Next generation sequencing of the hepatitis C virus NS5B gene reveals potential novel S282 drug resistance mutations

Hezhao Ji a,1, Robert A. Kozak b,1, Mia J. Biondi c, Richard Pilon a, Dominic Vallee a, Ben Binhua Liang d, David La d, John Kim a, Gary Van Domselaar d, Lynne Leonard e, Paul Sandstrom a, James Brooks a,⁎

a National HIV & Retrovirology Laboratories, National Microbiology Laboratory, Public Health Agency of Canada, Ottawa, Canada
b Department of Pathobiology, University of Guelph, Guelph, Canada
c Arthur Labatt Family School of Nursing, Western University, London, Canada
d Bioinformatics Core, National Microbiology Laboratory, Public Health Agency of Canada, Ottawa, Canada
e Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada

ARTICLE INFO
Article history:
Received 19 August 2014
Returned to author for revisions
11 November 2014
Accepted 23 December 2014
Available online 17 January 2015

Keywords:
HCV
NS5B
Pyrosequencing
Drug resistance mutation
In silico molecular modeling

ABSTRACT

Identifying HCV drug resistance mutations (DRMs) is increasingly important as new direct acting antiviral therapies (DAA) become available. Tagged pooled pyrosequencing (TPP) was originally developed as cost-effective approach for detecting low abundance HIV DRMs. Using 127 HCV-positive samples from a Canadian injection drug user cohort, we demonstrated the suitability and efficiency of TPP for evaluating DRMs in HCV NS5B gene. At a mutation identification threshold of 1%, no nucleoside inhibitor DRMs were detected among these DAA naïve subjects. Clinical NS5B resistance to non-nucleoside inhibitors and interferon/ribavirin was predicted to be low within this cohort. S282T mutation, the primary mutation selected by sofosbuvir in vitro, was not identified while S282G/C/R variants were detected in 9 subjects. Further characterization on these new S282 variants using in silico molecular modeling implied their potential association with resistance. Combining TPP with in silico analysis detects NS5B polymorphisms that may explain differences in treatment outcomes.

Crown Copyright © 2014 Published by Elsevier Inc. All rights reserved.
not collected. The average duration of intravenous drug use was
were co-infected with HIV-1. Data on previous HCV treatment was
duals were infected with genotype 1a (GT1a), although genotypes
cohort studied are summarized in Table 1. The majority of indivi-
HCV genotypes (Pilon et al., 2011). Estimated cost for genotyping for
derived reads aligned to reference sequences of the pre-determined
Cohort characteristics
Results
Cohort characteristics
All 127 specimens were successfully TPP sequenced and the
derived reads aligned to reference sequences of the pre-determined
HCV genotypes (Pilon et al., 2011). Estimated cost for genotyping for
each specimen was $45 (Ji et al., 2013b). The characteristics of the
cohort studied are summarized in Table 1. The majority of individ-
uals were infected with genotype 1a (GT1a), although genotypes
1b, 2a, 2b and 3a were also found. Twenty-one (16.5%) subjects
were co-infected with HIV-1. Data on previous HCV treatment was
not collected. The average duration of intravenous drug use was
more than two decades.

TPP performed with high coverage and low error rate
The average TPP read redundancy was 3120 for nucleotides 7603–
9378 (AA 220–345) in the NS5B gene fragment. The cumulative error
rate from PCR and the pyrosequencing process, as determined by TPP
results from the control plasmid, was 0.0062 changes/nucleotide
position. Not surprisingly, higher error rates were observed in the
homopolymeric nucleotide runs and their flanking loci (data not
shown). However, none of the detected NS5B variants were affected
by homopolymeric errors. Error rates at variant positions were
significantly lower (0.0013 changes/nucleotide position) than those
found at other positions. Based on these observations, a conservative
frequency cutoff of 1% was set for low abundance variant identification
in this study.

