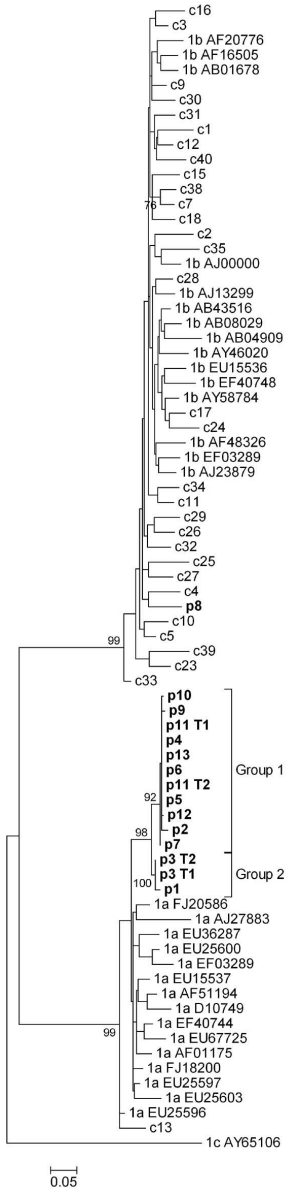
# Phylogenetic Investigation Of Nosocomial Transmission Of Hepatitis C Virus In An Oncology Ward

Ágnes Dencs, Andrea Hettmann, Tibor Martyin, Csilla Jekkel, Tivadar Bányai, Mária Takács

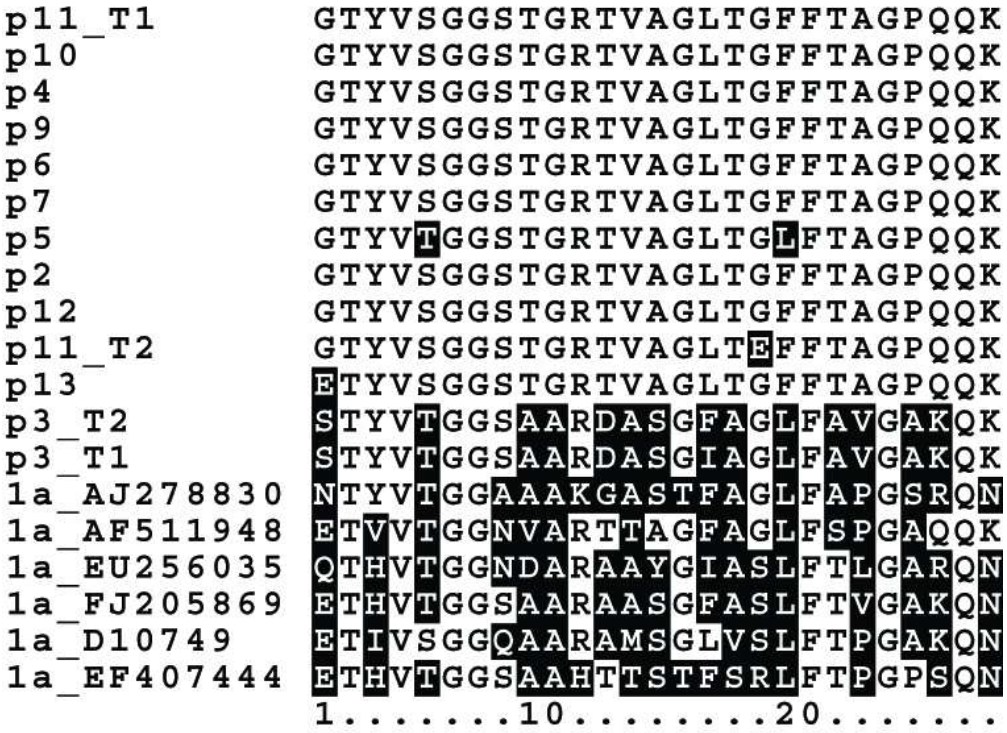
Journal of Medical Virology

# Phylogenetic Investigation Of Nosocomial Transmission Of Hepatitis C Virus In An Oncology Ward

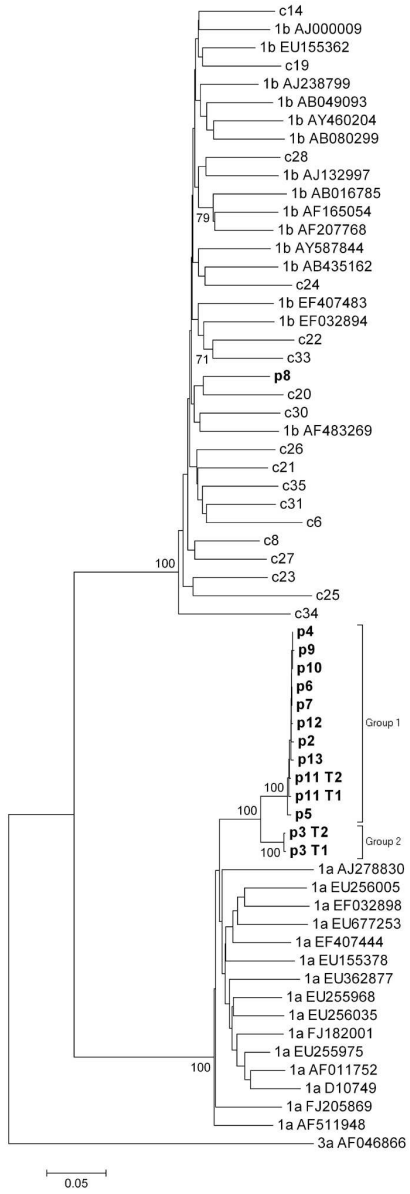
| | |
|---|---|
| Journal: | *Journal of Medical Virology* |
| Manuscript ID: | JMV-10-1843.R2 |
| Wiley - Manuscript type: | Research Article |
| Date Submitted by the Author: | 17-Sep-2010 |
| Complete List of Authors: | Dencs, Ágnes; National Center for Epidemiology, Division for Virology<br>Hettmann, Andrea; National Center for Epidemiology, Division for Virology<br>Martyin, Tibor; Pándy Kálmán Hospital,, Department of Infectology<br>Jekkel, Csilla; National Center for Epidemiology, Division for Virology<br>Bányai, Tivadar; Pándy Kálmán Hospital,, Department of Infectology<br>Takács, Mária; National Center for Epidemiology, Division for Virology |
| Keywords: | hepatitis C, NS5B region, envelope region, nosocomial infection, phylogenetic analysis |
| | |




