

Molecular Epidemiology of a Hepatitis C Virus Outbreak in a Hemodialysis Unit

Maria Alma Bracho,¹ María José Gosalbes,¹ David Blasco,² Andrés Moya,¹ and Fernando González-Candelas^{1*}

Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, València, Spain,¹ and Direcció General de Salut Pública, Conselleria de Sanitat i Consum, Generalitat Valenciana, València, Spain²

Received 4 November 2004/Returned for modification 1 February 2005/Accepted 24 February 2005

We analyzed a hepatitis C virus (HCV) transmission case in the hemodialysis unit of a private clinic by sequencing two genome regions of virus isolates from a number of patients attending this unit and some external controls. The analysis of 337 nucleotides (nt) in the NS5B region did not provide enough resolution to ascertain which patients were actually involved in the outbreak and the potential source. Nevertheless, this region allowed the exclusion of several patients as putative sources of the transmission case based on their genotypes and phylogenetic relationships. On the other hand, the analysis of several 472-nt-long clone sequences per sample in a more rapidly evolving region of the HCV genome, coding for the envelope proteins and encompassing hypervariable region 1, allowed us to establish the existence of at least two independent transmission events involving two different source patients and three recipients. The direction of the transmissions was further corroborated by different measures of genetic variability within and among samples.

During the spring of 2002, three patients who attended the hemodialysis unit of a private clinic in Vinaròs (Castelló, Spain) tested positive for hepatitis C virus (HCV) despite having tested negative shortly before. Since the same unit was regularly attended by other HCV-positive patients and given that the hemodialysis procedure has a high risk for HCV transmission (10, 13), it was necessary to study whether these patients had been infected in the clinic and, eventually, to determine the source of infection in order to facilitate the adoption of further safety measures to prevent any new HCV transmission.

Usually, molecular epidemiology analysis is based on the sequencing and comparison of a single product in a relatively conserved, i.e., slowly evolving, region of the target genome, such as NS5B or core in HCV (3, 12). This is a valid approach as long as there has been enough time for nucleotide differences to accumulate so that differentiation among ancestral and derived genomes has occurred. However, given the extraordinarily high evolutionary rate of HCV in some regions, such as hypervariable region 1 (HVR1), located in the amino terminus of the envelope 2 protein gene, it is possible to gain much better resolution of the evolutionary, and consequently epidemiological, relationships by using sequence information from these regions for cases in which a short time has elapsed between transmission from the source and sampling for analysis (2, 4, 6, 15).

In this report, we provide an illustrative case in which the analysis of a fast-evolving viral region allows better elucidation of a transmission case than the use of a single sequence from a more slowly evolving genome region.

* Corresponding author. Mailing address: Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, Apartado Oficial 22085, 46071 Valencia, Spain. Phone: 34 963543653. Fax: 34 96 3543670. E-mail: fernando.gonzalez@uv.es.

MATERIALS AND METHODS

Samples. Serum samples were obtained from nine HCV-infected patients who attended the hemodialysis clinic and from four unrelated HCV-positive patients (Table 1) from the nearest reference hospital (Hospital General de Castelló) to be used as local controls. The patients were diagnosed using an enzyme-linked immunosorbent assay (ORTHO HCV 3.0 ELISA TestSystem with Enhanced SAve; Ortho-Clinical Diagnostics, Cambridge, United Kingdom) and further confirmed by line immunoassay (INNO-LIA HCV II; Innogenetics N.V., Ghent, Belgium). The HCV genotypes of the samples were assigned by sequencing of an NS5B gene fragment and comparison with reference sequences from GenBank. Additional unrelated sequences from an independent molecular epidemiology study of HCV patients in our region (8, 20) were included in the analyses. Samples were taken in May 2002 and were stored frozen at -70°C until they were processed in July of the same year.

