10^{-3} (4.53×10^{-4}) s/s/y for HCV-1a and 3.29×10^{-3} (6.36×10^{-4}) s/s/y for HCV-1b. Those rates were estimated by the analysis of external calibration datasets (Supplementary data II) resulting in the range of previous estimations for almost the same HCV genomic region (Magiorkinis et al., 2009; Gray et al., 2013). Under this setup, the analyses produce plots of the sequence diversity (scaled as population size $\langle\langle N_e \times \tau \rangle\rangle$) versus time. For simplicity, these plots are assumed to represent the population dynamic of the virus, but it is worth noting that both, the actual population size (i.e. the number of infected hosts) and the ratio of the transmission events may influence the $N_e \times \tau$ value (Frost and Volz, 2010).

The Bayes factor test for these analyses supported the uncorrelated lognormal relaxed molecular clock for both subtypes. The models of nucleotide substitution selected by jModelTest were TIM2 + Γ + I and TPM2uf + Γ + I for HCV-1a and HCV-1b datasets respectively.

The estimated t_{MRCA}, expressed in calendar years, for HCV-1a was 1962 (between 1943 and 1977) whereas for HCV-1b it was 1929 (between 1895 and 1953). In addition, the resulting demographic history represented by the GMRFS Plots showed different profiles for each subtype. A quick growth in the population size between ~1985 and ~2000 was observed for HCV-1a (Fig. 2A) whereas a slow but steady growth in the population size between 1929 and 1980 was observed for HCV-1b (Fig. 2B). A more detailed picture of the dispersion process could be achieved by examining the time-annotated trees for each subtype (Fig. 3A and B for HCV-1a and 1b respectively). For HCV-1a, it can be observed that the diversification of ST-1 and ST-2 started between 1970 and 1980. For ST-2, the diversification occurred quickly in the second half of the '80s decade generating highly related sequences that turned out to form the unsupported Argentinean clade in the phylogenetic analysis. ST-1 showed a more gradual diversification that started in 1980 but continued up to the first half of the '90s decade.

For HCV-1b the picture is more complex because of the existence of several supported and unsupported clusters. The intermingled sequences had the oldest ancestors dating up to about 1900.

