Chapter 2.3

Multiple hepatitis C virus infections in human immunodeficiency virus-infected men who have sex with men
Multiple hepatitis C virus infections in human immunodeficiency virus-infected men who have sex with men

Xiomara V. Thomas¹, Bart P.X. Grady², Jan T.M. van der Meer³, Cynthia Ho¹, Joost W. Vanhommerig¹,², Sjoerd P. Rebers¹, Marc van der Valk³, Menno D. de Jong¹, Maria Prins³, Maria Prins³, Richard Molenkamp¹, Janke Schinkel¹

¹Academic Medical Center, Department of Medical Microbiology, Section of Clinical Virology, Amsterdam, The Netherlands; ²Public Health Service of Amsterdam, Department of Research, Cluster of Infectious Diseases, Amsterdam, The Netherlands; ³Department of Internal Medicine, Division of Infectious Diseases Amsterdam, The Netherlands.

Submitted for publication

Abstract

High rates of hepatitis C virus (HCV) reinfection among human immunodeficiency virus- (HIV-) infected men who have sex with men (MSM), following clearance of a primary infection, suggest absence of protective immunity. We investigated the incidence of HCV super- and reinfections in 85 HIV-infected MSM who had incident HCV infection. The occurrence of multiple infections was identified by serial sequencing of a fragment of NS5B and the HCV envelope. If the primary genotype was still present at the most recent viremic time, as indicated by the NS5B sequence analysis, serial envelope glycoprotein 2/hypervariable region 1 sequence analysis was performed to distinguish a new infection with the same genotype (clade switch) from intrahost evolution. Incidence rate and cumulative incidence of secondary infections were estimated and the effect of primary genotype (1a versus non-1) on the risk of acquiring a second infection with the same genotype was determined using Cox proportional hazards regression model. In 85 patients with a median follow-up 4.8 years, incidence rate of secondary infections was 5.39 cases/100 person-years (95% confidence interval, 3.34 to 8.26 cases/100 person-years). Cumulative incidence of genotype-switches (26.7%) was markedly higher than the cumulative incidence of clade switches (4.8%) at 5 years. In patients who had primary infection with HCV-1a, the risk of acquiring another HCV-1a infection was reduced compared with patients who had a primary non-HCV-1a infection and subsequently acquired HCV-1a (hazard ratio, 0.25; 95% confidence interval, 0.07 to 0.93).

Conclusion: The risk of acquiring a secondary infection with the primary genotype was markedly reduced compared with the risk of acquiring a secondary infection with a different genotype.
Introduction
Sexual transmission of hepatitis C virus (HCV) is rare [1]. However, during the past decade, there is increasing evidence for sexual transmission of HCV in human immunodeficiency virus- (HIV) infected men who have sex with men (MSM) [2,3]. With the reduced spontaneous clearance in HIV-infected patients [4] and more rapid progression to liver fibrosis and cirrhosis [5], the increased incidence of HCV in this population is of great concern. Nevertheless, the infection can be cured, especially during the acute phase, with reported success rates > 60% [6,7]. With new antiviral agents becoming available, success rates are expected to increase, but this may be associated with substantial cost [8]. After viral eradication, patients may become reinfected, as observed in people who inject drugs [9-12]. It is unknown whether clearance of a primary infection, either treatment-induced or spontaneously, results in protection against reinfection because studies have had inconsistent findings [12-16]. Few studies have evaluated the rate of HCV reinfection in HIV-infected MSM. We previously reported a high reinfection rate of 15.2 per 100 person-years in patients who were successfully treated for acute HCV infection [17]. This finding was confirmed and extended in a recent study from the United Kingdom that investigated the rate of reinfection in treatment-induced and spontaneous clearers of acute HCV infection [18]. The true incidence of secondary infections in HIV-infected MSM may be underestimated for several reasons in previous studies. Depending on interval of testing, secondary infections that were rapidly cleared may have been missed. In addition, patients with persistent infection were not included in both studies, whereas persistent viremia may have been caused by a superinfection with a new virus, with or without observed clearance of the primary virus. Furthermore, without genetic analysis of the relapsing virus, recurrence of viremia within 24 weeks after clearance of viremia may have been misclassified as a relapse [9,19]. Increasing the knowledge about the incidence of multiple infections and assessing the genetic relatedness of primary and successive viral strains would provide more insight into correlates of immunity against HCV and is important for vaccine development and targeted preventive strategies. Therefore, we studied the occurrence of multiple HCV infections in persistently infected patients and patients who cleared the infection with or without treatment. The presence of new infections with a different genotype (genotype switch) or different strain from the same genotype (clade switch) was assessed by sequencing the virus present during the entire viremic period.

