Shared HCV Transmission Networks Among HIV-1–Positive and HIV-1–Negative Men Having Sex With Men by Ultradeep Sequencing

Thuy Nguyen, Constance Delaugerre, Marc-Antoine Valantin, Corinne Amiel, Emmanuelle Netzer, Thomas L'Yavanc, Michel Ohayon, Nadia Valin, Nesrine Day, Georges Kreplak, et al.

1 **Shared HCV Transmission Networks among HIV-1 Positive and Negative Men Having**

2 **Sex with Men by Ultra-Deep Sequencing**

3 **Running head**: **Shared HCV Transmission Networks**

4 **Authors:** Thuy NGUYEN[1*], Constance DELAUGERRE[2,3], Marc-Antoine VALANTIN[4],

5 Corinne AMIEL[5], Emmanuelle NETZER[6], Thomas L'YAVANC[7,8], Michel OHAYON [7,8],

6 Nadia VALIN[9], Nesrine DAY[10], Georges KREPLAK[10], Gilles PIALOUX[8], Vincent

7 CALVEZ[1], Jean-Michel MOLINA [3,11], Anne-Geneviève MARCELIN[1], Eve TODESCO[1]

8 **Affiliations** :

9 [1] Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique

10 (iPLESP), AP-HP, Hôpital Pitié-Salpêtrière, Laboratoire de virologie, F-75013 Paris, France ;

11 [2] AP-HP, Hôpital Saint-Louis, Laboratoire de virologie, Paris, France ;

12 [3] INSERM UMR 941, Université de Paris Diderot, Sorbonne Paris Cité ;

13 [4] Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique

14 (iPLESP), AP-HP, Hôpital Pitié-Salpêtrière, Services de maladies infectieuses et tropicales,

15 F-75013 Paris, France ;

16 [5] Sorbonne Université, Centre d'Immunologie et de Maladies Infectieuses (CIMI) UMRS

17 CR7, Persistent Viral Infection (PVI) Team, Inserm U1135, APHP, Groupe Hospitalier Paris

18 Est, Hôpital Tenon, Laboratoire de virologie, F-75020 Paris, France ;

19 [6] INSERM SC10, Villejuif, France ;

20 [7] Centre de santé sexuelle Le 190, Paris, France ;

21 [8] Sorbonne Université, APHP, Hôpital Tenon, Department of Infectious Diseases, Paris,

22 France;

23 [9] Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique

24 (iPLESP), AP-HP, Hôpital Saint Antoine, Department of Infectious Diseases, F-75012 Paris,

25 France;

26 [10] Cerballiance Laboratory, Paris, France;

27 [11] AP-HP, Hôpital Saint-Louis, Department of Infectious Diseases, Paris

28 [*]**Corresponding author**

29 Thuy NGUYEN, PharmD, PhD

30 Service de virologie, Bât CERVI, Hôpital Pitié-Salpêtrière, 83 Bd de l'Hôpital, 75013 Paris

31 Email: thithu-thuy.nguyen@aphp.fr or thuthuynguyend07@gmail.com

32 Fax: +33 1 42 17 74 11; Phone: +33 1 42 17 58 42

39

40

ABSTRACT (250 words)

Objective: Several studies reported HCV transmission networks among men having sex with men (MSM) in Europe and the spread of HCV strains from HIV-HCV co-infected toward HCV mono-infected MSM. We aimed to investigate HCV transmission dynamics among HIV-positive and HIV-negative MSM by ultra-deep sequencing (UDS).

Design and Methods: NS5B fragment (388 bp) was sequenced from virus of 50 HIV-positive and 18 HIV-negative patients diagnosed with recent HCV infection. UDS data were analysed by Geneious (version 10.3.2). Phylogenetic trees were constructed by FastTree (version 2.1) and submitted to ClusterPicker (version 1.2.3) for transmission chain detection at different thresholds of maximum genetic distance (MGD) (3% for Sanger, 3% and 4.5% for UDS).