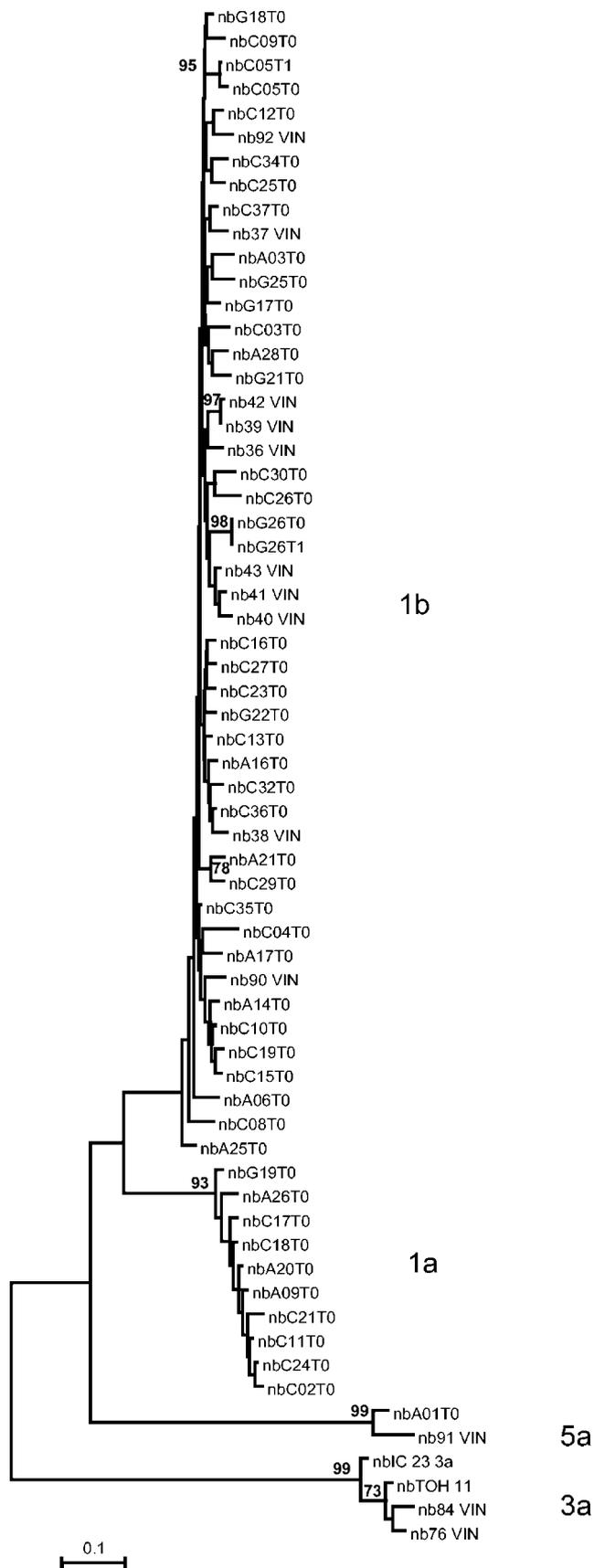
RNA extraction and RT-PCR. Viral RNA was obtained from 200 μl of serum for each sample using a High Pure Viral RNA kit (Roche Diagnostics GmbH,

TABLE 1. Serum samples received for study from HCV-positive patients attending the hemodialysis unit at the clinic where the suspected transmission case was reported

Patient	Sex	Geno- type	Date of HCV diagnosis ^b	Date of last HCV- negative test ^b	Attending the clinic since ^b
53876	Male	3a	03/10/01	02/05/01	20/01/00
36	Male	1b	PH	NA	23/01/02
37	Male	1b	PH	NA	05/06/02
38	Female	1b	PH	NA	05/08/00
39	Male	1b	PH	NA	28/09/01
40	Male	1b	14/05/02	11/01/02	05/07/00
41	Male	1b	09/07/02	25/06/02	13/01/97
42	Female	1b	11/04/02	11/01/02	20/06/01
43	Male	1b	PH	NA	01/10/99
29584 ^a	Unknown	3a	NA	NA	
29590 ^a	Unknown	1b	NA	NA	
29591 ^a	Unknown	5a	NA	NA	
29592 ^a	Unknown	1b	NA	NA	

^a Received from Servei de Microbiologia-Hospital General de Castelló and used as local controls.

^b Day/month/year. NA, not available; PH, previous to hemodialysis.



Mannheim, Germany). Reverse transcriptions (RT) were carried out in 20- μ l volumes containing 5 μ l of eluted RNA, 500 μ M of deoxynucleoside triphosphate (dNTP), 1 μ M of hexamers, 100 units of Moloney murine leukemia virus reverse transcriptase (Promega Corp., Madison, WI), and 20 units of rRNasin RNase inhibitor (Promega). Reaction mixtures were incubated at 42°C for 60 min, followed by 2 min at 95°C.