RBV/IFN resistance mutations identified by TPP
Although the treatment status of our cohort was unknown, the
proportion of inner city residents receiving HCV treatment in Canada
has historically been very low (Grebel et al., 2009). Inferring that
> 95% of this cohort would be treatment naïve; we sought to use TPP
to evaluate the prevalence of polymorphism that affect outcomes
with RBV/IFN treatment. While no single mutation resistance has
been associated with resistance to interferon (Kuntzen et al., 2008),
12 RBV/IFN associated mutations were detected by TPP at frequencies
ranging from 1.2% to 100% with significant variation between
genotypes (Table 2). For example, D310N (Asahina et al., 2005)
detected in 26/28 (92.9%) of GT3a and 22/84 (26.2%) of GT1a, but was
not observed in isolates from any other genotype. In another
instance, T329I was detected in 4/28 (14.3%) of GT3a and 1/4 (25%)
of GT1b, while not observed elsewhere. K254R was reported pre-
viously to be co-selected by RBV with mutations such as D375G,
F415Y, and Q544R (Young et al., 2003). In our analysis the K254R
phenotype was detected in 11.8% of the studied subjects, including all
group GT1b isolates and 11/84 (13.1%) of GT1a (Table 2). We also
observed mutations correlated with a sustained virologic response
(Asahina et al., 2005; Hamano et al., 2005). For example, D244N was
observed in 25/28 (89.3%) of GT3a subjects, while A333E was only
detected in all GT2a and GT2b subjects. Q309R was observed in 27/28
(96.4%) of GT3a and 29/84 (34.5%) of GT1a, while not found in the
other three genotypes. S326G was identified in only 14/84 (16.7%)
of GT1a and 1/28 (3.6%) of GT3a. These data demonstrate that there is
a high prevalence of, and significant inter-genotypic variation in,
polymorphisms associated with differences in RBV/IFN treatment
outcomes. No significant difference in the frequency of HCV poly-
morphisms, between the HIV-HCV co-infected and the HCV infected
subjects, was observed.

TPP did not detect NNI resistance mutations in AA 220–345 of NS5B
It has been predicted that mutations in the allosteric binding
pockets targeted by NNIs will be found more frequently, as they
have less of an effect on replicative fitness (McCown et al., 2008).
Amino acid C316 is within the allosteric binding pocket of the
polymerase and is of particular interest, as mutations at this position
confer resistance to two distinct classes of NNIs, the benzimida-
zoles and the benzofurans (Mayhoub, 2012); and more
recently has been found in patients who have failed treatment with
sofosbuvir (Donaldson et al., 2015). However, consistent with the
findings of Kuntzen et al., who analyzed the same region with
Sanger sequencing, we did not detect minority NNI mutations
C316N/Y/S, which are classically associated with resistance to NNIs
such as HCV-796 and ABT-333 (Alota et al., 2012; Howe et al., 2008;
Kuntzen et al., 2008; Poordad et al., 2014). In contrast, others have
reported mutations at this position ranging in frequency from 0.19%
to 91.6% (Alves et al., 2013; Bartels et al., 2013; Jaspe et al., 2012;
Kuntzen et al., 2008); however these discrepancies may be due to
the genotypic structure of the circulating viruses within particular
geographical locations.

Table 1

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>107</td>
</tr>
<tr>
<td>Female</td>
<td>20</td>
</tr>
<tr>
<td>Age (years)</td>
<td>43 ± 8.9</td>
</tr>
<tr>
<td>Duration of drug use (years)</td>
<td>22.0 ± 10.6</td>
</tr>
<tr>
<td>HCV genotypes</td>
<td></td>
</tr>
<tr>
<td>1a</td>
<td>66.1% (N=84)</td>
</tr>
<tr>
<td>1b</td>
<td>3.1% (N=4)</td>
</tr>
<tr>
<td>2a</td>
<td>1.6% (N=2)</td>
</tr>
<tr>
<td>2b</td>
<td>7.1% (N=9)</td>
</tr>
<tr>
<td>3a</td>
<td>22.0% (N=28)</td>
</tr>
</tbody>
</table>
Variants detected at codon S282

The highly sensitive TPP platform enables detection and quantification of HCV LDRVs present at a frequency of 1% or greater. Even with this technique, S282T, the primary mutation associated with resistance to sofosbuvir and other HCV NIs (Aloia et al., 2012), was not detected in any specimens. However, several other S282 variants were detected at differing frequencies. For example, S282R and S282C were each detected in one of the GT1a subjects at 4.42% and 4.87%, respectively. S282G was detected in one GT3a patient at a frequency of 7.77%, and an additional six GT1a samples at frequencies ranging from 1.15% to 1.82%. With the majority of S282G/R/C variants identified in GT1a subjects, we also examined the clade identities of the those 7 patients positive for S282G/R/C (Pickett et al., 2011). Phylogenetic analysis using both the Sanger sequencing data and the TPP consensus sequences (TPP20) showed that both 1a1 and 1a2 clades were present in this group (1a1:1a2 = 4:3). There was no specific pattern of association between S282 variation and specific 1a clades. In addition,

Table 2
Inter-genotype difference of HCV variants detected by TPP (frequency > 1%).