Results: Ten, seven-teen, and eight-teen HCV transmission chains were identified by Sanger at 3%, UDS at 3% and at 4.5% of MGD, respectively. Of 68 subjects enrolled, 38 (55.9%), 38 (55.9), and 43 (65.3%) individuals were involved in transmission networks found by Sanger at 3%, UDS at 3% and at 4.5% of MGD, respectively. Mixed transmission chains including HIV-positive and HIV-negative subjects were detected for 8/10 chains by Sanger at 3%, for 9/17 by UDS at 3%, and for 10/18 by UDS at 4.5% of MGD. Overall, the number of HIV-negative individuals clustering with HIV-positive ones was 9/18 by Sanger, 9/18 by UDS at 3%, and 10/18 by UDS at 4.5% of MGD.

Conclusions: HIV-positive and HIV-negative MSM shared HCV transmission networks, which emphasizes the need for HCV surveillance and prevention measures in these communities regardless of the HIV status.

Key words: shared HCV transmission chains, recent HCV infection, men having sex with men, ultra-deep sequencing

Word counts for the abstract: 250, Word counts for the text: 2694

I.   INTRODUCTION

HCV infection among HIV-infected men who have sex with men (MSM) has become an outbreak since several years particularly in urban centres in Europe, Australia, and the United States [1–5]. Indeed, a remarkable increase in HCV incidence among HIV-positive MSM was reported by studies of longitudinal cohorts. In France, data from the large Dat'AIDS cohort showed a regular increase of HCV incidence from 4.3 to 11.1 per 1000 person-years (PY) from 2012 to 2016 in French HIV-positive MSM despite a high HCV treatment coverage and cure rate [6]. Data from a meta-analysis of 28 studies revealed a pooled incidence of HCV at 7.8 per 1000 person-year (PY) in HIV-positive MSM while it was only 0.4 per 1000 PY in HIV-negative MSM in resource-rich countries such as Europe, Australia, the United States, and Canada [7]. Of note, a high incidence of HCV infection in HIV-negative MSM (14 per 1000 PY) was seen in individuals eligible for Pre-Exposure Prophylaxis (PrEP), probably because of their high-risk behaviours [8].

Sequencing and phylogenetic analyses are powerful tools to understand transmission dynamics at molecular level. Individuals are considered to share the same transmission chain if their viral populations are more genetically similar to each other than expected by chance and demonstrated by a tight cluster on phylogenetic trees satisfying requirements of branch support value and genetic distance threshold [9]. For example, a collaborative study by phylogenetic approach enrolling HIV-positive MSM recently diagnosed with HCV infection from England, Netherlands, France, Germany, and Australia (n= 226) highlighted a large European MSM specific HCV transmission network [2]. Moreover, several studies showed the spread of HCV strains from HIV-positive toward HIV-negative MSM [10,11]. HCV antibody testing is therefore recommended for MSM at high risk of HIV infection and included in PrEP programs.

To date, phylogenetic studies are often based on Sanger sequencing method to identify transmission chains [12,13]. However, Sanger sequencing, with only one bulk or consensus sequence generated, is unable to fully characterize intra-host genetic diversity especially for RNA viruses such as HCV [14]. Moreover, HCV infections are considered to be frequently established through transmission of minority variants [15–17]; a consensus sequence cannot therefore reliably capture such transmissions. Ultra-deep sequencing (UDS) with a high throughput of sequencing data allows detecting minority viral populations down to 1% and is able to characterize in-depth viral population. Therefore, in this study, we aimed firstly to identify and characterize HCV transmission chains and secondly to detect closely related HCV transmission events by UDS among HIV-positive and HIV-negative MSM with recent HCV infection.

II.   MATERIALS AND METHODS

2.1.Study design and patients

Fifty-five patients with recent HCV infection (50 HIV-positive and 5 HIV-negative), followed at the Pitié-Salpêtrière, Saint-Antoine and Tenon hospitals, Paris, France and 13 HIV-negative patients from the ANRS IPERGAY study (Intervention for prevention of HIV acquisition by antiretroviral therapy for PrEP among gay men at high risk of HIV-1 infection) [18,19] were enrolled. Overall, six patients were enrolled between July 2012 and December 2013 and 62 between March 2014 and May 2016. All of them reside in Paris except one patient from the IPERGAY study.