Amplification and direct sequencing of NS5B. Direct sequences of PCR products were obtained for a 337-nucleotide (nt)-long fragment of the NS5B gene. PCR was performed in a 50- μ l volume containing 5 μ l of RT product, 100 μ M of dNTP, 200 nM of each primer, and 2.5 units of *Taq* polymerase (Amersham Biosciences, Piscataway, NJ). Amplified products were purified with a High Pure PCR Products Purification kit (Roche). Direct sequencing of purified PCR products was performed on an 8- μ l volume, including 1.0 μ l of the PCR-amplified DNA, with the ABI PRISM BigDye Terminator v3.0 Cycle Sequencing Ready Reaction kit in an ABI 3700 automated sequencer (Applied Biosystems, Foster City, CA). Sequences were verified, and both strands were assembled using the Staden package (1). The sense and antisense primers used for amplification and direct sequencing of this region were 5'-TATGATACYCGCTGY TTYGACTC-3' and 5'-GTACCTRGTCATAGCCTCCGTGAA-3'.

Cloning and sequencing of E1-E2 region. A 472-nt fragment of the E1-E2 region containing HVR1 and HVR2 regions was amplified by nested PCR. The first amplification was performed in a 100- μ l volume containing 10 μ l of the RT product, 10 μ l of 10 \times PCR buffer, 200 μ M (each) dNTP, 400 nM (each) primer (sense, 5'-CGCATGGCYTGGGAYATGAT-3'; antisense, 5'-GGYSGTART GCCARCARTA-3'), and 2.5 U of *Pfu* DNA polymerase (Stratagene, La Jolla, CA). When necessary, a second PCR was performed with a nested sense primer (5'-GGGATATGATRATGAAATGGTC-3') and the same antisense primer indicated above. In all cases, PCR was performed in a Perkin-Elmer 2400 thermal cycler with the following thermal profile: 94°C for 3 min; 5 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 3 min; 35 cycles at 94°C for 30 s, 52°C for 30 s, and 72°C for 3 min; and a final extension at 72°C for 10 min. A single amplified product was observed after electrophoresis on a 1.4% agarose gel stained with ethidium bromide.

Amplification products were purified with a High Pure PCR Products Purification kit (Roche) and then directly cloned in EcoRV-digested pBluescript II SK(+) phagemid (Stratagene). Plasmid DNA was purified with a High Pure Plasmid Isolation kit (Roche). Recombinant clones were sequenced by the use of KS and SK primers (Stratagene) and the same procedure described above for the NS5B gene.

Sequence analysis. NS5B sequences were analyzed with a panel of 51 additional HCV sequences from the same genome region, including genotypes 1a (10 sequences), 1b (38 sequences), 3a (2 sequences), and 5a (1 sequence). Similarly, the 122 E1-E2 region sequences derived from recombinant clones corresponding to the 10 patients infected with HCV genotype 1b were analyzed, along with 73 sequences of the same genotype and HCV genome region.

For both HCV genome regions, multiple sequence alignments were obtained using ClustalW (19). The model that accounted best for the observed evolutionary pattern was ascertained using Modeltest 3.5 (11) and PAUP* v4.0b10 (18). A phylogenetic tree was obtained by maximum likelihood using PHYML (7) and employing the evolutionary model derived in the previous step. Support for the phylogenetic tree was obtained by bootstrap resampling (5) with 2,000 pseudoreplicates and neighbor-joining clustering (17) using the same evolutionary model as above, as implemented in MEGA2 (9).

Genetic variability estimates for sequences derived from related samples (see below) were obtained with DnaSP 3.95 (16).

Nucleotide sequence accession numbers. The HCV sequences obtained in this study have been deposited in GenBank under accession numbers AJ880694 to AJ880757 for NS5b region sequences and AJ866035 to AJ866183 for clones of the E1-E2 region.

FIG. 1. Maximum-likelihood phylogenetic tree obtained for the NS5B regions of 14 samples analyzed in this study (identified by suffix VIN) and 51 unrelated sequences. Among the latter, there are two pairs of sequences (nbC05T0-nbC05T1 and nbG26T0-nbG26T1) which were derived from the same patients at two time points separated by a 6-month interval. Nodes with bootstrap support higher than 70% are indicated. The scale bar represents genetic distance (substitutions per site).