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Percentage of mutant positive subjects in examined genotype groups (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>239V</td>
</tr>
<tr>
<td>1a (N=84)</td>
<td>7.1</td>
</tr>
<tr>
<td>1b (N=4)</td>
<td>—</td>
</tr>
<tr>
<td>2a (N=2)</td>
<td>—</td>
</tr>
<tr>
<td>2b (N=9)</td>
<td>11.1</td>
</tr>
<tr>
<td>3a (N=28)</td>
<td>67.9</td>
</tr>
</tbody>
</table>

—: not detected.

Fig. 1. Workflow of TPP based HCV NS5B genotyping.
In silico modeling identifies potential novel drug resistance mutations

We identified novel mutations S282C/G/R for which no cell-culture or biochemical NS5B resistance data has been published. Our data demonstrates that these mutations can be found in patients, which is in agreement with the similar finding of Le Pogam et al. (2010). We sought to further investigate their potential resistance profile utilizing a freely available program, I-TASSER, which can model a protein sequence onto a specified template. We submitted a genotype 1a sequence to be modeled on the NS5B template 1NB6. The 1NB6 template offers the advantage that its structure contains the sofosbuvir analog UTP at the active site. This structure contains divalent metal ions, and therefore demonstrates how the incoming UTP may access the active site, forming a ternary complex similar to that, which is seen with HIV-1 reverse transcriptase (O’Farrell et al., 2003). We compare these results to those obtained with the 4E7A structure, which demonstrates NS5B in a pretranslocated state, with a short primer-template and a terminal 2’,3’-ddC (Mosley et al., 2012). Although both structures are considered less than ideal (Gotte, 2014; Karam et al., 2014), in the absence of a high-resolution genotype 1a ternary structure chain-terminated with UTP, we felt that the selection of these structures as templates was appropriate. Having structures demonstrating a spectrum of potential conformational changes may even be preferred considering the elusive nature of how substitutions at S282 confer decreased sensitivity to sofosbuvir.

Previous studies have demonstrated that S282T causes 9.5-fold resistance to sofosbuvir (Lam et al., 2012). Thus it was not surprising that although both serine and threonine have polar, uncharged side-chains, the addition of a methyl group within threonine (Fig. 2A) changes the orientation of the side-chain such that the impact of the hydroxyl may change the orientation of the uracil. When examining the S282R structure, our model demonstrates an obvious interaction between the positively charged arginine and the hydroxyl group at position 2 of the ribose, which (because of the flexibility of the inhibitor) could change the orientation of the specificity domain required to chain terminate (Fig. 2B). With respect to positioning, the cysteine of S282C (Fig. 2C) may have less of an effect than the threonine, although it is difficult to predict the contribution of a thiol group in place of a hydroxyl group. Finally, as glycine is obviously smaller than all previous side chains, we propose that S282G (Fig. 2D) may have more serious effects on the resistance profile of sofosbuvir. This would likely be due to the change in polarity that would occur in this region, and the potential effects this change may have on the orientation of the base of the inhibitor. These results are in agreement with those obtained using the 4E7A structure as a template (Supplemental Fig. 2).

We also sought to investigate the potential synergistic impact of a newly identified mutation, L159F (Donaldson et al., 2015; Tong et al., 2014). L159F emerged in 6 genotype 3a subjects on sofosbuvir therapy by deep-sequencing methods. It was also present at baseline in four genotype 1b patients, who failed therapy; and of the on-treatment failures, a 1b patient had L159F at baseline, and it emerged in a 2a and two 1a patients. Depending on the patient and trial, the mutation was classified as either a resistance-associated substitution as determined by the FDA Division of Antiviral Products criteria, or resistance-associated polymorphism present at baseline in subjects who failed treatment (Donaldson et al., 2015). Our modeling data demonstrates that when the leucine at 159 is replaced with a phenylalanine, the orientation of the residue does not significantly change, except in the case of S282R/L159F (Fig. 3). In fact based on our data, for both S282T/ L159F (Fig. 3A) and S282C/L159F (Fig. 3C), the substitution at L159F may lessen any potential interactions between sofosbuvir and a 282 variant, suggesting it may act as a compensatory mutation if found. However, in the case S282R/L159F our model suggests that a steric clash would occur between the arginine and the uracil, likely as a result of the effect of aromatic side-chain of phenylalanine on the charged arginine residue (Fig. 3B). In summary, based on our modeling data (irrespective of the template used), and other deep-sequencing