The study was carried out in accordance with the Declaration of Helsinki. This work was a retrospective non-interventional study with no addition to standard care procedures. Reclassification of biological remnants into research material after completion of the ordered virological tests was approved by the local interventional review board of the three hospitals.

116   According to the French Public Health Code (CSP Article L.1121-1.1) such protocols are

117   exempted from individual informed consent.

118   In our study, patients having a positive HCV serology, and/or a detectable HCV viral load

119   (VL) associated with a negative HCV serology within the previous 12 months or having a

120   detectable HCV VL beyond 24 weeks of a successful anti-HCV treatment or a spontaneous

121   HCV clearance with subsequent reinfection by a different HCV genotype were considered as

122   recent HCV infections. Patients with a detectable HCV VL with increase of alanine

123   aminotransferase (ALT) ≥ 10 upper limit of normal (ULN) without any other etiology of

124   hepatitis or a detectable HCV VL beyond 24 weeks following a successful anti-HCV

125   treatment or spontaneous clearance with subsequent reinfection by a same HCV genotype

126   were also enrolled and considered as possible recent HCV infections.

127   Sanger and UDS were performed on frozen plasma samples and HCV transmission network

128   was constructed on the 2 datasets to compare quantity and characteristics of transmission

129   chains identified by both techniques.

130   *2.2. Extraction, amplification, and deep-sequencing*

131   Eighty microliters of HCV RNAs were extracted from 1 ml of plasma using® easyMAG®

132   (bioMérieux Clinical Diagnostics). Extracted RNAs were reverse transcribed in

133   complementary DNAs, and NS5B fragment (position 8256 to 8644 compared to H77,

134   fragment of 388 bp) was amplified by PCR in a one-step process (Superscript III One-step

135   RT-PCR with platinum Taq kit; Invitrogen, Carlsbad, CA, USA) according to the

136   manufacturers' protocol, by Forward primer: *A*TATGAYACCCGCTGYTTTGACTC-3' and

137   Reverse primer : 5'-GCNGARTAYCTVGTCATAGCCTC-3'. Samples were then multiplexed

138   with and subjected to standard Illumina Miseq paired-end sequencing at 2x250 bp.

139   *2.3. UDS data analysis*

140 UDS data were analyzed by Geneious software (version 10.3.2, http://www.geneious.com) [20].

141 Paired reads were firstly merged, primer-removed and quality-trimmed using quality

142 threshold of 30. Sequences with good quality (quality scores of 30 on at least 95% of bases)

143 were error-corrected by BBNorm from the BBtools package included in Geneious. Corrected

144 reads of each sample were *de novo* assembled by Geneious assembler with custom sensitivity

145 following different thresholds of similarity to reduce the number of reads and time for further

146 analysis while maintaining viral population diversity. Firstly, reads were assembled at 100%

147 of similarity. Reads unable to assemble at this threshold were then assembled at 99% of

148 similarity. The process continued and finished at threshold of 97% of similarity where almost

149 all reads were assembled. All contigs and unassembled sequences produced in this step were

150 grouped in one file used for phylogenetic analyses.

151 ### *2.4. Phylogenetic analysis to study transmission chains*

152 Contig sequences retrieved from clustering process were aligned to a reference sequence

153 corresponding to the subtype of the sample by "Map to Reference" function in Geneious.

154 Phylogenetic trees were constructed on Sanger dataset (all genotypes in one tree) and on UDS

155 dataset (separate tree for each genotype for better visualization) by FastTree software (version

156 2.1) [21] using generalized time-reversible as mathematics model. Transmission chains were

157 picked up by ClusterPicker software (version 1.2.3) [22] if branch support value calculated by

158 Shimodaira-Hasegawa test was superior to 0.80 and maximum genetic distance (MGD)

159 among individuals (>= 2 individuals) satisfied different levels: 3% for Sanger, and 3% or

160 4.5% for UDS.