37x149mm (400 x 400 DPI)




96x71mm (200 x 200 DPI)




44x129mm (400 x 400 DPI)

# Phylogenetic Investigation of Nosocomial Transmission of Hepatitis C Virus in an Oncology Ward

**Ágnes Dencs[1], Andrea Hettmann[1], Tibor Martyin[2], Csilla Jekkel[1], Tivadar Bányai[2], Mária Takács[1]**

[1]*National Center for Epidemiology, Division of Virology, H-1097, Budapest, Gyáli út 2-6.*

[2]*Pándy Kálmán Hospital, Department of Infectology, H-5700, Gyula, Semmelweis u. 1.*

Corresponding author: Mária Takács

Address: National Center for Epidemiology, Division of Virology, H-1097, Budapest, Gyáli út 2-6.

Telephone: +36-1-476-1392

Fax: +36-1-476-1368

E-mail: takmar@gmail.com

Running title: Transmission of HCV in an oncology ward

## Abstract

Nosocomial hepatitis C virus (HCV) infections have been reported from different health-care settings worldwide. Twenty patients, treated at the same oncology department, with no previous record of hepatitis C infection, tested positive for anti-HCV antibodies between November 2007 and June 2008. Twelve of the newly infected patients were found to be HCV RNA positive. The common origin of the infections was assumed. To investigate the relatedness of the detected viral strains phylogenetic analyses were performed using sequences from the NS5B and E1/E2 genome regions. A patient carrying HCV for years, was also involved in the study. She was treated at the same oncology department and was considered a possible infectious source. The previous HCV carrier harboured subtype 1b, while all other patients were infected with subtype 1a. Sequences from the 12 newly infected patients formed two groups. The viral sequences within the groups were very closely related. A greater evolutionary distance was observed between the two groups; however, their relatedness could be demonstrated by sequences from both regions with high statistical support. The results indicated that nosocomial transmission occurred. The phylogenetic analyses suggested that the viruses originated from a common source, possibly a patient carrying highly divergent variants. This presumed infectious source could not be identified in the course of this study. The genotype distribution of Hungarian control sequences included in the analysis confirmed this conclusion, since HCV genotype 1a was found to be relatively uncommon.


Keywords: hepatitis C, NS5B, envelope region, nosocomial infection, phylogenetic analysis

## Introduction

Hepatitis C virus (HCV) is a member of the *Hepacivirus* genus within the *Flaviviridae* family. HCV infection is a major public health issue worldwide. An estimated 170 million people are infected chronically (WHO, 1997). Chronic infection with HCV can lead to chronic liver disease, cirrhosis and hepatocellular carcinoma. In countries where the screening of donors has been introduced, acute hepatitis C virus infections related to the transfusion of blood or blood products and transplantation of solid organs have become rare. In developed countries, injection drug use accounts for most new HCV infections (Alter, 1999, Esteban et al., 2008). Other potential risk factors include birth to an infected mother, occupational exposure, sex with an infected partner and commercial tattooing (Indolfi and Resti, 2009, Henderson, 2003, Wejstal, 1999, Haley and Fischer, 2003). Health-care associated HCV infections also occur. Outbreaks of hepatitis C have been reported from haemodialysis units, oncology and hematology wards and different inpatient and ambulatory care settings. Most of these involved patient to patient transmission, occasionally affecting a large number of patients (CDC, 2003, Forns et al., 2005, Pekova et al., 2007, Sepkowitz, 2004). Usually, infections are caused by breaches in basic infection control practices by medical staff, such as the reuse of disposable needles and syringes, which can lead to the contamination of infusion equipment or multiple-dose medication vials (Williams et al., 2003). Sepkowitz reviews epidemics in oncology wards and he raises the possibility that HCV can be found in saliva and tears, and might have been transmitted via exposure to these secretions. The possibility that an infected health care worker might serve as the source of infection was examined in most studies and was excluded in all cases. Blood transfusion was also excluded as the source of the epidemics (Sepkowitz, 2004).