Fig. 2. Modeling of active site S282 variants in NS5B genotype 1a. Overlay of UTP-bound (orange), 1NB6 (O’Farrell et al., 2003), and an I-TASSER generated genotype 1a HCV NS5B containing (A) S282T, (B) S282R, (C) S282C, and (D) S282G. Mutations generated in I-TASSER are shown in green. The catalytic triad coordinates the two manganese ions shown in purple.
studies (Donaldson et al., 2015) we would prioritize further investigating S282G and S282R as single-mutants, and S282R/L159F with respect to a potential double-mutant.

Minor HCV NI DRMs were detected at different frequencies among genotypes

V321I, known to enhance resistance to NIs such as PSI-352938 and PSI-353661, was observed in three of 84 GT1a patients, with frequencies at 5.89%, 1.27% and 1.17% respectively. This is in agreement with Margeridon-Thermet et al. (2014) who found V321I in one GT1a sample (Margeridon-Thermet et al., 2014). It has also recently been shown that V321A may be associated with treatment failure in sofosbuvir-experienced patients (Donaldson et al., 2015). However, none of the other associated variants that may synergize with V321I such as R222Q, C223Y/H and L320I, were observed (Lam et al., 2011). It has been shown that L320F/C, A300T, and I239L may be co-selected with S282T in vitro using PSI-6130 and R7128 (Ali et al., 2008; Tong et al., 2014). In the case of L320F, the mutation alone confers a 2.7-fold increase in in vitro resistance and a 41-fold increase when present in conjunction with S282T in vitro (Tong et al., 2014). Our results demonstrate that the A300 site is highly polymorphic and that the phenotype at this site is highly genotype-specific. T300 is the dominant phenotype in all GT1b and GT3a subjects, but the majority of GT1a subjects have Q300, and all GT2a and 2b showed L300. Similarly, the I239 site is also polymorphic with significant inter-genotype variations. I239L was observed in only one GT1a subject at frequency of 2.10%, while I239V variants were detected in 26 subjects with HCV GT3a, GT1b or GT1a (Table 2). The 320F/C phenotypes were not detected in any of the studied subjects.

High concordance was observed between TPP and Sanger sequencing

Results from SS and TPP20 consensus sequences were highly concordant with the prevalence of mutations within the cohort being nearly identical (Table 3). The only exceptions were mutations at D244N, K254R and Q309R, which were each identified by SS; however the measured TPP mutation frequencies were 16.2%, 15.6% and 12.6% respectively. Thus, if the threshold for TPP reporting was set at 20%, these mutations would not have been identified; however, this did not significantly affect the prevalence of HCV DR within the cohort.

Discussion

Surveillance of populations to identify the baseline prevalence of HCV DRMs, such as S282T, is important for providing data upon which to make therapeutic and policy decisions. Although deep sequencing has been utilized to study NS5B regions, these studies either failed to look for DRMs; used small cohorts of HCV/HIV co-infected individuals; or were limited by HCV subtype homogeneity (Franco et al., 2013; Maimone et al., 2014; Margeridon-Thermet et al., 2014; Wu et al., 2013). Recent work by Tong et al. (2014) identified