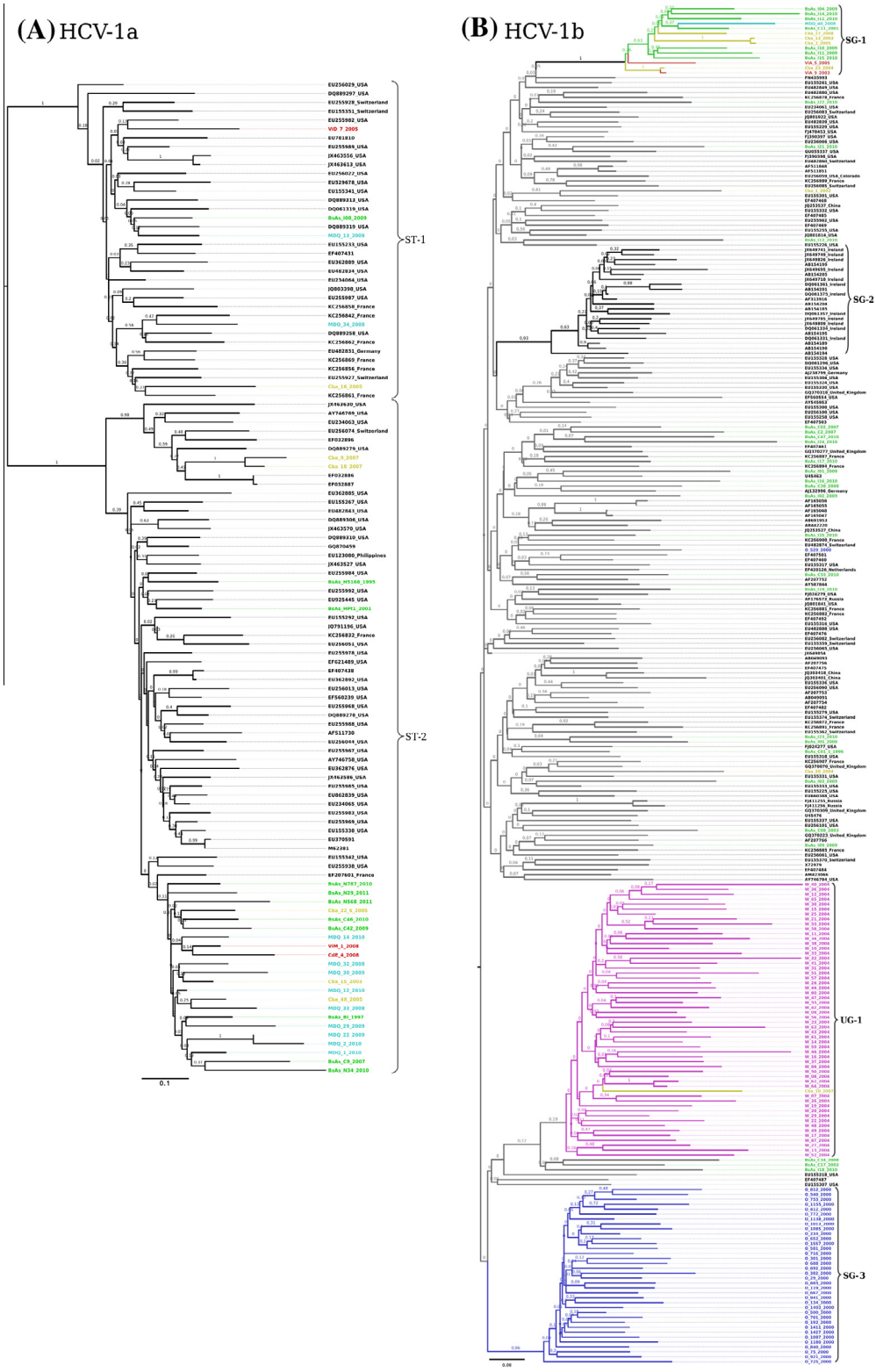

Fig. 1. Maximum likelihood tree for the E2 regions of HCV-1a (A) and HCV-1b (B) obtained with PhyML and RAxML respectively and using GTR + Γ + I as a model of nucleotide substitution for both datasets. The colored names represent sequences from different locations in Argentina: green for Buenos Aires City; gold for Córdoba City; blue for General O'Brien, cyan for Mar del Plata; magenta for Wheelwright and red for other cities and towns of Córdoba Province. The names in black represent sequences downloaded from Genbank after performing BLAST searches with the Argentinean sequences. The numbers above the branches represent the bootstrap proportion (over 1000 pseudoreplica). Highlighted groups: (A) subtrees 1 and 2 (ST-1 and 2): the two already described variants of HCV-1a. (B) UG-1: group with low support value (<70%); SG-1 to 3: groups with high support value (>90%). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

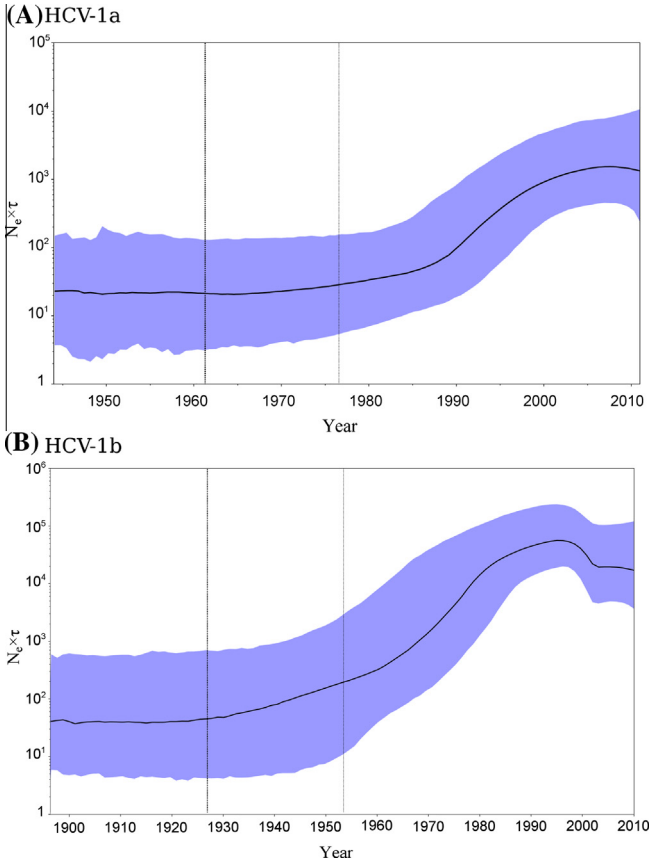

Fig. 2. Bayesian GMRF plots for demographic history reconstruction using E2 sequences for HCV-1a (A) and HCV-1b (B). *X* axis: date in calendar years; *Y* axis: estimated population size (effective number by generation time); bold dashed line: median time of most recent common ancestor (t_{MRCA}); light dashed line: upper high posterior probability density 95% (HPD95%) of t_{MRCA}; bold line: estimated population size (median); shaded area: HPD95% of the estimated population size.

Interestingly, the three clusters with Argentinean sequences, SG-1, SG-3 and UG-1, had estimated t_{MRCA} means within a relatively short time span, between ~1944 and ~1958. These clusters also showed similar dispersion period, since the diversification occurred in the first 15 years.