161 ClusterPicker software was also used to detect individuals belonging to closely related

162 transmission evens, defined as a less than 0.5% of MGD among numerous sequences of

163 different individuals. Adjacent samples in the same PCR or sequencing plate were not

164 considered and presented as closely related transmission events for risk of contamination.

165 Samples were considered only when they came from different experiments or when they were

166 non-adjacent in the same experiment without any sign of contamination in other samples

167 between them.

168 Trees were visualized in MEGA7 [23].

169 **III.   RESULTS**

170 *3.1.Patients' characteristics and sequencing results*

171 The patients' characteristics are presented in table 1. Briefly, the median age was 38.5 years

172 (IQR, 30.5-46.0) and the majority were MSM (85.3%). Among the 68 patients enrolled, 15

173 were cases of HCV reinfection, three were considered as possible recent HCV infections, 50

174 (73.5%) were HIV-coinfected. Significant difference among HIV-positive and HIV-negative

175 patients was observed only for age. HCV sequencing showed genotype 1a (47.1%), 4d

176 (41.2%), 3a (8.8%), and 2k (2.9%).

177 A median of 2389 sequences (interquartile range [IQR], 1851-2960) per sample was obtained

178 after quality trimming step.

179 *3.2.Comparison of HCV transmission chains identified by UDS and Sanger*

180 At 3% of MGD, Sanger detected 10 transmission chains in which a median of 3 subjects

181 (min-max = 2-6) was identified while UDS at 3% of MGD detected 17 chains (median = 2

182 subjects, min-max = 2-5) and UDS at 4.5% of MGD detected 18 chains (median = 2 subjects;

183 min-max=2-6). The number of subjects identified within each transmission chain was not

184 statistically different among Sanger and UDS at 3% and 4.5% of MGD (p value > 0.31 using

185 the Wilcoxon signed rank test in JASP software) [24].

186 In particular, UDS allowed detection of hidden transmission chains through minority variants.

187 UDS at 3% and 4.5% of MGD allowed detection of three and four additional transmission

188 chains, respectively which were not detected by Sanger (table S2 of supplementary data). One

189 transmission chain among these was formed from one individual living in Paris and another

8

190 residing outside of Paris (from IPERGAY study). Moreover, four subjects were additionally

191 detected by UDS to be included in transmission chains (table S1 of supplementary data).

192 However, Sanger sequencing also allowed detection of one transmission chain (chain 5 in

193 table S1 of supplementary data) which was not noticed at all by UDS.

194   *3.3.Individuals inside and outside HCV transmission chains*

195 Out of 68 individuals enrolled, 38 (55.9%), 38 (55.9), and 43 (65.3%) were detected to be part

196 of transmission chains by Sanger and UDS at 3% and at 4.5% of MGD, respectively.

197 Regarding characteristics of individuals inside and outside transmission chains, statistical

198 analyses showed no significant difference between two groups for age, HIV co-infection and

199 HCV reinfection rates, proportions of MSM, and proportions of individuals infected with

200 HCV-GT1a or HCV-GT4d whatever the technique used (table 2).

201   *3.4.HCV transmission chains including HIV-positive and HIV-negative individuals*

202 HCV transmission chains including HIV-positive and HIV-negative individuals were

203 observed in 8/10 (80%) chains by Sanger, in 9/17 (52.9%) by UDS at 3%, and in 10/18

204 (55.6%) by UDS at 4.5% of MGD. Overall, among 18 HIV-negative MSM included in this

205 study, the number of HIV-negative individuals clustering with HIV-positive ones was 9 by

206 Sanger, 9 by UDS at 3%, and 10 by UDS at 4.5% of MGD. By UDS at 4.5% of MGD, 8 out

207 of 13 HIV-negative individuals (61.5%) from IPERGAY trial enrolled in this study were

208 detected to belong to transmission chains.

