Identifying and characterizing Hepatitis C virus mixed infections using a novel deep sequencing-based approach

Andrea D. Olmstead1, Vincent Montoya1, Winnie Dong1, Celia K. Chui1, Jeffrey B. Joy1,2, Thuy Nguyen1, Weiyan Dong1, Vera Tai1, Art F. Y. Poon1, Chanson J. Brumme1, Marianne Martinello6, Gail V. Matthews5, Gregory J. Dore6, Tanya L. Applegate4, F. Richard Harrigan4, Jason Grebely2, Anita Y.M. Howe1
1BC Centre for Excellence in HIV/AIDS, Vancouver BC, Canada, 2University of British Columbia, Vancouver, BC, Canada, 3The Kirby Institute, UNSW Australia, Sydney NSW, Australia

Background
Hepatitis C virus (HCV) infection does not elicit a complete protective immune response; thus, individuals can be infected with multiple genotypes or subtypes. Mixed HCV infections have been observed in 0% to 39% of high-risk populations2 and pose a challenge to genotype-specific therapies and vaccine design. More reliable and standardized methods are needed to increase our understanding of the prevalence and impact of mixed HCV infections.

Research Goals
• Optimize deep sequencing strategy to detect mixed infections and accurately determine relative genotype proportions
• Develop strict cutoff criteria to eliminate false positive results
• Apply methodology to high-risk HCV infected cohorts

Samples
• Artificially mixed samples: Mixed, de-identified clinical samples of known HCV genotypes
• ACTIVATE: International trial of response-guided PegIFN/RBV therapy in people with HCV G2/3 (96% history of IDU, 73% within 6 months and 59% within one month of enrollment; no HIV)
• Dare-C II: Australian trial of 6 week sofosbuvir and ribavirin treatment in acute/recent HCV infection (84% history of IDU, 58% within 6 months of screening; 74% HIV)
• Canadian Co-infection Cohort (CCC): HIV/HCV co-infected individuals in Canada (mixed infection suspected)

Table 1 - Artificially mixed genotype samples used for testing and validation of mixed genotype methodology

<table>
<thead>
<tr>
<th>Genotype Mixes</th>
<th>Ratios</th>
<th>Sample Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a:1b, 1a:2b, 1a:3, 1a:6</td>
<td>2:1, 4:3, 4:6, 5:6, 2:3</td>
<td>10:0:10, 10:90, 98:2</td>
</tr>
</tbody>
</table>

Table 2 - Clinical samples used for application of mixed-genotype methodology

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Sample Number</th>
<th>Number of Individuals</th>
<th>Genotypes (Sample Number)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTIVATE</td>
<td>17</td>
<td>90</td>
<td>1a (100), 2a (2); 2b (1)</td>
</tr>
<tr>
<td>Dare-C II</td>
<td>20</td>
<td>18</td>
<td>1a (15), 2a (2); 2b (2); 1a:1b rec (1); 1a:1b mix (1)</td>
</tr>
<tr>
<td>CCC</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

NSSB Amplicon sequencing preferentially amplified G1b, G2, G3 and G5 in artificially mixed samples. No genotype bias was observed with Random Primer and Capture Probe sequencing (Figure 2).

Based on ROC Curve Analysis (Figure 4), mixed infection HCV genotypes should have a minimum percent depth of 1.8% (out of total HCV depth), minimum percent coverage of 98.2% and an average depth of 39.2 reads per position.

NSSB PCR
Nucleic Acid Extraction
Random primer
Probe capture

Random primer
Primers

Probe capture
Hybridization libraries to HCV NS5B, NS5A and NS5B capture probes (fluorescent)
Capture probes with beads
Amplify with low cycle PCR

Figure 2. Mean proportion of expected and observed HCV reads for various genotypes in artificially mixed samples sequenced following N55B PCR, randomly primed cDNA synthesis and probe capture.

Figure 3. Random primer and capture probe sequencing results in artificially mixed HCV samples. A) HCV reads/ sample and HCV read depth; B) % HCV reads and % coverage.

Figure 4. Receiver Operating Characteristic (ROC) curve analysis of 64 artificially mixed genotype HCV samples. AUC = Area Under the Curve.

Figure 5. Random primer and probe capture sequencing of ACTIVATE and Dare-C II samples. A) HCV reads vs viral load; B) coverage and depth.

On average, 2% and 1% of HCV reads mapped to a secondary genotype for Random Primer and Probe Capture sequencing, respectively.

None of the secondary genotypes in ACTIVATE or Dare-C II samples passed the mixed infection cutoff criteria.

Conclusions
• Template independent sequencing methods more accurately capture mixed infection proportions compared to PCR
• Establishing cutoff criteria can rule out false assumptions of mixed genotype infections
• Mixed infections may not be as prevalent in high-risk populations as previously assumed, thus future studies warrant use of methods with standardized cutoff criteria

References