RESULTS

The HCV genotype (Table 1) was determined by sequencing of an NS5B gene fragment and comparison with reference sequences of known genotype (Fig. 1). Eight of the nine sequences derived from patients attending the case clinic shared the same HCV genotype 1b, and the other was genotyped as 3a. Among the four control samples, two also harbored HCV-1b viruses. This information allowed us to exclude genotypes other than 1b as a potential source for the outbreak, and hence, the ensuing analyses were restricted to samples of genotype 1b.

The evolutionary model that best accounted for the data in the NS5B region corresponds to transversional model distance (14), with a gamma distribution accounting for heterogeneity in evolutionary rates among sites (Shape parameter = 0.5910) and a proportion of invariable sites (Pinvar = 0.3826). In the phylogenetic tree for this region (Fig. 1), two different monophyletic groups, each including three sequences from the case, can be observed. One group includes sequences from two of the patients whose seroconversion prompted the study (patients 40 and 41) and one patient previously known to be HCV infected (patient 43). The other group included the sequence from the third patient prompting the study (patient 42) and the sequences from two other known HCV-positive patients from the same clinic (patients 36 and 39). In this group, sequences from patients 39 and 42 were almost identical in this region and formed a well-supported subgroup (bootstrap support [BS] = 97%). Apart from the well-supported nodes defining the HCV genotypes and subtypes included in the study, the phylogenetic tree for the NS5B region presents only four nodes with bootstrap support higher than 70% (Fig. 1). Two of these correspond to two samples taken at different times (separated by 6 months) from the same patients (G26, BS = 98%, and C05, BS = 95%, respectively), one corresponds to two unrelated sequences (C29 and A21; BS = 78%), and the last one is the already-mentioned group comprising samples 39 and 42.

The analysis of the NS5B region does not provide a clear-cut answer to the relevant question of how many patients, if any, were infected with HCV at the case clinic. To obtain a better-resolved picture of the phylogenetic relationships among the involved patients, we proceeded by analyzing a more rapidly evolving region of the HCV genome. We determined the sequences of 122 cloned fragments derived from the 10 patients, 8 from the case clinic and 2 controls, in the E1-E2 region of HCV (Fig. 2). These sequences were analyzed with the sequences of 73 cloned fragments from the same region and genotype, a number of which were also derived from the same patient at different serial times (separated by 6 or 12 months). In this case, the best model of evolution corresponded to GTR (general time reversible) (14), with a gamma distribution ac-

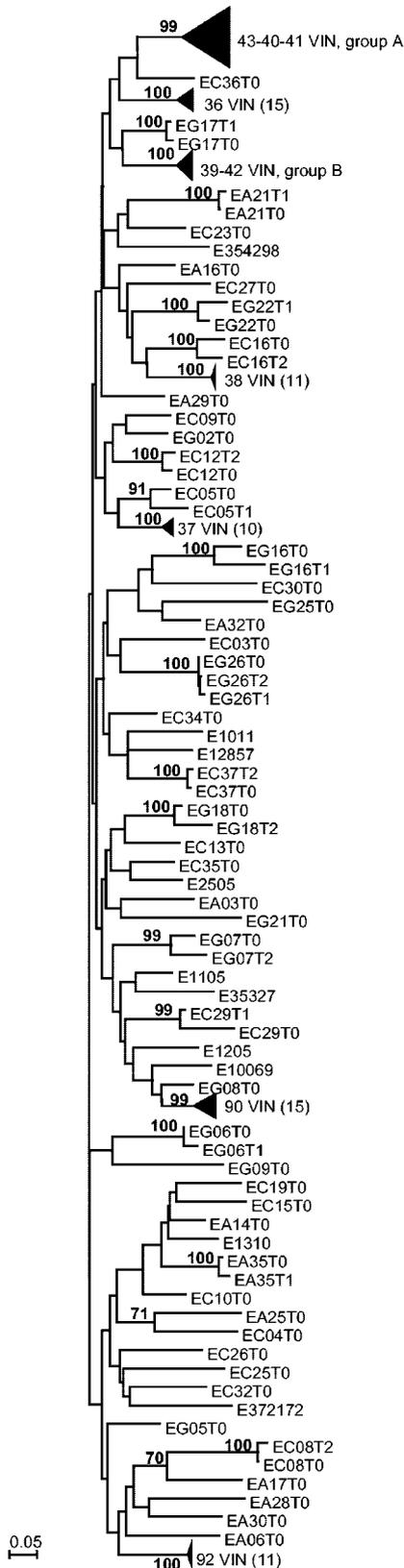


FIG. 2. Maximum-likelihood phylogenetic tree obtained with 122 clone sequences from the E1-E2 region derived from the 10 samples with genotype 1b included in this study and 73 unrelated sequences. Among the latter, there are several pairs of sequences which were derived from the same patients at two time points separated by a