Phylogenetic analysis of viral sequences obtained following such outbreaks can provide molecular evidence for the relatedness of infections, identify or exclude the possible sources and may help elucidate the events that occurred. Many molecular epidemiological studies on hepatitis C outbreaks have utilized either the fast changing first hypervariable region (HVR1), or partial sequences of the envelope region containing the HVR1 (Izopet et al., 1999, Silini et al., 2002). The more slowly evolving NS5B region has been shown to be best suited for genotyping and subtyping HCV (Laperche et al., 2005), but can also provide enough information to confirm relatedness of viral strains (Carneiro et al., 2007). However, better resolution may be achieved by the analysis of the E2 region - including HVR1 – in some cases, where sequences from the recipient are obtained shortly after infection (Bracho et al., 2005). Some compare or combine the approaches above, or perform quasispecies analysis of envelope region sequences, a method which has significantly improved the understanding of the transmission of virus variants between donor and recipient (Esteban et al., 1996, Bracho et al., 2005, Hmaïed et al., 2007). Utilization of longer sequences of both regions has been suggested for improving statistical support of results (Casino et al., 1999). In this paper a case is presented when phylogenetic analysis of sequences outside the first hypervariable region reveal a more distant relationship between transmitted viral sequences.

In 2007 and 2008 several patients of the oncology department of a hospital in Hungary returned for treatment or control with elevated ALT levels. Patients treated at the department were screened for anti-HCV antibodies, and twenty patients with no previous history of HCV infection were found to be positive. The possibility of nosocomial transmission was investigated by phylogenetic analyses of partial sequences from both the NS5B and E1/E2 regions. The study was approved by an institutional review committee and performed in accordance with Hungarian law. All patients gave their informed consent in writing.

## Materials and methods

### Patients and samples

Serum samples of 21 anti-HCV positive patients of the oncology ward were received for the detection of viral RNA, genotyping and phylogenetic analysis. Twenty patients were tested positive for anti-HCV antibodies between November 2007 and June 2008 with no history of previous HCV infection, and one patient had been known to be anti-HCV positive at least since 1998. Serum samples were collected between March and November of 2008, when the patients returned to the hospital for treatment or control. All patients were female, their average age was 57.7 years (range 27-76). Samples from 40 Hungarian HCV carriers from different parts of the country were used as controls (19 male and 21 female, average age 50.2, range: 20-79), including 6 local controls within 150 kilometers from the hospital. All samples were stored frozen at -20°C until the analysis.

### Detection of HCV RNA

Nucleic acids were extracted from 200 µl serum using QIAamp MinElute Virus Spin Kit (Qiagen, Hilden, Germany). Viral RNA was reverse transcribed using ImProm-II Reverse Transcription System (Promega Corp., Madison, WI) with random hexamers and in the presence of 20 units of Rnase inhibitor. 10 µl of the cDNA was used for amplifications in each case.

Table 1 shows the regions and primers used for amplification and sequencing.

Nested primers published by Khan et al. (2004) were used with modifications to detect the conserved 5'UTR regions of all known genotypes of hepatitis C virus. Primers specific for the NS5B region of the HCV genome were designed to detect genotype 1 viruses. PCR specific for the envelope region was performed with primers described by Ray et al. (2000). The

primers amplify a 1026 nucleotide long product that contains the carboxy-terminus of the core gene, the complete E1 region and a segment of the E2 region including HVR1.

All bands were detected in 2% agarose gel with ethidium bromide staining under UV illumination.

Sequencing of PCR products

PCR products were purified with Viogene PCR Advanced kit (Viogene, Taipei, Taiwan) using the spin column method. Purified PCR products were sequenced using the DYEnamic ET Dye Terminator Cycle Sequencing Kit (Amersham Biosciences, Little Chalfont Buckinghamshire, UK). Purification of sequencing reactions were carried out according to the ethanol precipitation protocol supplied with the kit. Reaction products were separated and analyzed on a MegaBACE 1000 capillary sequencer (Amersham Biosciences, Little Chalfont Buckinghamshire, UK).

Phylogenetic analysis

The following hepatitis C virus genotype 1a and 1b sequences were included in the phylogenetic analysis: 1a isolates EU677253, AJ278830, EF032898, EU155378, FJ205869, EU255968, FJ182001, EU256005, EU256035, EU362877, AF511948, AF511949, D10749, EU256005, EF407444, and 1b isolates AY587844, AB435162, EF407483, AF165054, AB049093, EF032894, EU155362, AJ132997, AF483269, AJ000009, AY460204, AF207768, AB080299, AB016785, AJ238799. Sequences were aligned with the MultAlin program (Corpet, 1988). The nucleotide substitution model that described the data best was chosen by ModelTest (Posada and Crandall, 1998) using PAUP* (Swofford, 2002) as implemented in FindModel. Phylogenetic trees were constructed by the Maximum Likelihood method using PhyML 3.0 (Guindon and Gascuel, 2003). The starting tree was computed by

PhyML using BIONJ (Gascuel, 1997). Subtype 1c isolate AY651061 was used as an outgroup. Bootstrap analysis with 1000 replicates was performed to confirm tree topology. Evolutionary distance calculations (nucleotide substitution model: Kimura 2-parameter) were carried out using MEGA software version 4.0 (Tamura et al., 2007).

Patient and control sequences obtained in this study were deposited in GenBank under the following accession numbers: patient sequences from the NS5B region (FN666627-FN666641), E1/E2 region (FN667677-FN667690), control sequences from the NS5B region (FN666642- FN666672), E1/E2 region (FN667691-FN667708), UTR region (FN668896, FN668897).