<table>
<thead>
<tr>
<th>Mutant phenotype</th>
<th>TPP1</th>
<th>TPP5</th>
<th>TPP10</th>
<th>TPP20</th>
<th>Sanger sequencing</th>
</tr>
</thead>
<tbody>
<tr>
<td>I239V</td>
<td>20.5</td>
<td>7.1</td>
<td>1.6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>I239L</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D244N</td>
<td>19.7</td>
<td>19.7</td>
<td>19.7</td>
<td>18.9</td>
<td>19.7</td>
</tr>
<tr>
<td>K254R</td>
<td>11.8</td>
<td>7.9</td>
<td>7.1</td>
<td>6.3</td>
<td>7.1</td>
</tr>
<tr>
<td>A300T</td>
<td>25.1</td>
<td>25.1</td>
<td>25.1</td>
<td>25.1</td>
<td>25.1</td>
</tr>
<tr>
<td>Q309R</td>
<td>51.2</td>
<td>44.1</td>
<td>44.1</td>
<td>43.3</td>
<td>44.1</td>
</tr>
<tr>
<td>Q309K</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>D310N</td>
<td>37.8</td>
<td>37.8</td>
<td>37.8</td>
<td>37.8</td>
<td>37.8</td>
</tr>
<tr>
<td>V321I</td>
<td>2.4</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>S282G</td>
<td>11.8</td>
<td>1.6</td>
<td>0.8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>T229I</td>
<td>3.9</td>
<td>2.4</td>
<td>2.4</td>
<td>1.6</td>
<td>1.6</td>
</tr>
<tr>
<td>A333E</td>
<td>9.4</td>
<td>8.7</td>
<td>8.7</td>
<td>8.7</td>
<td>8.7</td>
</tr>
</tbody>
</table>

| Fig. 3. Modeling of active site S282 variants with L159F in NS5B genotype 1a. Overlay of UTP-bound (orange), 1NB6 (O’Farrell et al., 2003), and an I-TASSER generated genotype 1a HCV NS5B containing (A) S282T/L159F, (B) S282R/L159F, (C) S282C/L159F, and (D) S282G/L159F. Mutations generated in I-TASSER are shown in green. The catalytic triad coordinates the two manganese ions shown in purple.

Table 3 Average prevalence (%) of HCV variants in the cohort while varied frequency cutoffs were applied.
mutations that may play a role in resistance to NS5B inhibitors, and highlighted the need to identify potential novel DRMs that require surveillance. Additionally, Donaldson et al. (2015) recently identified several novel mutations observed in patients who failed sofosbuvir-containing therapy (Donaldson et al., 2015).

The S282T mutation is associated with relapse in sofosbuvir monotherapy treatment (Gane et al., 2013; Svarovskaia et al., 2014), however, as the baseline genotype was not reported, whether the S282T mutation existed prior to treatment remains undetermined. While a recent small study found a single patient with S282T following sofosbuvir plus ribavirin therapy (Osinski et al., 2014), larger trials involving patients with HCV genotypes 1, 2 and 3 failed to find the S282T mutation among treatment-experienced and naïve individuals (Gane et al., 2013; Jacobson et al., 2013; Lawitz et al., 2013a, 2013b; Applegate et al., 2014; Donaldson et al., 2015; Maimone et al., 2014; Margeridon-Thermet et al., 2014; Svarovskaia et al., 2014). Our data is in agreement with these large studies in that we did not find evidence of S282T through either SS or at lower frequencies or by using TPP sequencing. It has been shown in cell culture that S282T results in decreased viral fitness (Ludmerer et al., 2005) and our inability to detect the mutation in our cohort is consistent with this mutation being associated with a high fitness cost.