The t_{MRCA} of SG-1 was estimated to be a 1958 (between 1942 and 1977). This estimation is almost equal to those previously obtained for the coalescence analysis of the NS5B region with 16 Buenos Aires sequences included in the BA7+ cluster (Golemba et al., 2013), which resulted in a date of the MRCA of 1959 (between 1941 and 1978).

The t_{MRCA} for SG-3 sequences was estimated to be 1949 (between 1930 and 1969) which is earlier, but still compatible, with previous estimations made using only O'Brien sequences: 1961 (between 1949 and 1971) (Golemba et al., 2013). The differences may be related with the way in which the substitutions rates were estimated.

The UG-1 group had an ancestor dating 1944 (between 1922 and 1965). This result is coincident with that obtained by Golemba et al. (2010), in a similar coalescence analysis, using only sequences from Wheelwright with a t_{MRCA} in 1948 (between 1916 and 1981).

In summary, the analysis of the GMRFS Plots for HCV-1b suggested the superposition of two processes: (1) a gradual, and more ancient, process of diversification of those lineages of HCV-1b that now appear as intermingled sequences in the phylogenetic analysis, and (2) different outbreak-like processes occurred in a more defined geographical areas like Wheelwright (UG-1) and O'Brien (SG-3) as well as in the whole central region of Argentina (SG-1). The relatively narrow time-frame of the t_{MRCA} and diversification period of the outbreak-like processes suggested by the Bayesian analysis may be related to some change in the host behavior such as access to the medical practices, blood transfusions, surgical practices, or popularization in the use of injectable drugs (Magiorkinis et al., 2009; Alter, 2011; Gray et al., 2013).

3.3. Phylogeographical analyses

To describe the migration process of the HCV over the central region of Argentina a Bayesian phylogeographical analysis of the viral sequences was carried out using the continuous relaxed random walk model of viral dispersion.

3.3.1. HCV-1a

The history behind the dispersion of the analyzed samples was described by a complex model of intercity transmissions and local disseminations. To clearly explain these results, the HCV-1a population history (Fig. 2A) may be summarized as a three-staged process. In the first stage, from 1960 to 1985, the population was characterized by a slow but steady increment in size. In the second stage, between 1985 and 2005, the population was characterized by an exponential increment in size. Finally in the third stage, from 2005 up to 2011, the profile showed a constant size population.

For the first stage, the Bayesian analysis suggested a place for the MRCA (t_{MRCA} ~ 1962) in the middle east of Córdoba province, but provides a disjoint range that included both Córdoba and Buenos Aires cities (Fig. 4C). The first intercity transmission appeared to have occurred toward Córdoba City, but shortly after that (between 1965 and 1975) the transmission also occurred from Córdoba City to Buenos Aires City (Fig. 4C and D). In the beginning of the second stage (~1985) the analysis suggested transmissions between the cities of Córdoba and Mar del Plata. At the same time, local dissemination of HCV-1a was inferred for Córdoba and Buenos Aires cities (Fig. 4D). Then, by the end of this stage (between 1995 and 2005) exchanges between Mar del Plata and Buenos Aires and between Córdoba City and other places of the Córdoba province were inferred (Fig. 4E). The third stage was characterized by city-level dissemination of HCV (Fig. 4F).

These results suggest an association between the population size of HCV-1a and the dissemination process. While the local scale dissemination appeared to be related to periods of little or no expansion of the population, the colonization of new places (intercity transmissions) appeared to have driven the increment in the population size.

3.3.2. HCV-1b

The HCV-1b population history (Fig. 2B) may be summarized as a two-staged process. The first stage started with the t_{MRCA} around 1920 and showed a gradual increase in the population size (diversity) up to ~1990. The second, and final, stage showed a plateau (or even an unsupported decrease) in the population size from ~1990 up to 2010.

For HCV-1b the phylogeographical analysis suggested that the initial steps in the dissemination occurred in Buenos Aires City between the t_{MRCA} ~1920 until 1940 (Fig. 4A). Shortly after, HCV started to disseminate from Buenos Aires, reaching Wheelwright in ~1950, O'Brien in ~1960 (Fig. 4B) and Córdoba City in ~1970 (Fig. 4C and D). During this period and up to 1990 local (city-level) dissemination of HCV occurred. Then, between 1990 and 2000 new transmissions were suggested from Buenos Aires to Córdoba, O'Brien and Mar del Plata and from Wheelwright to Córdoba (Fig. 4E). Finally, in the last decade (2000–2010) no noticeable intercity transmissions of HCV were inferred (Fig. 4F).

