Ultra-deep Sequencing Reveals Dynamics of Drug Resistance-Associated Variants in Hepatitis C Viruses: Relevance to Treatment Outcome and Resistance Screening

Richard Barnard1, Abha Chopra2, Ian James3, Don Cooper2, John Blincoc4, Mark Watson2, Cassandra Jabara3, Stanley Lemon4, Simon Mallal2,5, Silvana Gaudieri2,6

1Antiviral Research, Merck and Co., West Point, PA, USA, 2Institute for Immunology and Infectious Diseases, Murdoch University, Western Australia, 3Department of Biology, & Linberger Comprehensive Cancer Center, University of North Carolina, USA, 4Division of Infectious Diseases, Department of Medicine, & Linberger Comprehensive Cancer Center, University of North Carolina, USA, 5Center for Translational Virology and Infectious Diseases, Department of Pathology, Microbiology and Immunology, Vanderbilt University, Nashville Tennessee, USA, 6School of Anatomy, Physiology and Human Biology, University of Western Australia

Background:
Hepatitis C virus (HCV) infection is a major global health issue with approximately 3% of the world’s population estimated to be infected with the hepatitis C virus (HCV). Inefficacies in treatment have led to development of direct-acting antivirals (DAAs) that specifically target HCV proteins involved in the virus’s life cycle. One of the major concerns arising from the use of the DAAs is the emergence of resistance-associated variants (RAVs) that affect the efficacy of the drugs. RAVs are generally associated with a fitness cost and the use of ultra-deep pyrosequencing technology has shown that in most treatment-naive subjects low frequency circulating strains carry RAVs.

Methods
The primer design was based on the Primer ID format as described in Jabara et al. 2011, in which a random eight nucleotide tag (‘barcode’) is inserted after the target sequence with the 5’ addition of a non-specific sequence to allow amplification by PCR (Figure 1). Ideally, each DNA template should be synthesized with a unique barcode tag, this was facilitated by keeping the starting viral copies (~20,000) much lower than the number of primer ID’s (65,536 IDs) allowing an accurate estimation of viral species in the sample pool and the elimination of PCR induced artefacts in the analysis pipeline (Figure 2).

Results

- PrimerID removes the effect of primer bias and different PCR efficiencies (Fig 4 and 5)
- 454-PrimerID technology improves the detection of clinically relevant low frequency RAVs compared to sanger sequencing (Table 1).
- Low frequency viral strains harboring RAVs can persist for up to two years post-treatment failure.
- Strains carrying multiple RAVs are common in breakthrough viruses.

Figure 1: Sense strand interaction with Primer ID. HCV target binding site (TBS), TAG stop subject tag, PrimerID of 45,536 unique primers,Genentech P65 2’3’ 17 by primer testing sites.

Figure 2: PCR amplification induced artefacts

Table 1: Comparison of RAV frequencies in longitudinal samples by population (Sanger) sequencing and 454-PrimerID.

- Evidence for mutation networks within NS3 protease and putative compensatory mutations (Fig 6)

Discussion
Despite the emergence of new promising drugs that specifically target HCV, the development of novel effective antiviral therapies for hepatitis C is still facing the challenging issue of the rapid selection of viral variants bearing drug resistance mutations. This study used 454-PrimerID to mitigate RT-PCR errors and issues surrounding re-sampling and pooling. The study shows the clinical relevance of RAVs below the detection threshold of sanger sequencing in clinical outcome. This data supports the previous observation that HCV strains with RAVs at lower than the level detectable by sanger-based sequencing can reduce drug sensitivity. Given the success of this technique in studying viral quasi-species before, during and after VF in these subjects, 454-PrimerID sequencing should become the standard approach by which to perform temporal 454-sequencing studies.

Conclusions

1) Ultra-deep sequencing identifies clinically relevant low frequency RAVs in circulating quasi-species at levels not detectable using Sanger sequencing.
2) PCR amplification induced artefacts are eliminated using the primer ID methodology and analysis pipeline.
3) Our findings show the clinical relevance of RAVs below the detection threshold of sanger sequencing in clinical outcome.
4) This data supports the previous observation that HCV strains with RAVs at lower than the level detectable by sanger-based sequencing can reduce drug sensitivity.
5) In the majority of subjects, RAVs were detected at high levels at VF but this did not predict which RAVs would be present after 1-2 years of follow-up.
6) The low frequency of naturally occurring drug-resistance mutants suggests that resistant variants are not as fit as wild type viruses. However, some strains that harbor RAVs can overcome that replication deficiency most likely via the support of compensatory mutations.

References

Figure 3: Sequencing Workflow

Figure 4: Correlation between variation observed from PrimerID analysis and our normal variant analysis without bar-coding (Y-axis), r² = 0.8975

Figure 5: Percentage difference between variant counts using ultra-deep PrimerID method and method without for synonymous and nonsynonymous sites

Acknowledgments: We would like to thank our co-authors for their ongoing assistance with developing this project, and all staff at IIDD for their assistance and input. This work was funded by Merck and Co.