## Results

The serum samples of 21 anti-HCV positive patients were tested for the presence of HCV RNA using nested PCR with primers specific for the conserved 5'UTR region. Viral RNA could be detected in the sera of 12 of the newly infected oncology patients. The patient with previous HCV infection (patient 8) was also HCV RNA positive, thus was considered a putative source patient in the possible transmission event(s). Antiviral treatment was begun in the case of six HCV RNA positive patients. The other patients could not be treated because of metastasis or other diseases. Details of the 13 HCV RNA positive patients are shown in Table 2.

For the purpose of genotyping and phylogenetic analysis, a segment of the NS5B region from the HCV RNA positive patients was amplified and sequenced. Analysis of a 500 nucleotide long segment of the NS5B region was performed using sequences obtained from the 13 patients of the oncology, 31 unrelated controls from different regions of Hungary and 30 genotype 1a and 1b isolates from GenBank (Fig. 1). A rooted tree was constructed using the General Time Reversible model (gamma shape parameter: 0.397, proportion of invariant sites: 0.265). The tree showed that all 12 *de novo* infected patients were infected with subtype 1a viruses. Patient 8 was infected with subtype 1b, and was excluded as a possible source patient. The genotype 1a sequences from the outbreak formed two groups with high statistical support. Group 1 consisted of the closely related viral sequences from 10 newly infected patients (mean evolutionary distance: 0.003) (Table 3a). Five sequences were completely identical, and only the sequence from patient 2 differed from the others by more than 3 nucleotide positions. The sequences from the two other patients were clustered on a separate branch (Group 2). These sequences were separated from Group 1 with a mean evolutionary distance of 0.021. They differed from each other in three nucleotide positions, all three

changes also appearing on the amino acid level. Groups 1 and 2 were joined together on the tree by a common branch, which appeared in 98% of bootstrap replicates, suggesting a possible relatedness. The 12 sequences appeared as a cluster that was separated from subtype 1a isolates used for constructing the tree by a mean evolutionary distance (0.061) similar to those observed between pairs of independent HCV isolates of the same subtype (mean for subtype 1a: 0.050).

To see whether another, faster evolving region also supports this grouping, PCR was performed to amplify a portion from the more variable E1/E2 region of the virus, which included HVR1. Products could be obtained for only 12 patients. Patient 1 could not be included in this part of the study, because the sample did not give a positive result with the E1/E2 primer set even after several attempts. PCR products from 12 other samples were sequenced.

The alignment of the 81 nucleotide long HVR1 sequences, often used to demonstrate phylogenetic inference, from the oncology patients infected with 1a subtype strongly supported the existence of Group 1 with only 5 sporadic substitutions in the 10 sequences, 4 of them nonsynonymous. No close relationship of the viruses infecting Group 1 patients and patient 3 could be observed. Their HVR1 sequences differed in 12-15 amino acid residues, similarly to unrelated control and GenBank isolates (Fig. 2).

However, when the complete determined sequence (positions: 887-1813 nt) was used to construct a phylogenetic tree, it showed a similar picture to the one based on NS5B (Fig. 3). Again, the General Time Reversible model fit the data best (gamma shape parameter: 0.597, proportion of invariant sites: 0.266). The sequence from patient 8 was clustered together with 1b subtype HCV isolates. Among 1a isolates from the oncology the same grouping could be observed with high statistical support. Table 3b shows the sequence statistics of the envelope region. Sequences in Group 1 were very close genetically, the average distance within the

group was only 0.003. Thirteen variable nucleotide positions were observed, 5 of them within HVR1. These resulted in three variable amino acid residues in the E1 protein (3/192, 1.56%), 4 within HVR1 (4/27, 14.81%) and one in the short flanking segment of E2 (1/80, 1.25%). The mean evolutionary distance between Groups 1 and 2 (patient 3) was 0.049 on the nucleotide level. In 39 nucleotide positions the occurring bases were characteristic for each group, on the amino acid level 18 differences were unique.

The pairwise distances separating unrelated GenBank and Hungarian control sequences of the same subtype were also calculated and - like in case of NS5B sequences - they were found to be typically much greater than between the two patient groups (mean for subtype 1a: 0.115).

The NS5B and E1/E2 regions from patients 3 and 11 were sequenced twice, 3 months apart (sequences marked T1 and T2). The NS5B regions were found to be unchanged, in the envelope protein region, however, 2 and 3 substitutions could be observed on the nucleic acid level for patients 3 and 11, respectively, causing one change in the amino acid sequence in each patient.

The nucleotide sequences from 40 Hungarian HCV carriers determined in this study served as unrelated controls, but also provided information on the genotype and subtype distribution of Hungarian hepatitis C isolates. 38 of 40 viruses (95%) belonged to genotype 1 based on 5'UTR sequences. Subtypes within genotype 1 were determined based on NS5B and/or the envelope region. Where both regions were sequenced (11 controls), the genotyping results were in complete concordance. A single subtype 1a virus was found among the 38 genotype 1 isolates (1/40, 2.5%), all others were subtype 1b. In the case of two control patients only the 5'UTR region could be amplified and these viruses were genotyped based on the nucleotide sequences of that PCR product. One of them was shown to belong to genotype 3, the other to genotype 4. Six of the controls were local controls originating from the South-Eastern part of

the country, within a 150 km radius of the hospital (c24, c26, c27, c38-40). They all carried subtype 1b viruses.