Methods
Study population
The study population consisted of 85 HIV-infected MSM attending the HIV-treatment clinic of the Academic Medical Center in Amsterdam, the Netherlands, who acquired HCV infection sexually between 1994 and 2011. Acute HCV infection was defined by a positive HCV RNA test preceded by a negative HCV RNA test in patients without evidence of past HCV infection. Patients were either prospectively identified during acute infection or retrospectively by determining HCV RNA in anti-HCV-positive patients in stored serum samples from earlier times. Most patients were participants of the MSM Observational Study of Acute Infection with hepatitis C (MOSAIC), a prospective cohort study on acute hepatitis C infection in HIV-infected MSM [6]. Informed consent was obtained, and the study was conducted according to hospital ethical guidelines and the 2011 Dutch code of conduct for responsible use of human tissue and medical research.

Virologic assessments
The presence of HCV RNA was assessed by transcription-mediated amplification (VERSANT HCV RNA Qualitative Assay, Siemens, Malvern, PA, USA) or bDNA (VERSANT HCV RNA 3.0 Assay, Siemens). For genotyping, a 389-base pair fragment spanning positions 8616 to 8275 relative to the H77 strain (AF009606) of the NS5B region was amplified and sequenced as previously described [20]. For the detection of new infections with the same genotype (clade typing), a 590-base pair fragment from the envelope, spanning positions 1295 to 1885 relative to the H77 strain, which included the hypervariable region 1 (designated E2/HVR1), was amplified and directly sequenced. The E2/HVR1 sequences were submitted to a genetic sequence database (GenBank, National
Institutes of Health, Bethesda, MD, USA), under accession numbers KP399220 to KP399593; all NS5B sequences were submitted to the same database under accession numbers KP398885 to KP399219.

**Analysis of multiple infections**

For each patient, the first and most recent RNA positive samples were selected for NS5B genotyping. If a genotype switch was observed between the first and last RNA-positive sample, the samples taken between these 2 times were genotyped to determine the interval of genotype switch. At the most recent time, when the original genotype still was present, sequence analysis of E2/HVR1 was performed longitudinally on the stored serum samples to detect secondary infections with the same genotype (clade switch).

Phylogenetic analysis was used to determine whether the evolving primary virus still was present or replaced by a new viral strain from the same genotype. Sequences were analyzed using sequence assembly software (Codoncode, version 3.7.1, CodonCode Corp., Centerville, MA, USA), aligned using multiple sequence alignment software (Clustal X, version 2.0.11, Conway Institute, University College Dublin, Dublin, Ireland), and edited (GeneDoc, version 2.7, InFORMer Technologies, Inc., San Francisco, CA, USA). For visual inspection of sequences, maximum likelihood trees were constructed for each genotype under a Hasegawa-Kishino-Yano evolutionary model with invariant sites plus a gamma distribution of among-site rate heterogeneity as implemented in specialized software [21]. This substitution model exhibited the best-fit evolutionary substitution model for this fragment, using the Akaike information criterion in software programs [22,23]. Trees were unrooted and bootstrap values were determined from 1000 bootstrap resamplings of the original data.

To define an objective criterion for distinguishing intrahost evolution from a clade switch, nucleotide substitution rates were estimated for the E2/HVR1 genomic region for all patient-specific internal branches using a Bayesian Markov Chain Monte Carlo approach (BEAST, version 1.7.4, available at http://beast.bio.ed.ac.uk) [24]. For each genotype, the model was run separately by applying an uncorrelated relaxed lognormal molecular clock [25] and a coalescent exponential population size with a random starting tree [26,27]. The Markov Chain Monte Carlo chains were run for 10⁸ states to obtain good convergence and effective sample size > 200. The model was run while enforcing monophyly for each patient, thereby forcing the model to impose an unrealistically high nucleotide substitution rate for branches with sequences derived from a heterologous virus. Substitution rates for all branches were extracted from the maximum clade credibility tree (TreeAnnotator version 1.7.4) and constructed (FigTree version 1.4.0, both available at http://beast.bio.ed.ac.uk). Mean substitution rates from all internal branches from patient-specific monophyletic clades were calculated and a rate exceeding the mean plus 3-times the standard deviation (SD) was considered as evidence for the occurrence of a clade switch between the 2 times.

