Title

HIV co-infection is associated with increased transmission risk in patients with chronic hepatitis C virus

Journal

Journal of Viral Hepatitis, 26(11)

ISSN

1352-0504

Authors

Ragonnet-Cronin, Manon
Hostager, Reilly
Hedskog, Charlotte
et al.

Publication Date

2019-11-01

DOI

10.1111/jvh.13160

Peer reviewed

HHS Public Access

Author manuscript

J Viral Hepat. Author manuscript; available in PMC 2020 November 01.

Published in final edited form as:

J Viral Hepat. 2019 November ; 26(11): 1351–1354. doi:10.1111/jvh.13160.

HIV co-infection is associated with increased transmission risk in patients with chronic hepatitis C virus

Manon Ragonnet-Cronin[1,2], **Reilly Hostager**[1], **Charlotte Hedskog**[3], **Ana Osinusi**[3], **Eugenia Svarovskaia**[3], **Joel O. Wertheim**[1,*]

[1]Department of Medicine, University of California San Diego, San Diego, California, USA

[2]Current affiliation: Department of Infectious Disease Epidemiology, Imperial College London, London, UK

[3]Gilead Sciences, Foster City, California, USA

Abstract

Molecular epidemiological analysis of viral pathogens can identify factors associated with increased transmission risk. We investigated the frequency of genetic clustering in a large dataset of NS34A, NS5A, and NS5B viral sequences from patients with chronic hepatitis C virus (HCV) genotypes 1–6 infection. Within a subset of patients with longitudinal samples, Receiver Operator Characteristic (ROC) analysis was applied which identified a threshold of 0.02 substitutions/site as most appropriate for clustering. From the 7,457 patients with chronic HCV infection included in this analysis, we inferred 256 clusters comprising 541 patients (7.3%). We found that HCV/HIV co-infection, young age, and high HCV viral load were all associated with increased clustering frequency, an indicator of increased transmission risk. In light of previous work on HCV/HIV co-infection in acute HCV cohorts, our results suggest that patients with HCV/HIV co-infection may disproportionately be the source of new HCV infections and treatment efforts should be geared towards elimination in this vulnerable population.

Introduction

Molecular cluster analysis of viral pathogens can identify groups of patients with evidence of a shared, recent transmission history. The composition of these clusters can identify correlates of elevated transmission risk, as has been shown for both the human immunodeficiency virus (HIV) and hepatitis C virus (HCV) [1–4]. A previous molecular investigation into a cohort of patients with acute HCV infection found high frequencies of clustering (21%) [1], and patients in this cohort who were co-infected with HIV at the time of their HCV diagnosis were significantly more likely to be clustered in their analysis. Network-informed simulations across this cohort suggested that the use of curative, direct acting antivirals (DAAs) targeted at HCV acute cases with HIV co-infection would prevent 2.5 as many future HCV infections as non-targeted treatment among patients with acute

[*]**To whom correspondence should be addressed:** jwertheim@ucsd.edu.

Conflicts of interest

None declared.

HCV infection. However, it is unclear whether a similar strategy would be beneficial in a population of patients with chronic HCV infection.

Here, we explored molecular clustering in a population of 7,457 patients across 24 countries chronically-infected with HCV genotypes 1–6 to determine the factors associated with higher HCV transmission.

Methods

Data acquisition

HCV sequences were obtained from patients with chronic HCV infection treated with sofosbuvir with ribavirin in Phase II/III clinical trials run by Gilead Sciences (Foster City, CA). The clinical trials were conducted in accordance with the Declaration of Helsinki, Good Clinical Practice guidelines, and local regulatory requirements. All patients provided written informed consent. Patients were sampled before the initiation of treatment. Sequencing of HCV NS34A (n=1,605), NS5A (n= 6,462), and NS5B (n=6,645) was conducted as previously described and genotypes and subtypes were assigned through BLAST alignment to a set of reference sequences [3]. A subset of patients was sampled longitudinally while off therapy (n=211), with sequences generated at various timepoints (follow-up weeks 12, 24, 36, 48, 96, and/or 144). Median time between first and last visit was 96 weeks, and mean time between visits was 27 weeks. Sequences were aligned for each genomic segment separately using MAFFT [5]. All alignments are available as online Supplementary Material.

Alongside sequence data, epidemiological and clinical data were available for each patient, including sample date, country of sampling (24 countries total), age, sex, race/ethnicity (i.e., Asian, white, African American, Hispanic, other), HIV co-infection, and HCV viral load.