209   *3.5.Closely related HCV transmission events*

210 In a second analysis based on UDS data, we described individuals belonging to closely related

211 transmission events because numerous sequences of different samples were identical or

212 almost identical (MGD <0.5%). Five events considered as closely related transmission were

213 detected in transmission chains number 3, 4, 7, and 8 in table S1 of supplementary data. In

214 detail, we detected this event between individuals 9 and 10 (2 months of difference in date of

215 HCV infection); 13 and 14 in chain 3 (14 months), among individuals 15, 16, 18, and 19 in

216 chain 4 (14 months), among individuals 27, 28, and 29 in chain 7 (5 months), and among

217 individuals 30, 33, 34, and 35 (39 months) in chain 8.

218 Examples of two phylogenetic trees constructed from UDS sequences from two and three

219 individuals considered as closely related transmission events are shown in figures S1 and S2

220 of supplementary data.

221 **IV.   DISCUSSION**

222 In this work, we identified HCV transmission chains in MSM either co-infected by HIV or at

223 high risk of HIV acquisition in Paris by UDS and Sanger sequencing. Our study revealed a

224 high HCV clustering rate (from 56% to 65%) whatever the techniques used signifying a

225 dynamic transmission among them. Moreover, one patient under PrEP living outside Paris

226 was enrolled and this patient was found to be part of a transmission chain by UDS. Therefore,

227 in case of HCV infection, early initiation of treatment should be carried out in this population

228 to rapidly prevent further spread of the virus.

229 Transmission chains were identified at cut-off of 3% of MGD by Sanger and at two different

230 cut-offs of 3% and 4.5% of MGD by UDS. Indeed, few studies have conclusively established

231 the cut-off of MGD to identify a transmission chain among HCV-infected people by UDS,

232 varying from 2 to 4.5% [11,25–28]. It may be difficult to compare Sanger and UDS techniques at

233 the same cut-off of MGD because UDS allows a much deeper characterization of viral

234 diversity. In this work, NS5B deep sequencing improved the discrimination of transmission

235 chains versus Sanger sequencing which was in line with results from a study of Montoya *et al.*

236 [26]. UDS at both thresholds of MGD identified a median of two subjects within a transmission

237 chain compared to a median of three subjects by Sanger sequencing but this difference was

238 not statistically significant. Importantly, UDS allowed establishing more solid transmission

239 events and the transmission dynamics among individuals within each chain could be further

240 evaluated through detection of numerous clustered viral strains. For example, we detected five

241 transmission events considered very closely related i.e. individuals harboured viruses with

242 numerous overlapped sequences (MGD < 0.5%). Importantly, among them, some harboured

243 viruses with multiple identical sequences (MGD = 0%) suggesting direct transmission events

244 [29]. However, it is not possible to confirm direct transmission from one person to another using

245 molecular data alone. Indeed, both could be infected from a third source, or they could be

246 connected indirectly through a transmission chain including one or more intermediaries.

247 Although transmission directionality was not inferred due to lack of specific epidemiological

248 data, patients included in these events should be followed more closely including the

249 communities around them to assure a rapid intervention.

250 In this study, UDS also detected hidden transmission chains by identifying transmission

251 linkages through minority viral strains. However, the deeper characterization of viral

252 variability is also the reason why UDS did not detect one transmission chain found by Sanger

253 technique. The MGD among sequences of the three individuals involved in this transmission

254 chain was 2.64% with Sanger sequencing while the genetic distance among viral sequences of

255 the three individuals is higher than the MGD threshold of 3% and 4.5% with UDS. Therefore,

256 UDS did not capture the transmission linkage among these individuals as Sanger did. Further

257 studies would be necessary to determine the most suitable MGD cut-off for transmission

258 chain identification by UDS. Thereby, UDS would be interesting to deeply characterize

259 transmission patterns such as directness or directionality among individuals; however, it is not

260 more useful than Sanger sequencing in term of large-scale prevention and rapid intervention.

