The Journal of Infectious Diseases
EDITORIAL COMMENTARY

Digging Deeper Into Hepatitis C Virus Outbreaks

Andrea D. Olmstead

University of British Columbia and BC Centre for Disease Control, Vancouver, Canada

(See the major article by Campo et al on pages 957–65.)

Keywords. hepatitis C virus; outbreaks; deep sequencing; transmission networks; end-point limiting-dilution PCR; genetic heterogeneity.

Methods for investigating outbreaks of hepatitis C virus (HCV) have evolved since the discovery of the virus in 1989. Early investigations focused on epidemiological evidence such as likelihood of exposure, risk factors, temporal information, and serological evidence to identify members of an outbreak [1–3]. Streamlined nucleic acid sequencing methods provided an important means of verifying suspected transmission events, and, over time, sequencing has become a standard outbreak investigation tool [4–6]. When the viral sequences from different individuals are found to be identical or very similar, the probability is very high that those individuals are part of the same transmission chain. Unfortunately, interpreting the relationship between viral sequences in different individuals is not always straightforward.

RNA viruses such as HCV mutate very rapidly causing the sequences between the source and recipient of a transmission to quickly diverge. The magnitude of this divergence depends on the amount of time elapsed and on the region of the virus sequenced. This story is further complicated by the fact that, because of this rapid evolutionary rate, individuals are actually infected with entire swarms of nonidentical but closely related HCV variants, termed quasispecies [7]. Quasispecies are composed of one or a few majority variants, along with several minority variants. Sanger-based consensus sequencing can be used to identify the major HCV variants in an individual, but it is not always the major variants that establish infection in a new host following transmission [8, 9]. These features complicate HCV outbreak investigations, hampering efforts to develop standard methods for differentiating outbreaks from sporadic or unrelated infections.

The report by Campo et al in this issue of *The Journal of Infectious Diseases* attempts to tackle these issues by using an impressive sample set comprising 127 cases from 32 previously well-characterized HCV outbreaks. In the past, many HCV outbreaks have been investigated using sequencing; however, the viral genes sequenced, the analytical methods used, and the criteria used to define members of an outbreak varied from study to study [4, 5, 10, 11]. The study by Campo et al is the first comprehensive sequencing-based analysis that attempts to evaluate a large number of well-defined HCV outbreaks to identify standard genetic criteria that can be applied to investigate future outbreaks.

In addition to this rare sample set, the authors used end-point limiting-dilution polymerase chain reaction analysis to obtain sequences from multiple clones of the HCV hypervariable 1 region in each individual (approximately 40 clones/individual). This resulted in a data set that represents the quasispecies population in each individual more accurately than consensus sequencing. The relationship between each person's population of sequences was evaluated using 3 measures of genetic distance: (1) Hamming distance, (2) Nei-Tamura distance, and (3) patristic distance. The minimal distance between all variants, the average distance, and the distance between major variants was determined for each of the 3 measures. The goal was to determine which criteria allowed for the greatest differentiation between members of an outbreak versus unrelated cases.

As expected, many sequences (20.86%) from different people within the same outbreak were identical; however, many sequences differed by 1 or more nucleotides. When the genetic distances between different sequence populations were compared, minimal Hamming distance and minimal patristic distance were found to have the greatest ability to differentiate outbreak from nonoutbreak cases. This finding has important implications for future outbreak investigations. Firstly, the finding that Hamming distance (ie, the number of nucleotide differences between 2 sequences) performs just as well as patristic distance and simplifies the process of future outbreak investigations. Currently, many sequence-based outbreak investigations use phylogenetic methods, including consideration of patristic distance, to investigate the relationship between

Received and accepted 11 November 2015; published online 17 November 2015.

Correspondence: A. D. Olmstead, The University of British Columbia and The BC Centre for Disease Control, 655 W 12th Ave, Vancouver, BC V5Z4R4, Canada (andrea.olmstead@bccdc.ca).