6-month interval (T0, T1, and T2). Monophyletic clusters of clone sequences are represented by filled triangles whose areas are proportional to the genetic diversity among the sequences included therein. The number of sequences is reported in parentheses after the corresponding sample name. See Fig. 3 for more details of the groups labeled A and B. Nodes with bootstrap support higher than 70% are indicated. The scale bar represents genetic distance (substitutions per site).

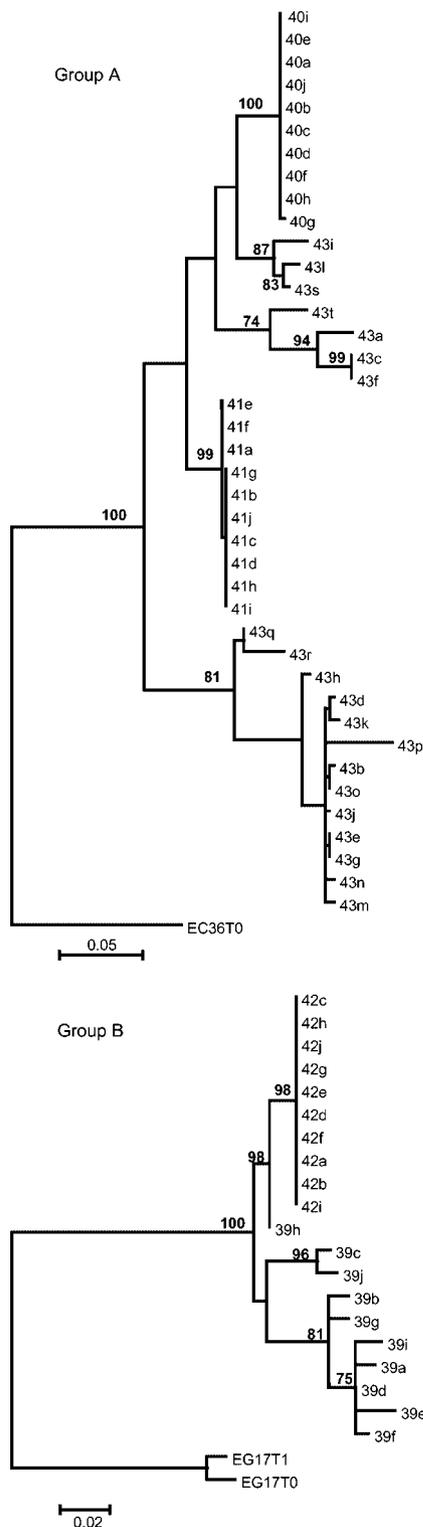


FIG. 3. Detailed representation of portions of the maximum-likelihood tree for E1-E2 sequences depicted in Fig. 2. (Top) Group A, including clone sequences from samples 40, 41, and 43. (Bottom) Group B, including clone sequences from samples 39 and 42. In both cases, the sister branch has been included as a reference. Nodes with bootstrap support higher than 70% are indicated. The scale bar represents genetic distance (substitutions per site).

counting for heterogeneity in evolutionary rates among sites (Shape parameter = 1.0455) and a proportion of invariable sites (Pinvar = 0.3358). The resulting maximum-likelihood phylogenetic tree is shown in Fig. 2. As in the previous tree, most phylogenetic groups do not reach a BS of >70%, and those that do can be divided into two categories. One category includes all pairs of serial samples from the same patient, with BSs of >90% and >99% in most cases; the second category includes nodes with 70% < BS < 90% that encompass samples from epidemiologically unrelated patients.