## Discussion

After several patients at the oncology department of a hospital in Hungary with elevated ALT levels and no previous record of HCV infection were found to be anti-HCV positive, molecular epidemiological investigation of a possible nosocomial outbreak was performed. The oncology ward had a high patient turnover with several thousand patients per year. They were not routinely screened for hepatitis B or C markers, but blood tests including liver function tests were performed regularly to control side effects of cytostatic treatment. Cases of high liver function were further investigated. A notable increase in the incidence of elevated ALT levels among patients returning for control was noticed in January and February of 2008. Until October of 2008, a total of 45 patients were screened for anti-HCV antibodies, and 21 of them were found to be positive. Twenty of them had no history of hepatitis C infection. One patient had been a known virus carrier for years (patient 8), and was considered a possible common source of infections.

Serum samples from 21 anti-HCV positive patients were sent for hepatitis C RNA detection, genotyping and phylogenetic analysis. Thirteen samples were HCV RNA positive, including patient 8, the long time carrier. Genotyping showed that *de novo* infected patients were infected with subtype 1a. However, patient 8 carried a subtype 1b virus, and thus was not the infectious source. Phylogenetic analysis of the NS5B region showed that the viruses of 10 newly infected patients were closely related (Group 1) and very likely had a common origin. The viruses infecting the two other patients were similarly closely related to each other (Group 2), but were clearly separated from the other 10 isolates by a distance less than usually observed between unrelated strains. The twelve sequences formed a statistically supported group that appeared monophyletic, which suggested they may be associated. Since NS5B evolves relatively slowly, a large segment of envelope region including HVR1 was also

amplified and analyzed. Samples of only 12 of the patients gave positive PCR results with this primer set. This occurred either because of a mutation in the primer binding regions, or the lower sensitivity of this method. The results obtained were similar and provided further evidence of a more distant, but common ancestor. GenBank controls also demonstrated that even unrelated isolates may show relatively low genetic distances, especially in the NS5B region. However, pairs of reference sequences with the smallest distances did not appear as a monophyletic group (e.g. isolates FJ182001 and EU255975). The rapidly changing HVR1 supported the existence of Group 1, but the only sequences that could be obtained from Group 2 (from patient 3) showed no similarity to Group 1 and differed as much from Group 1 sequences as unrelated isolates. Two possible interpretations of these results may be considered. One of these is that the common source of infections was a long time HCV carrier, who carried highly divergent quasispecies. A small inoculum was probably transferred from the source patient to the recipients, which may have only contained a fraction of these variants. Different variants became dominant in the different recipients, and diverged further from their ancestral sequence in the months between the infection and the collection of samples. This may be an example of selective transmission of HCV quasispecies, which has been first described in chimpanzee and *in vitro* experiments, where a total of three chimpanzees were infected with the same HCV containing plasma. Of the quasispecies present in the inoculum, the major variant became dominant in two animals and a minor one in the third animal (Hijikata et al., 1995, Sugitani et al., 1998). Selective transmission has been also documented in humans. After a needle stick accident the minor clone in the donor was found to be the most closely related to the major clone detected in the recipient (Saito et al., 2004). In an outbreak at a haemodialysis unit sequences from two recipients showed greatest similarity to a small subset of variants found in the source patient (Bracho et al., 2005). A HCV carrier cardiac surgeon infected several of his patients during surgery, and the

sequences recovered from the different recipients were related to different variants in the surgeon (Esteban et al., 1996). The transmission of a single or a small group of variants was observed also in perinatal infection (Manzin et al., 2000).

The common source of the infections at the oncology ward may have been another HCV carrier patient (or a health care worker) not identified by the field investigation, thus not included in the analysis. This patient was probably missed by the screening for anti-HCV antibodies either because of normal liver functions, or because of remission from or death caused by the patient's primary disease. Several other patients infected recently at the hospital with HCV could have also been missed by the investigation for the same reasons. Unfortunately, there are not sufficient epidemiological data available to determine the order the infections occurred and trace the patient-to-patient route of transmissions.

Another possible explanation of the results is that the sequences in Groups 1 and 2 are not related directly. In this case the reason for the homology detected between them is that they originate from a common population, and have an older common ancestor. To investigate this possibility, the viruses of 40 HCV carrier control patients from Hungary unrelated to the outbreak were also genotyped, including 6 local controls from within a 150 km radius of the hospital. Genotyping of the controls showed that subtype 1b is predominant in Hungary in the clinical material studied (92.5%), which is a similar result to that observed by Gervain (Gervain et al., 2003). Only a single subtype 1a virus was found, originating from a different part of the country and showing no notable sequence similarity to viruses of the oncology patients. These results do not support that two cases of nosocomial transmission may have occurred coincidentally from two unrelated HCV subtype 1a carriers both treated at the hospital.

In summary the results suggest that transmission of hepatitis C virus occurred in the oncology ward, where invasive procedures were frequent in the therapy of the patients. The infections

probably originated from a common source, although the source patient could not be identified. The precise mode of transmission could not be ascertained either, but contamination of a multiple-dose medication vial or infusion equipment could be suspected. After the outbreak was discovered, strict infection control measures were introduced and regularly monitored. After October of 2008 no more new cases were identified. These findings are similar to those published by Silini et al. (2002), Dumpis et. al. (2003), and Maillaird et al. (2009)

Silini et al. (2002) observed 13 newly admitted patients with anti-HCV seroconversion in a hematology ward. They excluded the role of HCV infected health care workers, and transfusional origin. Patient-to-patient transmission and the role of multidose vials were assumed as a cause of infection. Dumpis et al. (2003) tested 46 HCV infected patients on a paediatric oncology ward. They also excluded the role of HCV carrier health care workers and the role of the blood transfusion. It was concluded that infections could occur outside the patients' room and may relate to some malpractice by the medical staff. Mailliard et al. (2009) tested a large genotype 3a HCV outbreak at an oncology clinic and found that it was a single-source, patient-to-patient outbreak.