261 Importantly, depending on the techniques and MGD cut-off used, 53% to 80% of transmission

262 chains identified included both HIV-positive and HIV-negative subjects and more than 50%

263 of HIV-negative subjects enrolled in this study clustering with HIV-positive ones. The shared

HCV transmission networks among HIV-positive and HIV-negative MSM were also observed in two studies conducted in Amsterdam, the Netherlands and in Lyon, France [10,11]. Our results raise an alert for better screening, monitoring, and surveillance of HCV infection in this high-risk community regardless of the HIV status.

Even though the HCV reinfection rate in subjects inside transmission chains was not statistically higher than in subjects outside them, the need of follow-up for possible HCV reinfection and of patient support and education to prevent HCV reinfection and transmission arise in this high-risk population. Last but not least, 61.5% of HIV-negative individuals under PrEP enrolled in this study were detected to belong to transmission chains by UDS at 4.5% of MGD. Therefore, surveillance of HCV infection by HCV viral load instead of anti-HCV antibody test would be more advantageous to rapidly intervene and control transmission for those in PrEP programs.

An interesting study by Caro-Pérez *et al.* has showed an HCV outbreak in HIV-positive MSM in Barcelona related to a previously described European MSM transmission network [30]. For that reason, our high throughput of HCV sequencing data from HIV-positive and HIV-negative MSM in Paris will be further investigated to study the transmission network of these populations with HCV sequences of other MSM at European level [2].

One limitation of the study is that 14.7% of patients had unknown sexual orientation. However, they all engaged in risky behaviours for multiple HCV exposures as other MSM enrolled in the study, demonstrated by their HIV co-infection status, HCV reinfection rate, and other sexual transmission diseases discovered with their HCV infection. Furthermore, individuals with unknown sexual orientation were also involved in transmission chains and no significant difference was observed in proportions of MSM or unknown sexual orientation subjects among those inside and outside transmission chains. Another limitation in our study was the length of fragment NS5B sequenced. Indeed, a longer sequence could permit more

289   accurate differentiation of linked or unlinked virus and thus more exactly identify

290   transmission chains [31]. In our study, a quite short fragment of NS5B was amplified but it was

291   counterbalanced by a high depth of coverage by UDS. Furthermore, this strategy had been

292   applied in different settings [15,32]. In this study, we did not have enough epidemiological data

293   to confirm the true transmission events among individuals. However, the fact that almost all

294   individuals (67/68) were from Paris justified in some way their epidemiological connection as

295   well as the fact that transmission chains identified by UDS were established through multiple

296   clustered viral strains increased the likeliness of true transmission event identification.

297   In conclusion, in this study, a high clustering rate of HCV was observed in HIV-positive and

298   HIV-negative MSM communities in Paris, particularly those engaged in PrEP program.

299   Furthermore, HIV-positive MSM shared HCV transmission networks with HIV-negative

300   MSM. The more frequently screening and surveillance of HCV infection regardless of the

301   HIV status is essential to prevent the spread of HCV in these high-risk communities.

302

References

318 1. Götz HM, van Doornum G, Niesters HG, den Hollander JG, Thio HB, de Zwart O. A
319   cluster of acute hepatitis C virus infection among men who have sex with men--results
320   from contact tracing and public health implications. *AIDS Lond Engl*. 2005;19(9):969-
321   974.

322 2. van de Laar T, Pybus O, Bruisten S, et al. Evidence of a large, international network of
323   HCV transmission in HIV-positive men who have sex with men. *Gastroenterology*.
324   2009;136(5):1609-1617.