The Journal of Infectious Diseases® 2016;213:880–2
 DOI: 10.1093/infdis/jiv543

sequences. This requires a certain level of expertise, and the process can be very time-consuming depending on the software used and the number of sequences interrogated. Hamming distance, on the other hand, is relatively easy to calculate and appears to provide both the performance and speed required in an outbreak investigation. Secondly, the fact that the minimum distance between sequence populations performs best for differentiating outbreak from nonoutbreak cases exemplifies the importance of obtaining several sequences per individual as opposed to 1 sequence per individual. When the distance between major variants is the sole consideration (ie, with consensus sequencing), transmission events involving minority HCV variants can be missed.

A minimal Hamming distance of 3.77% was identified as the threshold that captured all members of an outbreak with 100% accuracy. In theory, any laboratory aiming to investigate an outbreak could apply this threshold to their data set, provided that they sequence the same genetic region used in this study. This method also provides a means for investigating the source of an outbreak. The authors found that, in cases where the source was known, the intrahost diversity in the source was always greater than in the other members of an outbreak. This is related to the fact that HCV diversity increases over time and that, in general, the individual in an outbreak who has been infected the longest will display the greatest intrahost diversity. However, caution must be exercised when interpreting these data because if the source of an outbreak is not sampled, the member of an outbreak with the greatest duration of infection could be falsely implicated as the source.

One potential limitation of the method presented by Campo et al is the requirement to obtain multiple sequences for each individual in an outbreak. Sequencing several clones for each sample can be a time-consuming task, depending on the number of people implicated. As an alternative, the authors determined that the 3.77% threshold was also applicable for identifying members of an outbreak, using deep-sequencing data. In fact, because of a greater sampling of variants in each individual, deep sequencing improved the sensitivity of transmission detection, owing to an increase in the number of variants found to have a distance below the threshold. Increased widespread use of deep sequencing is resulting in standardization of laboratory protocols and bioinformatics methods, dropping costs, and increasing numbers of samples that can be concurrently multiplexed [12, 13]. Thus deep sequencing is expected to play a more prominent role in future outbreak investigations.

The method presented in this study will undoubtedly have great utility in investigating nosocomial outbreaks of HCV infection, but perhaps its greatest impact could be to identify transmission networks of HCV among high-risk populations, such as in people who inject drugs [14]. With the approval of highly effective anti-HCV therapies, the opportunity to eliminate HCV is becoming a real possibility. Because of limited financial resources, these new treatment options are often prioritized to individuals with advanced liver damage. However, arguments have also been made for using HCV antivirals as a means of preventing onward HCV transmission (ie, for treatment as prevention) [15]. Indeed, reducing the incidence of HCV infection in high-risk populations will likely be one of the greatest challenges to HCV elimination.

Sequencing and phylogenetic methods may provide a means for identifying clusters of individuals at the greatest risk of HCV transmission [16–18]. However, phylogenetic methods are often limited by a lack of standardized criteria for defining transmission clusters. An outbreak investigation tool, such as that developed by Campo et al, combined with a population-level sequencing approach, could be used to pinpoint populations at high risk of HCV transmission for targeted treatment, care, and harm reduction.

Caution is always required when applying sequencing methods to outbreak investigations, as these tools also have the potential to breach patient confidentiality and falsely implicate individuals in an outbreak. Investigating outbreaks requires supporting epidemiological evidence, engagement of appropriate public health officials, and careful consideration of how to engage implicated individuals.

The method presented by Campo et al is an important addition to our outbreak investigation tool kit. It highlights the value of using well-characterized data sets to validate new experimental methods. It is also an excellent example of how modern sequencing-based approaches can be used to support and improve public health.