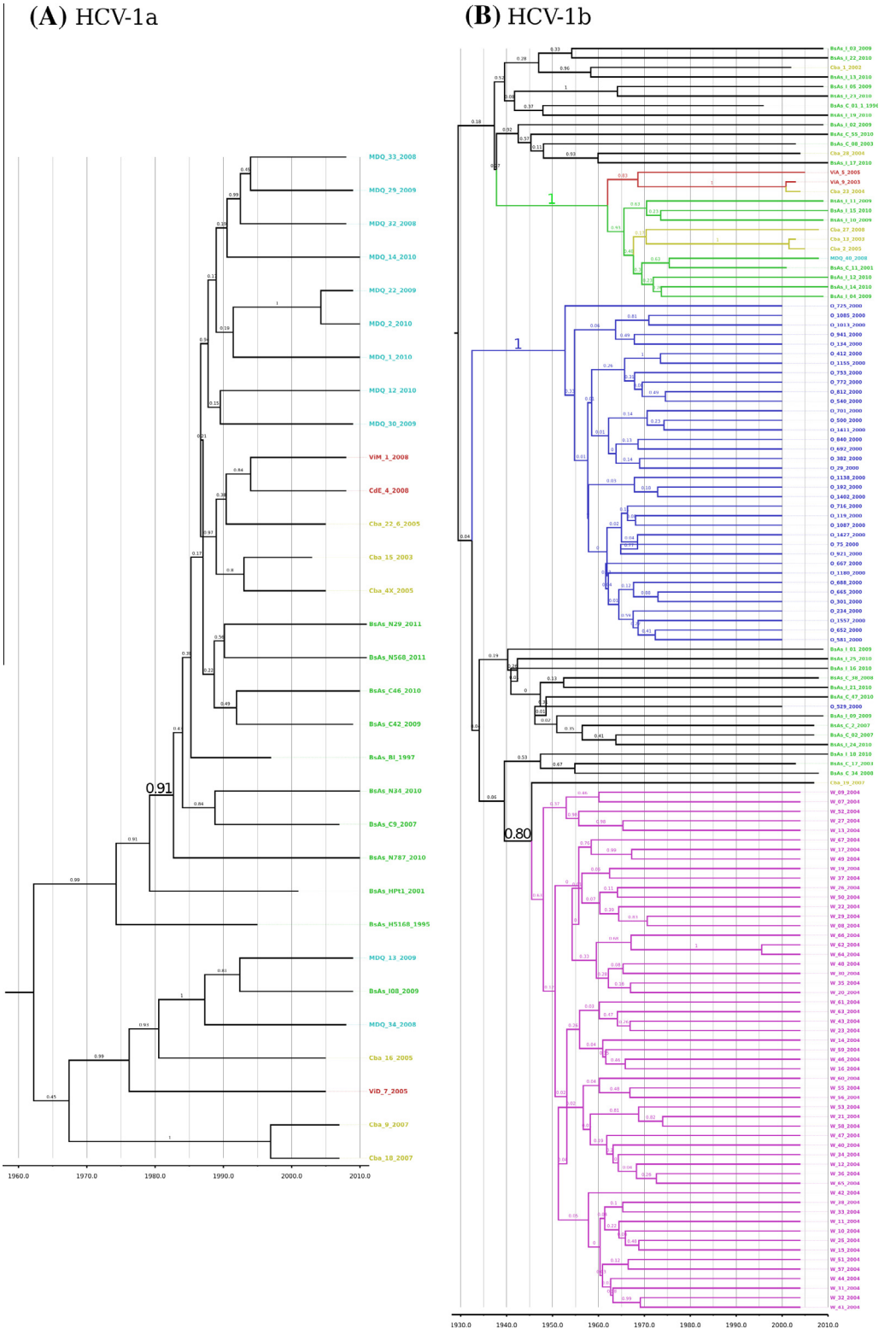

Fig. 3. Time-annotated maximum clade credibility trees of the E2 region of HCV-1a (A) and HCV-1b (B) in Argentina. The BEAST XML files were manually modified to set up the model selected for HCV-1a and HCV-1b datasets: TIM2 + Γ + I and TPM2uf + Γ + I respectively. The number on the right side of each node is the posterior probability of the clade. The posterior values for the groups referenced in the text were represented with a large font size. The *X* axis represents the time in calendar years.

Contrastingly to what was observed for HCV-1a, the changes in HCV-1b population size appeared not to be associated with intercity transmission or local dissemination events since both may have contributed to the gradual increase in the population size observed in the GMRFS Plot from the ~1920 up to ~1990. Also, both kinds of events were suggested to occur during the final constant/decreasing phase of the Plot (~1990–2010).