Using TPP we were able to detect minor variant species at the S282 locus, including S282G, S282R and S282C. Although in vitro studies often indicate reduced viral fitness of drug resistance mutations, Kuntzen et al. demonstrated viral fitness may be unimpaired in individual situations, perhaps as a result of compensatory mutations (Kuntzen et al., 2008). Reduced or weak enzymatic activity may also occur during biochemical analysis, as has been shown for S282R. This decrease in fitness was also the case when Le Pogam et al. (2010) detected S282G and S282R through clonal sequencing among patients on monotherapy with the nucleoside inhibitor RG7128. When assessing the fitness of these variants this group was unable to obtain replication competent-viruses when introducing these mutations through site-directed mutagenesis. They were, however, able to propagate isolated clones (Le Pogam et al., 2010). As these mutations have now been found by our group as well as others (including under drug selective pressure) (Donaldson et al., 2015; Franco et al. 2013; Maimone et al., 2014), we are confident that viruses containing S282R or S282G are fit enough to survive, and therefore warrant further investigation. We find it especially interesting, that Donaldson et al. (2015) detected S282R in a patient who failed a sofosbuvir-containing regimen. As S282C and S282G require only a single transition to from S282T, the importance of these variants may also include the potential to be springboard mutations (Powdrill et al., 2011). Selection experiments with Nls PSI-6130 and R7128 suggested that other mutations, including I239L/T, L320F/L/C and A300T, might be co-selected and synergize with S282T by enhancing resistance to the aforementioned compounds (Ali et al., 2008; Tong et al., 2014). Our data indicates that, in the absence of S282T, low frequency I239L is detected in only 1/127 subjects and there was no evidence of 320F/C in the cohort. These findings support the hypothesis that both mutations are co-selected with S282T during treatment with Nls (Ali et al., 2008). However, T300 was observed as the dominant phenotype in HCV GT1b and GT3a subjects, while the majority of GT1a subjects have Q300, and all GT2a and 2b showed L300. Therefore, we believe that A300T represents primarily an inter-genotype polymorphism rather than positively co-selected drug resistance variation. Nevertheless, further investigation is required to determine the exact role of these variants in viral fitness and drug resistance, especially when more compounds of this class are made available in the clinic. Our TPP data indicated that many of the patients within the cohort had at least one variation, or polymorphism, known to confer resistance to RBV/IFN combination (Asahina et al., 2005; Hamano et al., 2005; Jaspe et al., 2012). These findings may help to explain part of the variability observed in treatment outcomes using IFN based therapies. We employed TASSER as a predictive modeling method in order to better understand the mechanism of resistance of S282 polymorphisms, using two well-defined crystal structure templates. Despite differences in the genotype, deletions, and overall conformation (open versus closed) (Mosley et al., 2012; O’Farrell et al., 2003), the interaction between variants at S282 and the UTP (or terminal nucleotide) remains consistent, and allows for the prioritization of substitutions for further biochemical study. Residues in close proximity to S282 may help to coordinate UTP and, by corollary, sofosbuvir. The spatial and charge-related amino acid changes observed in the S282G/R/C variants may affect sofosbuvir binding in a similar manner to S282T. However, the specificity domain of sofosbuvir may also be able to easily reposition resulting in no change in the overall affinity of inhibitor binding. In order to definitively predict these types of outcomes expensive dynamic modeling software and high-processing computers would be required. In our study, generating static models of drug resistance mutations is a cost-effective, relatively simple first step that may help predict sites of drug resistance that are oriented around polymorphisms. This later point is more significant in that the S282G mutation was found among several sequences suggesting that it is stable and thus may be transmissible. From a clinical perspective, detecting this polymorphism as a minority variant previously undetectable by conventional methods, may provide more insight into cases of clinical failure. As the clinical use of sofosbuvir increases, the collective in vivo selective pressure will correspondingly increase and thus allow us to validate the drug resistance predications by using this technique.

In conclusion, we have successfully demonstrated that TPP is a cost-effective (Ji et al., 2010, 2013b), advanced NGS platform that can be used to evaluate HCV drug resistance. We have demonstrated that TPP allows population/cohort level HCV DR surveillance analysis by generating consensus sequences comparable to standard methods. The role of minority variants detected using TPP in treatment outcome remains to be determined. Through pooling multiple specimens in a single pyrosequencing run, HCV genotyping can be performed more efficiently and at cost significantly lower than conventional methods (Ji et al., 2010, 2013b). Recognizing potential limitations from pyrosequencing method, we have confirmed in a separate study that the TPP approach described here is fully adaptable for other NGS technologies, such as Illumina MiSeq (data not shown). Combining pooled NGS with molecular modeling provides an approach to explore potentially drug resistant polymorphisms that may manifest as clinical failure under drug pressure.

Materials and methods

Subjects and specimens

A total of 127 specimens were selected, from a previously described Ottawa cohort of IDUs based upon specimen availability and single HCV genotype (GT) infection (Pilon et al., 2011). No information was available on previous treatment history or HCV viral loads. This study was approved by the institutional Research Ethics Board.