Note

Potential conflict of interest. A. D. O. has previously collaborated with some of the authors of the editorialized manuscript, but the collaborations were unrelated to this manuscript, and there was no financial involvement in these collaborations. The author has submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. CDC. Epidemiologic notes and reports outbreak of hepatitis C associated with intravenous immunoglobulin administration—United States, October 1993-June 1994. MMWR Morb Mortal Wkly Rep **1994**; 43:505–9.
2. Sodeyama T, Kiyosawa K, Urushihara A, et al. Detection of hepatitis C virus markers and hepatitis C virus genomic RNA after needlestick accidents. Arch Intern Med **1993**; 153:1565–72.
3. Suzuki K, Mizokami M, Lau JY, et al. Confirmation of hepatitis C virus transmission through needlestick accidents by molecular evolutionary analysis. J Infect Dis **1994**; 170:1575–8.
4. Lanini S, Abbate I, Puro V, et al. Molecular epidemiology of a hepatitis C virus epidemic in a haemodialysis unit: outbreak investigation and infection outcome. BMC Infect Dis **2010**; 10:257.
5. Shemer-Avni Y, Cohen M, Keren-Naus A, et al. Iatrogenic transmission of hepatitis C virus (HCV) by an anesthesiologist: comparative molecular analysis of the HCV-E1 and HCV-E2 hypervariable regions. Clin Infect Dis **2007**; 45:e32–8.
6. Hmaied F, Mamou M, Dubois M, et al. Determining the source of nosocomial transmission in hemodialysis units in Tunisia by sequencing NS5B and E2 sequences of HCV. J Med Virol **2007**; 79:1089–94.
7. Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. Microbiol Mol Biol Rev **2012**; 76:159–216.
8. Saito T, Watanabe H, Shao L, et al. Transmission of hepatitis C virus quasispecies between human adults. Hepatol Res **2004**; 30:57–62.
9. Liu C-H, Chen B-F, Chen S-C, Lai M-Y, Kao J-H, Chen D-S. Selective transmission of hepatitis C virus quasi species through a needlestick accident in acute resolving hepatitis. Clin Infect Dis **2006**; 42:1254–9.

10. Bruguera M, Saiz JC, Franco S, et al. Outbreak of nosocomial hepatitis c virus infection resolved by genetic analysis of HCV RNA. J Clin Microbiol **2002**; 40:4363–6.
11. Ross RS, Viazov S, Gross T, Hofmann F, Seipp H-M, Roggendorf M. Transmission of hepatitis C virus from a patient to an anesthesiology assistant to five patients. N Engl J Med **2000**; 343:1851–4.
12. Lapointe H, Dong W, Lee GQ, et al. HIV drug resistance testing by high-multiplex "wide" sequencing on the Illumina MiSeq. Antimicrob Agents Chemother **2015**; 59:6824–33.
13. Quiñones-Mateu ME, Avila S, Reyes-Teran G, Martinez MA. Deep sequencing: becoming a critical tool in clinical virology. J Clin Virol **2014**; 61:9–19.
14. Hellard M, Doyle JS, Sacks-Davis R, Thompson AJ, McBryde E. Eradication of hepatitis C infection: the importance of targeting people who inject drugs. Hepatology **2014**; 59:366–9.
15. Martin NK, Vickerman P, Grebely J, et al. Hepatitis C virus treatment for prevention among people who inject drugs: modeling treatment scale-up in the age of direct-acting antivirals. Hepatology **2013**; 58: 1598–609.
16. Rolls DA, Sacks-Davis R, Jenkinson R, et al. Hepatitis C transmission and treatment in contact networks of people who inject drugs. PLoS One **2013**; 8:e78286.
17. Olmstead AD, Joy JB, Montoya V, et al. A molecular phylogenetics-based approach for identifying recent hepatitis C virus transmission events. Infect Genet Evol **2015**; 33:101–9.
18. CDC. Use of enhanced surveillance for hepatitis C virus infection to detect a cluster among young injection-drug users—New York, November 2004–April 2007. MMWR Morb Mortal Wkly Rep **2008**; 57:517–21.