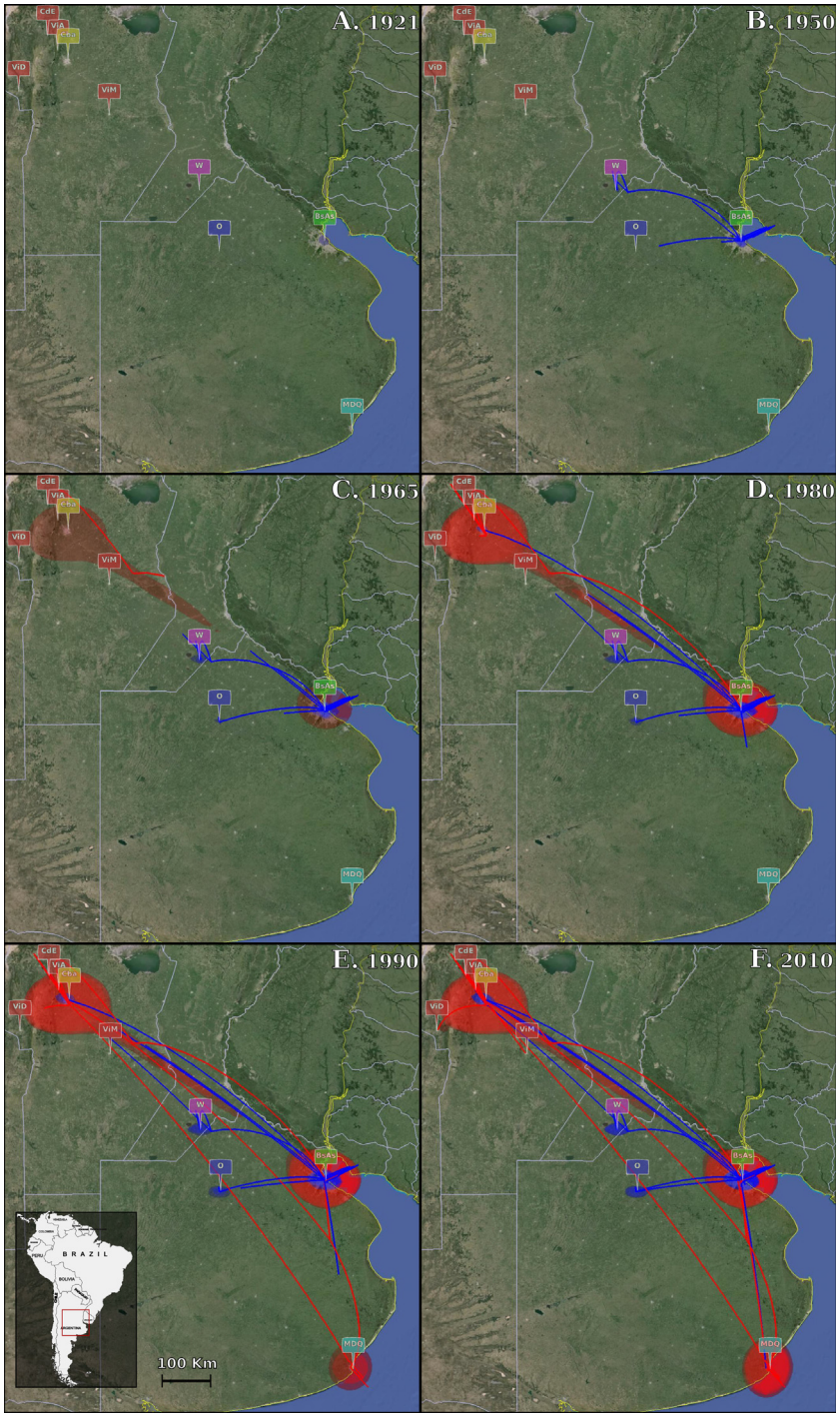

Fig. 4. Summary of the time-slices trees phylogeographical analysis. The same central region of Argentina is represented at different times: (A) 1921; (B) 1950; (C) 1965; (D) 1980; (E) 1990 and (F) 2010. The phylogenies of HCV-1a (red) and HCV-1b (blue) are plotted according to the inferred location of the ancestors at the respective times. The red and blue shadows over the maps correspond to the ancestor location HPD 80% for HCV-1a and HCV-1b respectively. Place markers: BsAs: Buenos Aires City; Cba: Córdoba City; CdE: Cruz del Eje; MDQ: Mar del Plata; O: General O'Brien; ViA: Villa Allende; ViM: Villa María; ViD; Villa Dolores; W: Wheelwright. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.4. The molecular epidemiology in the historical context

To test the credibility of the obtained estimations, the results of the coalescence and phylogenetic analyses were discussed together along with relevant historical events of the central region of Argentina.