**Definitions**

Cleared infections were defined as infections with [1] absent HCV RNA for ≥ 60 days in untreated patients or [2] a sustained viral response that was achieved 24 weeks after treatment discontinuation. If a reinfection occurred within 24 weeks after the end of treatment, the response was defined as a sustained viral response. A reinfection was defined as recurrence of HCV RNA with a viral strain other than the primary strain after a cleared infection. A superinfection was defined as an infection with the detection of a new viral strain other than primary virus during follow-up, without documented HCV RNA-negative times in between. A new infection was defined as either reinfection or superinfection.

**Statistical analysis**

Data analysis was performed with statistical software (R, version 2.8, Institute for Statistics and Mathematics, Wirtschaftsuniversität Wien, Vienna, Austria) (IBM SPSS Statistics for Windows, Version 19.0, IBM Corp, Armonk, NY, USA). The incidence and confidence interval (CI) of new infections after primary infections were calculated using person time methods and reported per 100 person-years. The mid-P test was used to compare incidence rates. The date of primary HCV infection was estimated as the midpoint between the dates of the last RNA-negative and first RNA-
positive samples. Follow-up was calculated from estimated time of infection until the date when either a genotype or clade switch occurred, or the last date of HCV genotyping when no viral switch had occurred. Cumulative incidence curves were estimated within a competing-risks framework to determine the incidence of the first genotype switch compared with the first clade switch during follow-up. Cox proportional hazards regression model and log-rank test were used to evaluate the effect of genotype (1a versus non-1a) at primary infection on the risk of a secondary infection with genotype 1a. During treatment, patients were considered not to be at risk for a new infection. Statistical significance was defined by \( P \leq .05 \).

Results

Patients
At HCV acquisition, the median age was 41.6 years (interquartile range [IQR], 36.2 to 46.8 y) and median CD4+ cell count was 500 cells/µL (IQR, 393 to 638 cells/µL). The HIV load was available for 80 of 85 patients; no HIV RNA was detectable at HCV infection in 39 patients (49%). Median interval between HCV RNA testing was 2.2 months (IQR, 0.9 to 4.8 months). During the acute stage of infection, 56 of 85 patients (66%) were treated with pegylated interferon and ribavirin, and 46 of the 56 patients (82%) had sustained viral response. There were 5 of 29 patients (17%) who were not on treatment and who cleared the primary infection spontaneously.

Genotype switches
The NS5B genotyping identified 18 genotype switches. A switch from HCV-4d to HCV-1a was the most common switch (n = 9). Other observed genotype switches were 1a to 4d (n = 3), 1a to 2b (n = 2), 1a to 3a (n = 1), 1b to 4d (n = 1), 1a to 1b (n = 1), and 3a to 1a (n = 1). Genotype switches occurred in treated and untreated patients after cleared or persistent infections. In these 18 genotype switches, 16 genotype switches were secondary infections and 2 genotype switches were tertiary infections.

Clade switches
Given the conserved character of NS5B (9), serial sequencing of the more variable E2/HVR1 region was performed in the 69 patients who did not have a genotype switch, to determine whether changes in E2/HVR1 over time were compatible with intrahost evolution or a new infection with the same genotype (clade switch). In total, 380 sequences were generated. Median interval between sequences was 0.57 year (IQR, 0.21 to 0.99 y).

Visual inspection of maximum likelihood E2/HVR1 trees indicated that sequences clustered per patient, with some intermingling of sequences from different patients at the beginning of the infection. However, in 6 HCV-1a-infected patients, variants were detected during follow-up that formed distinct phylogenetic clusters, suggesting replacement of the primary strain with another HCV-1a strain (Figure 1). These 6 clade switches also were observed in the NS5B phylogenetic tree (data not shown).

Intermingling at the beginning of infection also was observed in the HCV-4d phylogenetic tree. Sequences derived from 1 patient occasionally formed separate clusters during follow-up (Figure 2; patient 004 and 013), which impeded distinguishing intrahost evolution from a clade switch by visual inspection only.

To define a more objective criterion for the presence of a clade switch, we estimated nucleotide substitution rates across internal branches of the tree using a Bayesian coalescent analysis, as described in the methods section (24). The mean substitution rate across all branches was 9.44e-03 substitutions/site/year (SD, 1.56e-02). Using a cutoff of 5.62e-02 substitutions/site/year for clade switch designation, which is the mean of the substitution rate + [3 X SD], 7 outliers from 6 patients were identified (Figure 3), confirming that the clusters with large intrapatient genetic distances in the HCV-1a phylogenetic tree represented clade switches. Nucleotide substitution rates above the threshold were not present in patients infected with genotypes other than 1a. Clade switches occurred as a second infection (n = 3), third infection (n = 2), and fourth infection (n = 1).
Figure 1 (left). Maximum Likelihood Phylogeny of Hepatitis C Virus (HCV) Genotype 1a. Maximum likelihood tree of all E2/HVR1 HCV-1a sequences. Bootstrap values ≥ 75 indicative of well-supported clades are shown. Symbols are patient specific. Tip labels denote patient code followed by sampling date. Multiple HCV-1a infections in the same patients are indicated on the right of the phylogenetic tree. Number of asterisks behind patient code indicates the number of infection.