Optimization of clustering threshold

Among the patients with longitudinal samples, Receiver Operator Characteristic (ROC) analysis was applied to select the most appropriate genetic distance threshold for clustering, as has been performed previously [2]. The optimal threshold would maximize co-clustering of sequences from the same patients (true positives) and minimize clustering between patients (false positives). We excluded a single patient with evidence of super infection based on a change of viral subtype from 1b to 1a. We explored genetic distance thresholds between 0.005 through 0.05 substitutions/site (in 0.001 substitutions/site increments), each time counting the number of true positives, true negatives, false positives and false negatives to estimate sensitivity and specificity. The ROC analysis was conducted separately for each of the three genomic segments and the most common genotypes: 1–4. ROC curves for sensitivity and specificity indicated that 0.02 substitutions/site was the most appropriate distance threshold for clustering across all subtypes and genomic segments, with sensitivity ranging from 95% to 100% and specificity ranging from 96% to 100% (Supplementary Figure 1).

Cluster identification and statistical analysis

Maximum likelihood phylogenies were reconstructed independently for each genotype (1–6) and each HCV genomic segment (NS34A, NS5A and NS5B) in RAxML, under a GTR+I+G model [6] (phylogenetic trees are available in Supplementary Material). Clusters were identified using the Cluster Picker [7], with ambiguous bases treated as matches according to IUPAC nomenclature. Patients were classified as clustered if any genomic segment from their virus were ≤0.02 substitutions/site divergent from the same genomic segment of a virus from another patient. We did not enforce a bootstrap support threshold for defining a cluster in the Cluster Picker, because the HCV phylogeny lacks the long, internal branches that are necessary to make this type of inference. Multivariate and univariate logistic regression were used to determine epidemiological and clinical factors associated with clustering. Age and viral load were reclassified as categorical variables; countries were amalgamated into continents and collapsed with race/ethnicity. As sensitivity analyses, we explored the robustness of our findings at clustering thresholds of 0.01 and 0.03 substitutions/site. All statistical analyses (including ROC analysis) were conducted in R.

Results

Optimization of clustering threshold

The longitudinal dataset comprised 210 patients with chronic HCV infection: genotypes 1 (n=148), 2 (n=3), 3 (n=58), 4 (n=1). The median number of visits per patient was 5, ranging between 2 and 7. The most appropriate threshold for our clustering analysis was determined to be 0.02 substitutions/site, as described in the Methods section.

HCV clustering analysis

The dataset used for clustering comprised sequences from 7,457 patients (see Figure 1 and Supplementary Table 1 for epidemiological, clinical, and genotype breakdown). Countries were merged into regions and combined with race/ethnicity for analysis. The majority of sequences were genotype 1 (61.9%), from North America (60.6%), and from HIV-negative patients (94.7%). The subtype breakdown of sequences is shown in Supplementary Table 2.

At a threshold of 0.02 substitutions/site, 541 patients clustered (7.3%) into 256 clusters (Supplementary Figure 2). For 481/541 clustered individuals, more than one genomic region had been sequenced: 127 had 3 regions sequenced, and 354 had NS5A and NS5B sequences. Among these 481 individuals, only 29 clustered across all sequenced regions. NS5B sequences were more likely to cluster than sequences from other genomic regions (Supplementary Figure 3). Analyzing all genotypes in a combined multivariate regression analysis, the strongest correlates with clustering were HIV co-infection ($p<0.001$), younger age (age 12–19; $p<0.05$), and higher HCV viral loads ($>10^7$; $p<0.05$) (Figure 1). The same predictors were also associated with clustering in the univariate analyses (Supplementary Table 3). The effect of HIV co-infection was progressive with cluster size: individuals clustering with two or more individuals (n=65) were 3.02 times more likely to be HIV positive than those clustering with one or none ($p<0.01$) in a multivariate analysis. We then performed multivariate regression analysis independently on each of the most common genotypes. In the genotype 1-only analysis, we observed the association between clustering

and HIV co-infection ($p<0.001$) and higher viral loads ($p<0.05$), but not age—indicating that the age association was led predominantly by genotypes 2–6. However, for genotypes 2, 3, and 4, statistical associations were weak and unsubstantiated (data not shown) likely because of smaller sample sizes. Of 397 HIV co-infected patients, 388 (97.8%) were sampled in North America and 362 (91.2%) were infected with HCV genotype 1.