Sequences from the studied samples appear in two forms in the phylogenetic tree for the E1-E2 region. Samples 36, 37, 38, 90, and 92 show highly supported (BS > 99%) monophyletic clusters that include only sequences derived from each individual patient. However, sequences from samples 40, 41, and 43 on the one hand and 39 and 42 on the other are grouped in two separate, highly supported (BS = 100%) clusters, designated groups A and B, respectively. Within them, sequences from one patient each, 43 and 39, do not group monophyletically, as opposed to the sequences from the other patients in the corresponding groups (Fig. 3). In group A (Fig. 3), all the sequences from samples 40 and 41 are almost identical and form monophyletic groups with very high support (BS = 100 and 99%, respectively), whereas sequences derived from sample 43 are distributed in two main subgroups, one of them clearly related to sequences from the other two samples (BS = 81%). Furthermore, sequences from sample 43 are much more diverse than those from the other two patients. A similar pattern can be observed for the sequences of group B (Fig. 3), with all sequences being identical and forming a highly supported monophyletic group (BS = 98%) and with sample 39 presenting more diverse sequences that do not group in a single cluster, some of which are clearly related (occupying a basal position) to those from sample 42.

Table 2 presents a summary of the inpatient genetic variability derived from the comparison of clone sequences obtained for the E1-E2 regions of the 10 samples with HCV genotype 1b included in the study. Different measures of genetic variation are coincident in separating two clearly distinct groups, one comprising samples 40, 41, and 42, those derived from the three patients suspected to have been recently infected by HCV, and the other with the remaining samples, including the two controls of genotype 1b. The first group is characterized by very low genetic variability, with at most two haplotypes and a single mutation among the 10 clone sequences derived from each sample. This is also reflected in other genetic variability parameters that allow a better comparison with other samples, for which the number of clone sequences is occasionally different. The second group is characterized by much higher values of genetic variation, both in number of haplotypes (*h*), total number of mutations (η), substitutions per site (π), and differences among pairs of sequences (κ). These results are indicative of a recent origin of the viral population infecting the patients in the first group.

We further compared the levels of differentiation within and between two groups involved in the transmission, as detected in the previous phylogenetic analyses. For comparison, we also analyzed the differentiation between and within control samples and the two transmission groups. Table 3 summarizes these analyses. In the three comparisons between pairs of

TABLE 2. Summary of inpatient genetic variability for the E1-E2 region of HCV genotype 1b samples analyzed in this study

Group and sample	n^a	h^b	Hd ^c	S^d	η^e	π^f	κ^g
A							
40	10	2	0.200 (0.154) ^h	1	1	0.00042 (0.00033)	0.200 (0.268)
41	10	2	0.467 (0.132)	1	1	0.00099 (0.00028)	0.467 (0.443)
43	20	18	0.989 (0.019)	94	99	0.06116 (0.00828)	28.868 (13.183)
B							
39	10	10	1.000	42	43	0.02961 (0.00444)	13.978 (6.861)
42	10	1	0.000	0	0	0	0
Controls							
36	15	12	0.971 (0.033)	40	44	0.02236 (0.00279)	10.552 (5.096)
37	10	6	0.889 (0.075)	20	21	0.01618 (0.00281)	7.933 (4.032)
38	11	6	0.800 (0.114)	4	4	0.00362 (0.00289)	1.709 (1.078)
90	15	10	0.924 (0.053)	62	63	0.02999 (0.00543)	15.952 (7.541)
92	11	9	0.964 (0.051)	8	8	0.00444 (0.00084)	2.364 (1.393)

^a n , number of clone sequences.

^b h , number of different haplotypes.

^c Hd, haplotype diversity.

^d S , number of polymorphic sites.

^e η , total number of mutations.

^f π , nucleotide diversity.

^g κ , average number of nucleotide differences between pairs of sequences.

^h Standard deviations of relevant parameters are reported in parentheses.

suspect source and recipient patients (40-43, 41-43, and 39-42), the different parameters evaluating genetic differentiation indicate a much lower level of difference between the patients in these pairs than in the comparison between unrelated samples (between groups A and B, between controls and both groups, and between pairs of control samples). This is also reflected in the almost complete absence of polymorphic sites in the sequences derived from recently infected patients that are monomorphic in the corresponding sources and in the number of fixed differences, as opposed to what happens in the compar-

ison between sequences derived from epidemiologically unrelated patients.