Figure 2. Maximum Likelihood Phylogeny of Hepatitis C Virus (HCV) Genotype 4d. Maximum likelihood tree of all E2/HVR1 HCV-4d sequences. Bootstrap values ≥ 75 indicative of well-supported clades are shown. Symbols are patient specific.
Figure 3. Hepatitis C Virus (HCV) E2/HVR1 Substitution Rates Per Site Per Year in Human Immunodeficiency Virus-Infected Men Who Have Sex With Men. Dot plot of monophyly enforced substitution rates for all branches from all patients. The horizontal dotted line represents the cutoff (5.62e-02 substitutions/site/year [mean + (3 × SD)]). Patient 019 was infected with HCV-4d between his first and third HCV-1a infection. The apostrophe (’) indicates the infection order.

Figure 4. New Infections After Primary Hepatitis C Virus (HCV) Infection. Flow chart summarizing all observed genotype and clade switches during follow-up. Abbreviations: R, relapse; NR, nonresponse; SVR, sustained viral response. Bullets indicate the end of follow-up.
Figure 6. Cumulative Incidences for Genotype and Clade Switches and Kaplan-Meier Method for New Hepatitis C Virus (HCV) Infections After Primary HCV Infection. (A) Cumulative incidences, estimated within a competing-risks framework, for genotype and clade switches after primary HCV infection in 85 Human Immunodeficiency Virus-infected men who have sex with men. Lines represent new infections with the same genotype as the original HCV infection (clade switch, solid line) and new HCV infections with a genotype different than the genotype at primary HCV infection (genotype-switch, dashed line). (B) For all patients, Kaplan-Meier method for new HCV-1a infections (n = 11) after a primary infection with HCV-1a (solid line) (n = 51) or a non-HCV-1a primary infection (dashed line) (n = 34). Cox proportional hazards model was used to compare primary infection with HCV-1a to primary infection with non-HCV-1a.
Incidence of new infections: association with clearance and primary genotype

In total, 24 new infections, both super- and reinfections, were observed in 19 patients. Figure 4 summarizes all new infections in a flow chart in relation to treatment and treatment outcome. In 4 of 19 patients, multiple viral switches were observed. Most new infections were reinfections, following either spontaneous or treatment-induced clearance of the primary infection, with a total of 19 reinfections in patients who had total 64 previously cleared infections and 5 superinfections in 44 persistent infections. In the 5 spontaneously cleared infections, 3 reinfections occurred. Figure 5 illustrates the individual infection history of patients with multiple infections.

The overall incidence of a secondary infection was 5.39 cases per 100 person-years (95% CI, 3.34 to 8.26 person-years), with 19 secondary infections in 85 patients and median follow-up 2.87 years (IQR, 1.51 to 5.87 y) excluding the treatment period. The incidence of reinfection in patients who had a resolved primary infection (n = 51) was 14.5 per 100 person-years (95% CI, 8.41 to 23.34 person-years). In contrast, the incidence of superinfections was significantly lower, with 1.6 cases per 100 person-years (IQR, 0.5 to 3.9 person-years; P < .01.)

Cause-specific cumulative incidence curves showed that at 5 years after primary infection, the cumulative incidence of a clade switch was 4.8% (95% CI, 0.0% to 10.1%) and genotype switch was 26.7% (95% CI, 13.3% to 38.1%) (Figure 6). In addition, Cox proportional hazards regression model showed that patients with HCV-1a during primary infection had a decreased risk for acquiring HCV-1a again as secondary infection (hazard ratio, 0.25; 95% CI, 0.07 to 0.93) compared with patients who had a non-HCV-1a infection as the primary infection (log-rank test, P = .03) (Figure 6).

Discussion

With the large similarity of circulating viruses in an emerging epidemic, it can be difficult to identify new infections that have the same genotype as the original infection, especially in patients with persistent viremia or a relapse after treatment. In addition to sequencing the conserved NS5B region to identify genotype switches, we sequenced a genetically highly diverse fragment of the E2 gene. The intrahost nucleotide substitution rate of this region was determined using serial samples, and a threshold was defined for genetic divergence between 2 sequences in a specific time, enabling the distinction between new infections and intrahost evolution. This allowed us to precisely estimate the incidence of new infections with the original or different genotype in persistent or cleared acute HCV infections.