Phylogenetic analysis suggested that multiple viral introductions could explain a diverse ancestral population (~1920 for HCV-1b) in Buenos Aires City. The intense immigration policies carried out in the ending of the nineteenth century may have propitiated this feature of the early HCV-1b population. The subsequent dissemination of HCV-1b around 1940, occurred in a

relatively short period of time and may be explained by several changes in the host behavior that may have influenced the HCV and other blood-borne pathogen transmissions. However, due to the lack of full epidemiological information for all the samples, none of them could be proved or ruled out.

It is worth noting that, at least for Argentina, the suggested time of dissemination coincides with the development of the country (Salvatore, 2009), the growth in the free public health system (Isuani and Mercer, 1988) and the treatment with transfusions, and plasma derivatives, of a endemic disease (*Junin* virus; Enria et al., 2008). Additionally, other factors as dental practices, extra-hospital vaccinations and injectable treatments as well as the abuse of recreational drugs may have also propitiated the expansion of HCV-1b in this period in the same way that it was suggested for other countries (Alter, 2011).

Although the estimated t_{MRCA} of HCV-1a also falls in the same period depicted above, it may not be related to the introduction of this viral strain in the country since the current sequences form a polyphyletic group.

For both subtypes, the diversification and some internal transmissions occurred during the 1970 and 1980. By 1990, the HCV-1b viral population reaches a plateau until the present time, but for HCV-1a, between 1985 and 2005 the viral population suffered an exponential growth. The different population dynamics and time of dissemination exhibited by HCV-1a and 1b, were already described in world-wide datasets of these subtypes and it was suggested that such differences may be related with the route of transmission (Magiorkinis et al., 2009). In Argentina, as it is observed in several countries, HCV-1a and HCV-3a were associated with transmission in intravenous drugs users (Quarleri et al., 2007; Alter, 2011), whereas there is no conclusive information for HCV-1b which may have transmitted by the same or by other means. In summary, our results suggest that the differences in the subtypes population dynamic may be the result of their different histories of transmission. The early introduction of HCV-1b in many countries (including Argentina) allowed its expansion when the changes in the host behavior favored the transmission of blood borne pathogens, whereas other factors, such as the popularization of drug abuse, may have contributed to the latter dissemination of HCV-1a.

4. Conclusions

The phylogeographic analyses provided useful information for making hypothesis about the spreading of some infectious diseases. In the case of HCV-1a and 1b, the molecular epidemiology shows remarkable coincidences with known historical events of the country. However, two considerations may be done. First, the lack of full epidemiological information about the samples and the inexistent randomized epidemiological studies precludes the correlation of any historical event with the viral population dynamic. Second, the results only represent the history of the samples, which may not be the truth for the real whole viral population. Poor sampling, lineage extinctions, methodological issues (i.e. primer design), among other factors may bias the results.

Based on the results, a hypothetical history of the spreading of HCV-1a and 1b in the central region of Argentina was proposed.

For both subtypes diverse ancestral populations were inferred in the larger cities studied. Then, HCV was spread to the other cities. The analyses suggested that HCV-1b was introduced and dispersed first and it is currently in a stationary phase. On the other hand, HCV-1a was introduced later and is still in a growth phase.

Finally, while the transmission of HCV-1b appears to have been somehow prevented, the HCV-1a may still represent a concern in the public health. Further work should be carried out to address their current transmission rate (and its main transmission route) in the Argentinean population.

Acknowledgements

This study was carried out thanks to the financial support of the grants: PIP 112-200801-01169 (CONICET: Consejo Nacional de Investigaciones Científicas y Técnicas – Argentina), PICT 2006-00285 (ANPCyT: Agencia Nacional de Promoción Científica y Tecnológica – Argentina) and UBACyT 2008-B601 (UBA: Universidad de Buenos Aires – Argentina). The founders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.meegid.2014.05.008. These data include Google maps of the most important areas described in this article.

References

Alfonso, V., Mbayed, V.A., Sookoian, S., Campos, R.H., 2005. Intra-host evolutionary dynamics of hepatitis C virus E2 in treated patients. J. Gen. Virol. 86 (Pt 10), 2781–2786. http://dx.doi.org/10.1099/vir.0.81084-0.

Alter, M.J., 2011. HCV routes of transmission: what goes around comes around. Semin. Liver Dis. 31 (4), 340–346. http://dx.doi.org/10.1055/s-0031-1297923.

Ciccozzi, M., Ciccaglione, A.R., Lo Presti, A., Yalcinkaya, T., Taskan, Z.P., Equestre, M., Costantino, A., Bruni, R., Ebranati, E., Salemi, M., Gray, R., Rezza, G., Galli, M., Zehender, G., 2011. Reconstruction of the evolutionary dynamics of the hepatitis C virus 1b epidemic in Turkey. Infect. Genet. Evol. 11 (5), 863–868. http://dx.doi.org/10.1016/j.meegid.2011.02.006.