HCV NS5B sequencing with TPP approach

Fig. 1 outlines the TPP workflow for HCV NS5B sequencing (Ji et al., 2013b). In brief, HCV RNA was extracted from dried blood spot (DBS) specimens using the Nuclisens EasyMag system (Boom et al., 1990). The HCV NS5B gene fragment encompassing the palm and partial finger domains was reverse transcribed and amplified...
using Qiagen One step RT-PCR kit (Love et al., 2003; Pilon et al., 2011). Following RT-PCR, one round of nested-PCR was performed using customized fusion primers, employing the 454 Lib-L template while the reverse primers tagged with patient-specific MIDs (Ji et al., 2013b). The template-specific primer sequences for the fusion primers were: NS5B-in-for: IACYATCATGCAAARRYAGGT; and NS5B-in-rev: ACCTRGTATAGCTCCGTGAA (Pilon et al., 2011). The resulting ~600 bp amplicon spanned the majority of relevant mutation loci related to sofosbuvir and RBV/IFN treatments. The amplicons were purified, quantified, and then pooled at equimolar concentrations with those from other subjects (Ji et al., 2013b). The pooled templates were then clonally amplified using Roche emulsion PCR (Lib-L) kit, and loaded onto picotiterplate fitted with an 8-lane gasket for unidirectional pyrosequencing with the GS FLX Titanium kit (Roche, USA).

Sequence analysis

All TPP reads satisfying the GS FLX default quality screening criteria were then decoded by patient-specific MIDs using an in-house Perl Script. Sequence contigs were then generated for each individual subject using specimen specific HCV GT references (Genbank accession numbers: 1a: NC004102; 1b: AJ238800; 2a: AB047639; 3a: AY515261) (Pilon et al., 2011). All reads were subjected to an in-house second round of quality screening with TPP reads considered valid only if they were ≥100 bp in length and mapped to the reference sequence with both 60% overlap and ≥75% identity. The overall empirical TPP error rate was assessed by PCR amplifying and pyrosequencing in parallel a pedigree H77 Delta E1-p7 replicon plasmid construct (a kind gift from Dr. Selena Sagan, Stanford University).

Based on the translated sequence alignments, all amino acid (AA) changes previously been associated with anti-HCV regimes or compounds were examined and quantified using BioEdit 7.0.9 (Hall, 1999). The inter-group comparisons among HCV genotypes were further assessed for all AA variants. The frequencies of TPP-based AA variants were also compared with Sanger sequencing (SS) results obtained previously (Pilon et al., 2011). To facilitate the comparison of the TPP with SS results, we reported a variant or DRM if it was present in ≥20% of the TPP reads (TPP20). To demonstrate the capacity of TPP to detect LADR, TPP consensus sequences with three lower variant frequencies (TPP1, TPP2, and TPP3) were also examined.

In silico modeling

In order to assess the potential impact of sofosbuvir activity of low frequency polymorphisms located within the NS5B active site, we developed in silico models using NS5B crystal structures containing uridine-5′-triphosphate (UTP). We submitted full-length genotype 1a sequences to the I-TASSER online server, which predicts protein structure and functions (Roy et al., 2010, 2011; Zhang, 2008). Rather than allowing the server to use the most related crystal structure in the database as a reference, we stipulated it use the HCV NS5B structure 1NB6 (O’Farrell et al., 2003) to identify potential interactions. As a structure with sofosbuvir bound has not yet been published, we felt this was an appropriate substitute as there is the presence of a terminal UTP, representing what would be the specificity domain of the inhibitor (corresponding to the base region of UTP) is in close proximity to the S282 position. This type of modeling is unique in that it does not require sophisticated software. We validated our chosen template structure by demonstrating that the positions of the active site residues are in agreement with other accepted crystal structures that have been used to model potential relationships between S282 and sofosbuvir (Supplemental Fig. 1) (Donaldson et al., 2015), as well as with the more recently published structure 4ETA (Supplemental Fig. 2) (Mosley et al., 2012). Tertiary alignments were carried out using USCF Chimera.

Acknowledgments

This study was funded through Ontario HIV Treatment Network, Canada (Grant no: ROGB171) and the Public Health Agency of Canada. The authors are grateful for the critical review and constructive comments from Dr. Curtis Cooper, the Director of the Ottawa Hospital and Regional Hepatitis Program, Canada.

Appendix A. Supplementary information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2014.12.037.

References


Howe, A.Y., Cheng, J., Johann, S., Mullen, S., Chunduru, S.K., Young, D.C., Bard, J., Chopra, R., Krishnamurthy, G., Mansour, T., O’Connell, J., 2008. Molecular mechan-


Lack of the NS5B S282T mutation in HCV isolates from liver tissue of naive infected patients. J. Virol. 77, 7575–7581.