325 3. Giraudon I, Ruf M, Maguire H, et al. Increase in diagnosed newly acquired hepatitis C in
326   HIV-positive men who have sex with men across London and Brighton, 2002-2006: is
327   this an outbreak? *Sex Transm Infect*. 2008;84(2):111-115. doi:10.1136/sti.2007.027334

328 4. van de Laar TJW, van der Bij AK, Prins M, et al. Increase in HCV incidence among men
329   who have sex with men in Amsterdam most likely caused by sexual transmission. *J
330   Infect Dis*. 2007;196(2):230-238. doi:10.1086/518796

331 5. Gamage DG, Read TRH, Bradshaw CS, et al. Incidence of hepatitis-C among HIV
332   infected men who have sex with men (MSM) attending a sexual health service: a cohort
333   study. *BMC Infect Dis*. 2011;11:39. doi:10.1186/1471-2334-11-39

334 6. Pradat P, Huleux T, Raffi F, et al. Incidence of new hepatitis C virus infection is still
335   increasing in French MSM living with HIV. *AIDS Lond Engl*. February 2018.
336   doi:10.1097/QAD.0000000000001789

337 7. Ghisla V, Scherrer AU, Nicca D, Braun DL, Fehr JS. Incidence of hepatitis C in HIV
338   positive and negative men who have sex with men 2000-2016: a systematic review and
339   meta-analysis. *Infection*. 2017;45(3):309-321. doi:10.1007/s15010-016-0975-y

340 8. Gras J. HCV RNA and Antigen Detection for Diagnosis of Acute Hepatitis C Among
341   MSM on PrEP. Poster 585 presented at the: Conference on Retroviruses and
342   Opportunistic Infections (CROI); March 4, 2018; Boston, the USA.
343   http://natap.org/2018/CROI/croi_152.htm. Accessed July 10, 2018.

344 9. Hué S, Clewley JP, Cane PA, Pillay D. HIV-1 pol gene variation is sufficient for
345   reconstruction of transmissions in the era of antiretroviral therapy. *AIDS Lond Engl*.
346   2004;18(5):719-728.

347 10. Hoornenborg E, Achterbergh RCA, Schim van der Loeff MF, et al. MSM starting
348   preexposure prophylaxis are at risk of hepatitis C virus infection. *AIDS Lond Engl*.
349   2017;31(11):1603-1610. doi:10.1097/QAD.0000000000001522

350 11. Charre C, Cotte L, Kramer R, et al. Hepatitis C virus spread from HIV-positive to HIV-
351   negative men who have sex with men. *PloS One*. 2018;13(1):e0190340.
352   doi:10.1371/journal.pone.0190340

353 12. González-Candelas F, Bracho MA, Wróbel B, Moya A. Molecular evolution in court:
354   analysis of a large hepatitis C virus outbreak from an evolving source. *BMC Biol*.
355   2013;11:76. doi:10.1186/1741-7007-11-76

356   13.   Féray C, Bouscaillou J, Falissard B, et al. A novel method to identify routes of hepatitis
357   C virus transmission. *PloS One*. 2014;9(1):e86098. doi:10.1371/journal.pone.0086098

358   14.   Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. *Microbiol Mol Biol Rev*
359   *MMBR*. 2012;76(2):159-216. doi:10.1128/MMBR.05023-11

360   15.   Wang GP, Sherrill-Mix SA, Chang K-M, Quince C, Bushman FD. Hepatitis C virus
361   transmission bottlenecks analyzed by deep sequencing. *J Virol*. 2010;84(12):6218-6228.
362   doi:10.1128/JVI.02271-09

363   16.   Bull RA, Luciani F, McElroy K, et al. Sequential bottlenecks drive viral evolution in
364   early acute hepatitis C virus infection. *PLoS Pathog*. 2011;7(9):e1002243.
365   doi:10.1371/journal.ppat.1002243

366   17.   D'Arienzo V, Moreau A, D'Alteroche L, et al. Sequence and functional analysis of the
367   envelope glycoproteins of hepatitis C virus variants selectively transmitted to a new host.
368   *J Virol*. 2013;87(24):13609-13618. doi:10.1128/JVI.02119-13

369   18.   Molina J-M, Capitant C, Spire B, et al. On-Demand Preexposure Prophylaxis in Men at
370   High Risk for HIV-1 Infection. *N Engl J Med*. 2015;373(23):2237-2246.
371   doi:10.1056/NEJMoa1506273