The overall incidence of secondary infections was 5.39 cases per 100 person-years. However, in patients with cleared infections, the incidence of secondary infections was 14.5 per 100 person-years. A novel and important finding is that, for the most common genotype 1a, the risk of acquiring genotype 1a infection again was markedly reduced. This suggests that, even in HIV-infected individuals, partial, genotype-specific immunity is generated from the primary infection.

The high incidence of reinfection after spontaneous or treatment-induced clearance of the primary infection observed in this study is consistent with the high reinfection rate of 15.2 cases/100 person-years observed in an earlier smaller study about the incidence of reinfection after treatment induced clearance in HIV-infected MSM with acute HCV from 2 HIV-clinics [17]. The incidence is higher than observed in a recent study from the United Kingdom [18], which showed a reinfection rate of 8.0/100 person-years in HIV-infected MSM who cleared their primary HCV infection. The higher estimate in our study might be explained by our extensive sequence analysis, which enabled us to identify clade switches in patients who relapsed within 24 weeks after the end of treatment. This resulted in the identification of additional reinfections, whereas such patients were considered not to have cleared their primary infection and were excluded in the United Kingdom study. In addition, the smaller testing interval in our study might have resulted in the identification of reinfections that would have been missed otherwise [28].
A previous study reported a nonsignificant trend toward a lower incidence of reinfection in patients with spontaneous than treatment-induced clearance [18]. Our study population included only 5 spontaneous clearers in untreated patients, and 3 of them became reinfected, suggesting that spontaneous clearance of a previous infection does not reduce the risk of reinfection.

In the present study, the replacement of the primary virus by other viruses in persistently infected patients was investigated. We cannot prove that such replacements were superinfections because they could have been caused by dynamic changes in dominance when different viral strains were present at the same time. However, we believe that the newly identified viruses were true superinfections because of the overall high incidence of new infections [29]. However, another possibility is that such superinfections may have been reinfections in which aviremic times were missed; the likelihood of this scenario is small because of the small RNA testing interval in this study. Nevertheless, the incidence of such superinfections was low (1.6 cases per 100 person-years). There are several possible explanations for a lower incidence of superinfection in persistently infected patients compared with the incidence of reinfection in patients with cleared infection. Patients who were aware of their chronic HCV infection may have been less likely to engage in high-risk sexual behavior. In addition, in the HCV cell culture system, superinfection is excluded at a post-entry step [29,30]. Furthermore, cross-neutralizing antibodies are present in most patients during the chronic phase of infection [31]. These circulating antibodies may immediately neutralize any new virus, before a new infection can be established.

The present study showed a markedly reduced incidence of new infections with the original genotype compared with new infections with a different genotype. It is unknown whether the incidence of such new infections is truly reduced or whether these infections are cleared more rapidly and may be missed because of our testing interval of 2.2 months. Reinfections are characterized by lower viral load and a shorter duration of viremia, suggesting the existence of acquired immunity [15,32]. However, our study suggests that there is a strong genotype-specific component to this acquired immunity, resulting in partial protection against the genotype present in the primary infection. Further study may show whether such genotype-specific immunity is caused by B, T, or NK cells [33]. For vaccine development, this suggests that a strategy directed at generating genotype-specific responses may be more successful than pursuing a pangenotype vaccine for a variable virus such as HCV. Other studies are needed to confirm our findings because the ongoing epidemic in Amsterdam may convey immunity only against locally circulating variants of the same genotype, with saturation of infection with these variants in the population at risk.

In conclusion, the present study confirmed the high rate of HCV reinfection after primary infection in HIV-positive MSM, highlighting the need for public health interventions in this high-risk group. In addition, this study demonstrated that observational cohort studies with frequent sampling of individuals with acute HCV infection are important to better understand the correlates of immunity against HCV [34]. Such studies may contribute to the development of a protective vaccine, which may be the most effective public health intervention.

Acknowledgements
We thank all MOSAIC study participants and other members of the MOSAIC study team: J. Arends, D. van Baarle, K. Brinkman, M. van den Ende, L. Gras, D. Kwa, T. van de Laar, F. Lambers, J. Mulder, C. Smit, and W. van der Veldt. The authors are grateful to J. Karlas for coordinating HCV testing; M. Bakker for assistance with sample retrieval, G.R. Visser for data management, M. Mutschelknauss for data collection and patient recruitment, and R. Rose for assistance with data analysis.
References