Culasso, A.C., Aloisi, N., Candela, M., Campos, R.H., Baré, P., 2012a. Intrafamilial transmission of hepatitis C virus in patients with severe haemophilia A. Haemophilia 18 (5), 824–827. http://dx.doi.org/10.1111/j.1365-2516.2012.02792.x.

Culasso, A.C., Elizalde, M., Campos, R.H., Barbini, L., 2012b. Molecular survey of hepatitis C virus in the touristic city of Mar del Plata, Argentina. PLoS ONE 7 (9), e44757. http://dx.doi.org/10.1371/journal.pone.0044757.

Culasso, A.C.A., Baré, P., Aloisi, N., Monzani, M.C., Corti, M., Campos, R.H., 2014. Intra-host evolution of multiple genotypes of hepatitis C virus in a chronically infected patient with HIV along a 13-year follow-up period. Virology 449, 317–327. http://dx.doi.org/10.1016/j.virol.2013.11.034.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772. http://dx.doi.org/10.1038/nmeth.2109.

del Pino, N., Oubiña, J.R., Rodríguez-Frías, F., Esteban, J.I., Buti, M., Otero, T., Gregori, J., García-Cehic, D., Camos, S., Cubero, M., Casillas, R., Guàrdia, J., Esteban, R., Quer, J., 2013. Molecular epidemiology and putative origin of hepatitis C virus in random volunteers from Argentina. World J. Gastroenterol. 19 (35), 5813–5827. http://dx.doi.org/10.3748/wjg.v19.i35.5813.

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 (4), 619–627. http://dx.doi.org/10.1002/jmv.21112.

Di Lello, F.A., Neukam, K., Parra-Sanchez, M., Plaza, Z., Soriano, V., Cifuentes, C., Mira, J.A., Poveda, E., Pineda, J.A., 2013. Hepatitis C virus genotype 4 in Southern and Central Spain does not originate from recent foreign migration waves. J. Med. Virol. 85 (10), 1734–1740. http://dx.doi.org/10.1002/jmv.23657.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214. http://dx.doi.org/10.1186/1471-2148-7-214.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32 (5), 1792–1797. http://dx.doi.org/10.1093/nar/gkh340.

Enria, D.A., Briggiler, A.M., Sánchez, Z., 2008. Treatment of Argentine hemorrhagic fever. Antiviral Res. 78 (1), 132–139. http://dx.doi.org/10.1016/j.antiviral.2007.10.010.

Frost, S.D., Volz, E.M., 2010. Viral phylodynamics and the search for an 'effective number of infections'. Philos. Trans. R. Soc. Lond. B Biol. Sci. 365 (1548), 1879–1890. http://dx.doi.org/10.1098/rstb.2010.0060.

Golemba, M.D., Di Lello, F.A., Bessone, F., Fay, F., Benetti, S., Jones, L.R., Campos, R.H., 2010. High prevalence of hepatitis C virus genotype 1b infection in a small town of Argentina. Phylogenetic and Bayesian coalescent analysis. PLoS ONE 5 (1), e8751. http://dx.doi.org/10.1371/journal.pone.0008751.

Golemba, M.D., Culasso, A.C.A., Villamil, F.G., Baré, P., Gadano, A., Ridruejo, E., Martinez, A., Di Lello, F.A., Campos, R.H., 2013. Hepatitis C virus diversification in Argentina: comparative analysis between the large city of Buenos Aires and