DISCUSSION

A detailed analysis of the phylogenetic relationships and genetic variability and differentiation within and among sequences in a rapidly evolving region of the HCV genome provided a different, more accurate image of the relationships among infected patients involved in a hepatitis C virus trans-

TABLE 3. Divergence between clone sequences of relevant pairs of samples

Sample 1 group and sample no.	Sample 2 group or sample no.	S^a	κ^b	Dxy ^c	Da ^d	Fd ^e	Mono1 ^f	Mono2 ^g
Within group A								
40	41	27	25.80	0.05676 (0.00763) ^h	0.05605 (0.00763)	25	1	1
40	43	101	45.05	0.10262 (0.0120)	0.06990 (0.01379)	8	99	1
41	43	96	41.92	0.09500 (0.01118)	0.06200 (0.01293)	2	98	0
Within group B								
39	42	45	18.10	0.03946 (0.0067)	0.02427 (0.0072)	3	43	0
Group A	Group B	146	68.06	0.16023 (0.00813)	0.10932 (0.00833)	25	32	95
Controls	Group A	208	78.94	0.18956 (0.00598)	0.07707 (0.00617)	4	44	161
Controls	Group B	189	73.74	0.17563 (0.00772)	0.08608 (0.00788)	10	14	194
Avg between pairs of controls		103.5	76.10	0.18174 (0.02029)	0.16532 (0.02044)	59.1	16.2	36

^a S , total number of polymorphic sites.

^b κ , average number of nucleotide differences between samples.

^c Dxy, average number of nucleotide substitutions per site between samples with Jukes-Cantor correction.

^d Da, number of net nucleotide substitutions per site between samples with Jukes-Cantor correction.

^e Fd, number of fixed differences between populations.

^f Mono1, mutations polymorphic in sample 2 but monomorphic in sample 1.

^g Mono2, mutations polymorphic in sample 1 but monomorphic in sample 2.

^h Standard deviations of relevant parameters are reported in parentheses.

mission that occurred at a hemodialysis unit. The analysis revealed the occurrence of two independent transmission events, one involving one source (represented by sample 39) and one infected patient (patient 42), the other involving one common source (patient 43) and two recipients (patients 40 and 41). Both transmission events resulted in well-defined, highly supported monophyletic groups when several cloned sequences of the E1-E2 region of the HCV genome from the corresponding samples were analyzed using phylogenetic methods. The support for each of these groups as a whole was similar to that received by the equivalent monophyletic groups in which all the clone sequences from the control and unrelated samples appeared in the same analysis (Fig. 2). Similar levels of support from bootstrap resampling were observed only in clades formed by sequences from the same patient and, in consequence, can be considered a defining mark of very closely related sequences, derived either from the same individual at close time intervals or from very closely related samples, such as those represented by source and recipient individuals. On the other hand, sequences from unrelated samples either do not receive significant bootstrap support (lower than 70%) or the support is barely higher than this value and always lower than 90%.

The analysis of a more conserved, slowly evolving region of the HCV genome such as NS5B does not provide a well enough resolved picture of the previously described relationships. First, high statistical support by bootstrap analysis is obtained only for genotype and subtype clades (Fig. 1), for serially sampled sequences from the same patient, and for one of the two groups related to a recent transmission event, the one represented by samples 39 and 42. However, the other group, composed of sequences derived from samples 40, 41, and 43, is present in the phylogenetic tree but does not receive enough bootstrap support. The analysis of this region also presents an example of a potentially more disturbing result. There appears to be a group (patients 36, 39, and 42) which, although without statistical support, could be mistakenly thought to represent a different, larger transmission case. Nevertheless, the analysis of this region does provide evidence for a lack of association between sequences, even from the same genotype, not related by transmission events. In turn, this allows one to focus on a more labor-intensive analysis involving the cloning and later sequencing of a more rapidly evolving region, such as E1-E2 in the HCV genome.