We then conducted a sensitivity analysis by rerunning the multivariate regression analyses at cluster genetic distance cutoffs of 1% and 3%, in which 85 (1.1%) and 1805 (24.21%) patients clustered, respectively (Supplementary Table 4). The association between clustering and HIV co-infection, younger age, and higher HCV viral load were generally robust. The effect of clustering thresholds was the same across all genomic regions (Supplementary Figure 3).

Discussion

We investigated demographic and clinical factors associated with clustering in an HCV molecular transmission network in a dataset of sequences sampled internationally from chronically-infected HCV patients. The factors most strongly associated with clustering, an indicator of elevated transmission risk, were HIV co-infection, younger age, and higher HCV viral loads.

Of these associations with increased clustering frequency, HIV co-infection was the greatest in magnitude and the most consistent. Of the 397 HIV co-infected patients, 91.2% were infected with HCV genotype 1. Sensitivity analyses performed on the genotype 1-only dataset demonstrated that as the genetic distance threshold becomes more conservative, from 2% to 1%, the adjusted odds of clustering for patients with HIV co-infection increased from 2.43 to 3.49. Notably, 97.8% of these 397 HIV co-infected patients were from North America. Therefore, we are unable tease apart whether the association between HIV co-infection and HCV clustering was specific to genotype 1 or to North America.

We caution against overinterpreting the epidemiological relevance of higher clustering frequencies in younger patients. In HIV clustering studies, it has been shown that these types of approaches are biased towards clustering patients with shorter time since infection [8]. Given the decades that patients can live with chronic HCV infection, it is likely that the increased clustering frequency observed in patients age 12–19 reflects this shorter time since infection.

The frequency of clustering in the aforementioned Australian study of clustering in an HCV cohort of acutely infected (23.7%) [1] is similar to that in our analysis (24.2% when analyzed at 0.03 substitutions/site for consistency); however, a direct comparison of these clustering frequencies is potentially problematic as the Australian study analyzed the Core and E1 genes in acutely infected patients. In addition, cluster size distribution between the Australian study and the present study differed dramatically: in the Australian study nearly half of clusters comprised more than two subjects, in comparison with only 8% of our clusters, suggesting a different underlying epidemiological situation. Another HCV clustering study of a local cohort of chronically infected patients found substantially higher

rates of clustering [2]. We acknowledge that the high level of geographic diversity in our analysis, including patients from all over the world, likely biased clustering frequencies downward. Furthermore, comparisons across studies are complicated by the use of differing clustering algorithms, for example Welzel *et al.* defined cluster in the phylogeny based on inter-cluster distances [3], whereas we defined clusters in phylogenies based on maximum within cluster genetic distance [7]. Rose *et al* used both sequence-based genetic distance and a phylogeny-based distance method [2].

The Australian study also detected a strong association between clustering and HIV co-infection (adjusted odds ratio=4.56) [1]. Importantly, the HCV genotypes analyzed in that study were more varied than those from the HIV co-infected analyzed here. Although our clustering approach cannot distinguish between increased odds of being the source or recipient of HCV infection based on HIV co-infection status, the Australian study was able to conclude that HCV/HIV co-infected patients were more likely to be the source of new HCV cases. Therefore, the consistent prominence of HIV co-infection on clustering in both chronic and acute HCV infections suggests that there is a need to focus prevention/treatment efforts in patients coinfected with HCV/HIV as an important subgroup in the push towards HCV elimination. Future research should focus on whether the effect of HIV co-infections differs across modes of HCV transmissions (sexual vs. injecting drugs).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Ben Murrell for comments on the study design. MRC, RH, and JOW were supported by a research grant to their institution from Gilead Sciences. JOW was also supported in part by an NIH-NIAID K01 Career Development Award (K01AI110181) and an NIH-NIAID R01 (AI136056).