the small rural town of O'Brien. PLoS ONE 8 (12), e84007. http://dx.doi.org/10.1371/journal.pone.0084007.
Gray, R.R., Tanaka, Y., Takebe, Y., Magiorkinis, G., Buskell, Z., Seeff, L., Alter, H.J., Pybus, O.G., 2013. Evolutionary analysis of hepatitis C virus gene sequences from 1953. Philos. Trans. R. Soc. Lond. B Biol. Sci. 368 (1626), 20130168. http://dx.doi.org/10.1098/rstb.2013.0168.
Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52 (5), 696–704. http://dx.doi.org/10.1080/10635150390235520.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
Isuani, E.A., Mercer, H., 1988. La fragmentación institucional del sector salud: pluralismo o irracionalidad. Centro Editor de América Latina, Buenos Aires, ISBN: 9789502515106.
Kershenobich, D., Razavi, H.A., Sánchez-Avila, J.F., Bessone, F., Coelho, H.S., Dagher, L., Gonçales, F.L., Quiroz, J.F., Rodriguez-Perez, F., Rosado, B., Wallace, C., Negro, F., Silva, M., 2011. Trends and projections of hepatitis C virus epidemiology in Latin America. Liver Int. 31 (Suppl. 2), 18–29. http://dx.doi.org/10.1111/j.1478-3231.2011.02538.x.
Kühnert, D., Wu, C.-H., Drummond, A.J., 2011. Phylogenetic and epidemic modeling of rapidly evolving infectious diseases. Infect. Genet. Evol. 11 (8), 1825–1841. http://dx.doi.org/10.1016/j.meegid.2011.08.005.
Lauer, G.M., Walker, B.D., 2001. Hepatitis C virus infection. N. Engl. J. Med. 345, 41–52.
Lavanchi, D., 2009. The global burden of hepatitis C. Liver Int. 29, 74–81.
Lemey, P., Rambaut, A., Drummond, A.J., Suchard, M.A., 2009. Bayesian phylogeography finds its roots. PLoS Comput. Biol. 5 (9), e1000520. http://dx.doi.org/10.1371/journal.pcbi.1000520.
Magiorkinis, G., Magiorkinis, E., Paraskevis, D., Ho, S.Y., Shapiro, B., Pybus, O.G., Allain, J.P., Hatzakis, A., 2009. The global spread of hepatitis C virus 1a and 1b: a phylodynamic and phylogeographic analysis. PLoS Med. 6 (12), e1000198. http://dx.doi.org/10.1371/journal.pmed.1000198.
Njouom, R., Nerrienet, E., Dubois, M., Lachenal, G., Rousset, D., Vessière, A., Ayouba, A., Pasquier, C., Pouillot, R., 2007. The hepatitis C virus epidemic in Cameroon: genetic evidence for rapid transmission between 1920 and 1960. Infect. Genet. Evol. 7 (3), 361–367.
Picchio, G.R., Baré, P.C., Descalzi, V.I., Bussy, M.V., Soria, S.M., Raffa, M.P., Mazzencio, N.E., Etchehun, S., Cámera, J.A., Mosier, D.E., Villamil, F.G., 2006. High prevalence of infection with a single hepatitis C virus genotype in a small rural community of Argentina. Liver Int. 26 (6), 660–665. http://dx.doi.org/10.1111/j.1478-3231.2006.01290.x.
Pickett, B.E., Striker, R., Lefkowitz, E.J., 2011. Evidence for separation of HCV subtype 1a into two distinct clades. J. Viral Hepat. 18 (9), 608–618. http://dx.doi.org/10.1111/j.1365-2893.2010.01342.x.
Pybus, O.G., Charleston, M.A., Gupta, S., Rambaut, A., Holmes, E.C., Harvey, P.H., 2001. The epidemic behavior of the hepatitis C virus. Science 292, 2323–2325.
Quarleri, J.F., Bolcic, F.M., Bouzas, M.B., Laufer, N., Gómez Carrillo, M., Mammana, L., Kaufman, S., Pérez, 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 (2), 76–83.
Ré, V., Contigiani, M., Fumiko, C., Yoshida, T., Lampe, E., 2007. Identification of hepatitis C virus subtype 2c by sequencing analysis in patients from Córdoba, Argentina. Mem. Inst. Oswaldo Cruz 102 (8), 995–998. http://dx.doi.org/10.1590/S0074-02762007000800016.
Ré, V.E., Culasso, A.C., Mengarelli, S., Farías, A.A., Fay, F., Pisano, M.B., Elbarcha, O., Contigiani, M.S., Campos, R.H., 2011. Phylodynamics of hepatitis C virus subtype 2c in the province of Córdoba, Argentina. PLoS One 6 (5), e19471. http://dx.doi.org/10.1371/journal.pone.0019471.
Salvatore, R.D., 2009. Stature growth in industrializing Argentina: the Buenos Aires industrial belt 1916–1950. Explor. Econ. Hist. 46, 70–92. http://dx.doi.org/10.1016/j.eeh.2008.02.002.
Simmonds, P., 2004. Genetic diversity and evolution of hepatitis C virus – 15 years on. J. Gen. Virol. 85 (Pt 11), 3173–3188.
Smith, D.B., Bukh, J., Kuiken, C., Muerhoff, A.S., Rice, C.M., Stapleton, J.T., Simmonds, P., 2014. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and assignment web resource. Hepatology 59 (1), 318–327. http://dx.doi.org/10.1002/hep.26744.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22 (21), 2688–2690. http://dx.doi.org/10.1093/bioinformatics/btl446.
Zehender, G., Sorrentino, C., Lai, A., Ebranati, E., Gabanelli, E., Lo Presti, A., Vujoševic, D., Lauševic, D., Terzić, D., Shkjezi, R., Bino, S., Vratnica, Z., Mugosa, B., Galli, M., Ciccozzi, M., 2013. Reconstruction of the evolutionary dynamics of hepatitis C virus subtypes in Montenegro and the Balkan region. Infect. Genet. Evol. 17, 223–230. http://dx.doi.org/10.1016/j.meegid.2013.04.003.