Another relevant consideration emerging from this study is the need to incorporate as many possible control samples from the local population(s) as are available in the study of outbreaks or transmission cases. The usual evidence for a close relatedness between viruses derived from different samples is their grouping into a more or less well-supported monophyletic group. However, this is not necessarily the result of direct transmission, as it can also be due to both samples being derived from a common population with an older, more divergent common ancestor. If this is the case, then other samples from the same source population and epidemiologically unrelated to the case in question will reveal the true nature of this relationship, but if no such additional sample is included in the study, which will certainly happen if sampling of controls is not sufficient, the relationship will not be noticed. In this study, the availability of a large number of local samples of the same

genotype as those involved in the transmission case helped us to ascertain the existence of at least two different transmission chains with different origins. Nevertheless, the two sources are related, as they derive from a common local pool. Without incorporating a sufficient number of local control samples, both sources and the infected samples could be erroneously interpreted as deriving from a single source in the hemodialysis unit. As a consequence, we can conclude that whatever procedure or protocol was not handled properly and resulted in transmission, it was not a unique event and was repeated at least twice. This implies that a more thorough revision of the operating procedures in this hemodialysis unit is necessary.

ACKNOWLEDGMENTS

This work was supported by Conselleria de Sanitat i Consum, Generalitat Valenciana, and project Grupos03/204 from Agència Valenciana de Ciència i Tecnologia.

REFERENCES

1. **Andrews, P.** 1992. Evolution and environment in the hominoidea. *Nature* **360**:641–646.
2. **Casino, C., J. McAllister, F. Davidson, J. Power, E. Lawlor, P. L. Yap, P. Simmonds, and D. B. Smith.** 1999. Variation of hepatitis C virus following serial transmission: multiple mechanisms of diversification of the hypervariable region and evidence for convergent genome evolution. *J. Gen. Virol.* **80**:717–725.
3. **Echevarría, J. M., P. León, C. J. Domingo, J. A. López, C. Elola, M. Madurga, F. Salmerón, P. L. Yap, J. Daub, and P. Simmonds.** 1996. Laboratory diagnosis and molecular epidemiology of an outbreak of hepatitis C virus infection among recipients of human intravenous immunoglobulin in Spain. *Transfusion* **36**:725–730.
4. **Esteban, J. I., J. Gómez, M. Martell, B. Cabot, J. Quer, J. Camps, A. González, T. Otero, A. Moya, R. Esteban, and J. Guardia.** 1996. Transmission of hepatitis C virus by a cardiac surgeon. *N. Engl. J. Med.* **334**:555–560.
5. **Felsenstein, J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**:783–791.
6. **González-Candelas, F., M. A. Bracho, and A. Moya.** 2003. Molecular epidemiology and forensic genetics: application to a hepatitis C virus transmission event at a hemodialysis unit. *J. Infect. Dis.* **187**:352–358.
7. **Guindon, S., and O. Gascuel.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**:696–704.
8. **Jiménez, N.** 2004. Evolución del virus de la hepatitis C en muestras hospitalarias de la Comunidad Valenciana. Ph.D. thesis. Universitat de Valencia, Valencia, Spain.
9. **Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei.** 2001. MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**:1244–1245.
10. **Memon, M. I., and M. A. Memon.** 2002. Hepatitis C: an epidemiological review. *J. Viral Hepat.* **9**:84–100.
11. **Posada, D., and K. A. Crandall.** 2001. Selecting the best-fit model of nucleotide substitution. *Syst. Biol.* **50**:580–601.
12. **Power, J. P., E. Lawlor, F. Davidson, E. C. Holmes, P. L. Yap, and P. Simmonds.** 1995. Molecular epidemiology of an outbreak of infection with hepatitis C virus in recipients of anti-D immunoglobulin. *Lancet* **345**:1211–1213.
13. **Pradan, P., and C. Trépo.** 2000. HCV: epidemiology, modes of transmission and prevention of spread. *Baillieres Best Pract. Res. Clin. Gastroenterol.* **14**:201–210.
14. **Rodríguez, F., J. L. Oliver, A. Marín, and J. R. Medina.** 1990. The general stochastic model of nucleotide substitution. *J. Theor. Biol.* **142**:485–501.
15. **Ross, R. S., S. Viazov, and M. Roggendorf.** 2002. Phylogenetic analysis indicates transmission of hepatitis C virus from an infected orthopedic surgeon to a patient. *J. Med. Virol.* **66**:461–467.
16. **Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas.** 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**:2496–2497.
17. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
18. **Swofford, D. L.** 2002. PAUP*. Phylogenetic analysis using parsimony (* and other methods). Sinauer Associates, Sunderland, Mass.
19. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
20. **Torres-Puente, M.** 2004. Variabilidad genética y respuesta al tratamiento antiviral en el virus de la hepatitis C (VHC). Ph.D. thesis. Universitat de Valencia, Valencia, Spain.