References

1. Bartlett SR, Wertheim JO, Bull RA, Matthews GV, Lamoury FM, Scheffler K, Hellard M, Maher L, Dore GJ, Lloyd AR et al.: A molecular transmission network of recent hepatitis C infection in people with and without HIV: Implications for targeted treatment strategies. J Viral Hepat 2017, 24(5):404–411. [PubMed: 27882678]
2. Rose R, Lamers SL, Massaccesi G, Osburn W, Ray SC, Thomas DL, Cox AL, Laeyendecker O: Complex patterns of Hepatitis-C virus longitudinal clustering in a high-risk population. Infect Genet Evol 2018, 58:77–82. [PubMed: 29253674]
3. Welzel TM, Bhardwaj N, Hedskog C, Chodavarapu K, Camus G, McNally J, Brainard D, Miller MD, Mo H, Svarovskaia E et al.: Global epidemiology of HCV subtypes and resistance-associated substitutions evaluated by sequencing-based subtype analyses. J Hepatol 2017, 67(2):224–236. [PubMed: 28343981]
4. Hue S, Pillay D, Clewley JP, Pybus OG: Genetic analysis reveals the complex structure of HIV-1 transmission within defined risk groups. Proceedings of the National Academy of Sciences of the United States of America 2005, 102(12):4425–4429. [PubMed: 15767575]
5. Katoh K, Standley DM: MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 2013, 30(4):772–780. [PubMed: 23329690]
6. Stamatakis A: RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 2006, 22(21):2688–2690. [PubMed: 16928733]

7. Ragonnet-Cronin M, Hodcroft E, Hue S, Fearnhill E, Delpech V, Brown AJ, Lycett S, UK HIV Drug Resistance Database: Automated analysis of phylogenetic clusters. BMC Bioinformatics 2013, 14:317. [PubMed: 24191891]

8. Volz EM, Koopman JS, Ward MJ, Brown AL, Frost SD: Simple epidemiological dynamics explain phylogenetic clustering of HIV from patients with recent infection. PLoS Comput Biol 2012, 8(6):e1002552. [PubMed: 22761556]

Attribute	Value	Total	Clustered (%)	AOR
HIV co-infection	No	7060	471 (6.7%)	Ref.
	Yes	397	70 (17.6%)	2.1***
Age	12-19	12	3 (25%)	4.5*
	20-24	69	8 (11.6%)	2.1
	25-29	141	8 (5.7%)	0.9
	30-39	480	26 (5.4%)	0.8
	40-49	1182	97 (8.2%)	1.1
	50-59	3005	233 (7.8%)	Ref.
	60+	2081	140 (6.7%)	0.9
HCV genotype	1	4614	417 (9%)	Ref.
	2	863	16 (1.9%)	0.2***
	3	1574	68 (4.3%)	0.4***
	4	275	38 (13.8%)	1.8**
	5	67	2 (3%)	0.4
Sex	Male	4721	165 (3.5%)	Ref.
	Female	2736	376 (13.7%)	0.8*
Viral load (copies/ml)	$<10^4$	35	2 (5.7%)	0.4
	10^4-10^5	288	17 (5.9%)	0.7
	10^5-10^6	1718	98 (5.7%)	0.7
	10^6-10^7	4176	305 (7.3%)	Ref.
	$>10^7$	1240	119 (9.6%)	1.3

AOR, Adjusted odds ratio; CI, confidence interval

*** $p<0.001$; ** $p<0.01$; * $p<0.05$

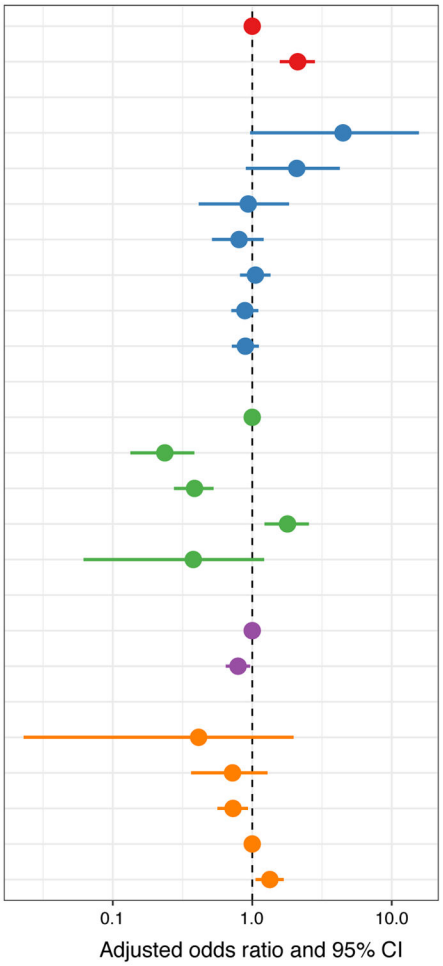

Figure 1:

Demographic, clinical, and genotypic breakdown for HCV-infected trial participants and their association with HCV clustering.