Molecular methods have been used widely in investigation of hepatitis C virus outbreaks. In the present study sequence analysis of the NS5B region gave the correct grouping of the isolates under investigation with good statistical support. The observed clusters were confirmed by the phylogenetic tree based on a large portion of the envelope region. It seems that in many cases determining partial sequences (about 500 bases or more) of NS5B is sufficient to show or exclude relatedness of virus strains as also applied in an earlier investigation (Dencs et al., 2009). This method has the added advantage that false negative PCR results due to mutations in the primer binding regions are less likely to occur than in case of the more variable E1/E2. The short HVR1 often used in phylogenetic analyses may

also be suitable when only a short time has passed since the infection. However, in this study the more distant relatedness of the virus groups did not emerge from the analysis of this region.

Acknowledgements

The authors thank Zoltán Kugler for his help in data collection and epidemiological investigations. The skilled technical assistance of Mrs Katalin Kulcsár, Mrs Mária Végh and Mrs Erzsébet Dubszki is very much appreciated.

References

Alter, M. J. 1999. HCV infection in the United States. J. Hepatol.31(Suppl. 1):88–91.

Bracho MA, Gosalbes MJ, Blasco D, Moya A, González-Candelas F. 2005. Molecular epidemiology of a hepatitis C virus outbreak in a hemodialysis unit. J Clin Microbiol. 43:2750-2755.

Carneiro MA, Teles SA, Lampe E, Espírito-Santo MP, Gouveia-Oliveira R, Reis NR, Yoshida CF, Martins RM. 2007. Molecular and epidemiological study on nosocomial transmission of HCV in hemodialysis patients in Brazil. J Med Virol. 79:1325-1333.

Casino C, McAllister J, Davidson F, Power J, Lawlor E, Yap PL, Simmonds P, Smith DB 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.

Centers for Disease Control and Prevention. Transmission of hepatitis B and C viruses in outpatient settings—New York, Oklahoma, and Nebraska, 2000–2002. Morbidity and Mortality Weakly Report 2003. 52:901–906.

Corpet F. 1988. Multiple sequence alignment with hierarchical clustering. Nucl Acids Res. 16:10881-10890.

Dencs Á, Hettmann A, Szűcs M, Rusvai E, Takács M. 2009. Phylogenetic analysis of a hepatitis C virus outbreak in a haemodialysis unit. 51st Annual meeting of the Hungarian Society of Gastroenterology, Tihany, June 13-16, 2009. Z. Gastroenter. 47: 463.

Dumpis U, Kovalova Z, Jansons J, Cupane L, Sominskaya I, Michailova M, Karayiannis P, Gardovska D, Viazov S, Ross S, Roggendorf M, Pumpens P. 2003. An outbreak of HBV and HCV infection in a paediatric oncology ward: epidemiological investigations and prevention of further spread. J.Med Virol. 69:331-338.

Esteban JI, Gómez J, Martell M, Cabot B, Quer J, Camps J, González A, Otero T, Moya A, Esteban R, Guardia, J. 1996. Transmission of hepatitis C virus by a cardiac surgeon. N Engl J Med. 334:555-560.

Esteban JI, Sauleda S, Quer J. 2008. The changing epidemiology of hepatitis C virus infection in Europe Journal of Hepatology. 48:148–162.

Forns X, Martínez-Bauer E, Feliu A, García-Retortillo M, Martín M, Gay E, Navasa M, Sánchez-Tapias JM, Bruguera M, Rodés J. 2005. Nosocomial transmission of HCV in the liver unit of a tertiary care center. Hepatology. 41:115-22.

Gascuel O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol. 14:685-695.

Gervain J, Simon Jr G, Simon J. 2003. Genotype distribution of hepatitis C virus in the Hungarian population with chronic viral hepatitis C. Eur J Gastroenterol Hepatol. 4:449-450.

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

Haley RW, Fischer RP. 2003. The tattooing paradox: are studies of acute hepatitis adequate to identify routes of transmission of subclinical hepatitis C infection? Arch Intern Med. 163:1095-8.

Henderson, DK. 2003. Managing Occupational Risks for Hepatitis C Transmission in the Health Care Setting. Clin Microbiol Rev. 16:546–568.

Hijikata M, Mizuno K, Rikihisa T, Shimizu YK, Iwamoto A, Nakajima N, Yoshikura H. 1995. Selective transmission of hepatitis C virus in vivo and in vitro. Arch Virol. 140:1623-8.

Hmaïed F, Ben Mamou M, Dubois M, Pasquier C, Sandres-Saune K, Rostaing L, Slim A, Arrouji Z, Ben Redjeb S, Izopet J. 2007. Determining the source of nosocomial transmission in hemodialysis units in Tunisia by sequencing NS5B and E2 sequences of HCV. J Med Virol. 79:1089-1094.