372   19.   Molina J-M, Charreau I, Spire B, et al. Efficacy, safety, and effect on sexual behaviour
373   of on-demand pre-exposure prophylaxis for HIV in men who have sex with men: an
374   observational cohort study. *Lancet HIV*. 2017;4(9):e402-e410. doi:10.1016/S2352-
375   3018(17)30089-9

376   20.   Kearse M, Moir R, Wilson A, et al. Geneious Basic: an integrated and extendable
377   desktop software platform for the organization and analysis of sequence data.
378   *Bioinforma Oxf Engl*. 2012;28(12):1647-1649. doi:10.1093/bioinformatics/bts199

379   21.   Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees
380   for large alignments. *PloS One*. 2010;5(3):e9490. doi:10.1371/journal.pone.0009490

381   22.   Ragonnet-Cronin M, Hodcroft E, Hué S, et al. Automated analysis of phylogenetic
382   clusters. *BMC Bioinformatics*. 2013;14:317. doi:10.1186/1471-2105-14-317

383   23.   Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis
384   Version 7.0 for Bigger Datasets. *Mol Biol Evol*. 2016;33(7):1870-1874.
385   doi:10.1093/molbev/msw054

386   24.   JASP Team. JASP (Version 0.9) [Computer software].2018.

387   25.   Olmstead AD, Joy JB, Montoya V, et al. A molecular phylogenetics-based approach for
388   identifying recent hepatitis C virus transmission events. *Infect Genet Evol J Mol*
389   *Epidemiol Evol Genet Infect Dis*. 2015;33:101-109. doi:10.1016/j.meegid.2015.04.017

390   26.   Montoya V, Olmstead A, Tang P, et al. Deep sequencing increases hepatitis C virus
391   phylogenetic cluster detection compared to Sanger sequencing. *Infect Genet Evol J Mol*
392   *Epidemiol Evol Genet Infect Dis*. 2016;43:329-337. doi:10.1016/j.meegid.2016.06.015

393 27. Campo DS, Xia G-L, Dimitrova Z, et al. Accurate Genetic Detection of Hepatitis C
394   Virus Transmissions in Outbreak Settings. *J Infect Dis*. 2016;213(6):957-965.
395   doi:10.1093/infdis/jiv542

396 28. Rose R, Lamers SL, Massaccesi G, et al. Complex patterns of Hepatitis-C virus
397   longitudinal clustering in a high-risk population. *Infect Genet Evol J Mol Epidemiol Evol*
398   *Genet Infect Dis*. 2018;58:77-82. doi:10.1016/j.meegid.2017.12.015

399 29. Romero-Severson EO, Bulla I, Leitner T. Phylogenetically resolving epidemiologic
400   linkage.   *Proc   Natl   Acad   Sci   U   S   A*.   2016;113(10):2690-2695.
401   doi:10.1073/pnas.1522930113

402 30. Caro-Pérez N, Martínez-Rebollar M, Gregori J, et al. Phylogenetic analysis of an
403   epidemic outbreak of acute hepatitis C in HIV-infected patients by ultra-deep
404   pyrosequencing. *J Clin Virol Off Publ Pan Am Soc Clin Virol*. 2017;92:42-47.
405   doi:10.1016/j.jcv.2017.05.008

406 31. Lamoury FMJ, Jacka B, Bartlett S, et al. The Influence of Hepatitis C Virus Genetic
407   Region on Phylogenetic Clustering Analysis. *PloS One*. 2015;10(7):e0131437.
408   doi:10.1371/journal.pone.0131437

409 32. Gonçalves Rossi LM, Escobar-Gutierrez A, Rahal P. Multiregion deep sequencing of
410   hepatitis C virus: An improved approach for genetic relatedness studies. *Infect Genet*
411   *Evol   J   Mol   Epidemiol   Evol   Genet   Infect   Dis*.   2016;38:138-145.
412   doi:10.1016/j.meegid.2015.12.020

413