Indolfi G, Resti M. Perinatal transmission of hepatitis C virus infection. 2009. J Med Virol. 81:836-843.

Izopet J, Pasquier C, Sandres K, Puel J, Rostaing L. 1999. Molecular evidence for nosocomial transmission of hepatitis C virus in a French hemodialysis unit. J Med Virol. 58:139-144.

Khan N, Aswad S, Shidban H, Aghajani M, Mendez R, Mendez R, Comanor L. 2004. Improved detection of HCV Infection in hemodialysis patients using a new HCV RNA qualitative assay: experience of a transplant center. J Clin Virol. 30:175-182.

Laperche S, Lunel F, Izopet J, Alain S, Dény P, Duverlie G, Gaudy C, Pawlotsky JM, Plantier JC, Pozzetto B, Thibault V, Tosetti F, Lefrère JJ. 2005. Comparison of hepatitis C virus NS5b and 5' noncoding gene sequencing methods in a multicenter study. J Clin Microbiol. 43:733-739.

Mailliard ME, Capadano ME, Hrnicek MJ, Gilroy RK, Gulizia JM. 2009. Outcomes of a patient-to-patient outbreak of genotype 3a hepatitis C. Hepatology. 50:361-368.

Manzin A, Solforosi L, Debiaggi M, Zara F, Tanzi E, Romanò L, Zanetti AR, Clementi M. 2000. Dominant role of host selective pressure in driving hepatitis C virus evolution in perinatal infection. Virol. 74:4327-4334.

Pekova LM, Teocharov P, Sakarev A. 2007. Clinical course and outcome of a nosocomial outbreak of hepatitis C in a urology ward. J Hosp Infect. 67:86-91.

Posada D., Crandall KA. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics. 14:817-818.

Ray SC, Mao Q, Lanford RE, Bassett S, Laeyendecker O, Wang YM, Thomas DL. 2000 Hypervariable region 1 sequence stability during hepatitis C virus replication in chimpanzees. J Virol. 74:3058-66.

Saito T, Watanabe H, Shao L, Okumoto K, Hattori E, Sanjo M, Misawa K, Suzuki A, Takeda T, Sugahara K, Ito JI, Saito K, Togashi H, Kawata S. 2004. Transmission of hepatitis C virus quasispecies between human adults. Hepatol Res. 30:57-62.

Sepkowitz KA. 2004. Risk to cancer patients from nosocomial hepatitis C virus. Infect Control Hosp Epidemiol. 25:599-602.

Silini E, Locasciulli A, Santoleri L, Gargantini L, Pinzello G, Montillo M, Foti L, Lisa A, Orfeo N, Magliano E, Nosari A, Morra E. 2002. Hepatitis C virus infection in a hematology ward: evidence for nosocomial transmission and impact on hematologic disease outcome. Haematologica. 87:1200-1208.

Sugitani M, Shikata T. 1998. Comparison of amino acid sequences in hypervariable region-1 of hepatitis C virus clones between human inocula and the infected chimpanzee sera. Virus Res. 56:177-82.

Swofford, D. L. 2002. PAUP*. Phylogenetic analysis using parsimony (* and other methods). Sinauer Associates, Sunderland, Mass.

Tamura K, Dudley J, Nei M & Kumar S 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution. 24:1596-1599.

Wejstal R. 1999. Sexual transmission of hepatitis C virus. J Hepatol. 31 Suppl 1:92-95.

Williams IT, Perz JF, Bell BP. 2004. Viral hepatitis transmission in ambulatory health care settings. Clin Infect Dis. 38:1592-1598.

World Health Organization. 1997. Hepatitis C: global prevalence. Wkly Epidemiol Rec. 72:341-344.

Table 1. Regions of HCV and primers used for PCR and sequencing

| Region | Step | Annealing temperature °C | No. of cycle | Direction | Position* | Sequence |
|---|---|---|---|---|---|---|
| 5'UTR | Outer | 60 | 35 | Sense | 43-68 | 5'-CCC TGT GAG GAA CTA CTG TCT TCA CG-3' |
| | | | | Antisense | 348-322 | 5'-TGC TCA TGG TGC ACG GTC TAC GAG ACC-3' |
| | Nested | 59 | 30 | Sense | 78-103 | 5'-TCT AGC CAT GGC GTT AGT ATG AGT GT-3' |
| | | | | Antisense | 313-288 | 5'-CAC TCG CAA GCA CCC TAT CAG GCA GT-3' |
| NS5B | Outer | 58 | 40 | Sense | 8312-8332 | 5'-GGA GKC AAT YTA CCA ATG YTG TGA C-3 |
| | | | | Antisense | 9220-9203 | 5'-GCC GSD ATT GGA GTG AGT TTR AG-3' |
| | Nested | 57 | 35 | Sense | 8622-8636 | 5'-ACG GAK GCT ATG ACY AGG-3' |
| | | | | Antisense | 9187-9171 | 5'-ACY GCC CAR TTR AAG AGG TA-3' |
| E1/E2 | Outer | 58 | 35 | Sense | 834-859 | 5'-GCA ACA GGG AAC CT TCCT GGT TGC TC-3' |
| | | | | Antisense | 2086-2064 | 5'-GGG CAG DBC ARR GTG TTG TTG CC-3' |
| | Nested | 59 | 30 | Sense | 843-868 | 5'-AAC CTT CCT GGT TGC TCT TTC TCT AT-3' |
| | | | | Antisense | 1868-1848 | 5'-GAA GCA ATA YAC YGG RCC ACA-3' |

*Nucleotide positions are based on the H77 isolate (accession number: AF009606).

Table 2. Details of the oncology patients included in the molecular epidemiological study

| Patient no | Age | Primary disease | First detection of elevated ALT | First detection of aHCV positivity | HCV genotype | Additional information |
|---|---|---|---|---|---|---|
| P1 | 57 | Breast cancer | October 2007 | November 2007 | 1a | |
| P2 | 53 | Breast cancer | April 2008 | May 2008 | 1a | |
| P3 | 70 | Ovarian cancer | June 2007 | January 2008 | 1a | |
| P4 | 70 | Breast cancer | February 2008 | August 2008 | 1a | |
| P5 | 58 | Breast cancer | December 2007 | June 2008 | 1a | |
| P6 | 67 | Breast cancer | October 2007 | May 2008 | 1a | |
| P7 | 56 | Breast cancer | August 2008 | June 2008 | 1a | |
| P8 | 76 | Hodgkin lymphoma | Before October 1998 | Before October 1998 | 1b | Known to be already a HCV carrier before 1998 |
| P9 | 65 | Ovarian cancer | October 2007 | December 2007 | 1a | Elevated ALT in 2003, but anti-HCV negative. Sudden deterioration of liver functions in 2007 |
| P10 | 55 | Breast cancer | January 2008 | January 2008 | 1a | Developed chronic eosinophilic leukaemia in the 8th week of therapy |
| P11 | 51 | Colonic adenocarcinoma | November 2007 | December 2007 | 1a | |
| P12 | 57 | Urinary bladder cancer | January 2008 | June 2008 | 1a | |
| P13 | 49 | Breast cancer | 2005 | August 2008 | 1a | |

Table 3a. Sequence statistics based on the determined 500 base long segment from the NS5B region

| Compared sequences | Nucleotide level | | | Amino acid level | | |
|---|---|---|---|---|---|---|
| | Number of variable positions | Avg. no. of different positions (SEM) | Mean genetic distance (SEM) | Number of variable positions | Avg. no. of different positions (SEM) | p-distance (SEM) |
| Within group 1 | 9 | 1.636 (0.535) | 0.003 (0.001) | 4 | 0.727 (0.354) | 0.004 (0.002) |
| Within group 2 | 3 | 3.000 (1.717) | 0.006 (0.003) | 3 | 3.000 (1.687) | 0.018 (0.010) |
| Between groups 1 and 2 | 19 | 10.455 (2.913) | 0.021 (0.006) | 7 | 1.864 (0.870) | 0.011 (0.005) |
| Among controls of the same subtype (1a/1b) | 96 / 189 | 23.792 (2.518) / 31.245 (2.489) | 0.050 (0.006) / 0.067 (0.006) | 18 / 53 | 3.392 (0.872) / 5.029 (0.857) | 0.021 (0.005) / 0.031 (0.005) |
| Between subtype 1a controls and groups 1-2 | 112 | 28.670 (3.632) | 0.061 (0.008) | 23 | 5.643 (1.867) | 0.034 (0.011) |

Table 3b. Sequence statistics based on the determined 927 base long segment from the envelope region

| Compared sequences | Nucleotide level | | | Amino acid level | | |
|---|---|---|---|---|---|---|
| | Number of variable positions | Avg. no. of different positions (SEM) | Mean genetic distance (SEM) | Number of variable positions | Avg. no. of different positions (SEM) | p-distance (SEM) |
| Within group 1 | 13 | 2.836 (0.805) | 0.003 (0.001) | 8 | 1.600 (0.569) | 0.005 (0.002) |
| Between group 1 and patient 3 | 53 | 43.591 (6.529) | 0.049 (0.007) | 25 | 21.000 (3.977) | 0.070 (0.014) |
| Among controls of the same subtype (1a/1b) | 326 / 471 | 97.419 (4.815) / 118.640 (5.465) | 0.115 (0.006) / 0.147 (0.008) | 92 / 162 | 34.267 (3.439) / 43.480 (3.764) | 0.111 (0.011) / 0.161 (0.014) |
| Between subtype 1a controls and groups 1-2 | 349 | 103.359 (6.426) | 0.123 (0.009) | 104 | 41.979 (4.737) | 0.137 (0.015) |

## Legends to the figures

Fig. 1. Phylogenetic analysis of the NS5B region nucleotide sequences. The maximum likelihood tree includes 30 subtype 1a and 1b isolates from GenBank and 31 unrelated sequences from Hungarian HCV carriers. Sequences from the patients of the oncology are displayed in bold (p1-13), Hungarian control sequences are marked by 'c' and a number. Viral sequences from two patients (p3 and p11) were determined at two different time points (T1 and T2).

Fig. 2. Multiple alignment of HVR1 amino acid sequences from patients of the oncology and of selected HCV 1a isolates obtained from GenBank. Positions differing from the first sequence are shown with a black background.

Fig. 3. Phylogenetic analysis of the E1/E2 region nucleotide sequences. The maximum likelihood tree includes 30 subtype 1a and 1b isolates from GenBank and 18 unrelated sequences from Hungarian HCV carriers. Sequences from the patients of the oncology are displayed in bold (p1-13), Hungarian control sequences are marked by 'c' and a number. Viral sequences from two patients (p3 and p11) were determined at two different time points (T1 and T2).