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

Richard Barnard[1], Abha Chopra[2], Ian James[2], Don Cooper[2], John Blinco[2], Mark Watson[2], Cassandra Jabara[3], Stanley Lemon[4], Simon Mallal[2,5], Silvana Gaudieri[2,6]

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

Background:

Hepatitis C is a global health issue with approximately 3% of the worlds' population estimated to be infected with the hepatitis C virus (HCV) Inefficiencies in treatment has led to development of direct-acting antivirals (DAAs) that specifically target HCV proteins involved in the virus's life cycle[1]. 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 naïve subjects low frequency circulating strains carry RAVs[2].

The aim of the study was to investigate i) the clinical relevance of low frequency RAVs; ii) the persistence of RAVs and iii) compensatory mutations in a subset of subjects who had failed boceprevir (SCH 503034; protease inhibitor).

Subjects

Subjects were enrolled in the long-term follow-up study P-05063 of previous treatment with boceprevir. (ClinicalTrials.gov Identifier: NCT00689390). These subjects all had detectable RAVs at the virologic failure (VF) time-point based on Sanger sequencing. Twelve subjects had three time-points available: at baseline before the commencement of boceprevir treatment; at VF; and follow-up (median 830.5 days from VF).

Methods

The primer design was based on the Primer ID format as described in Jabara et al. 2011[3], 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 cDNA 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).

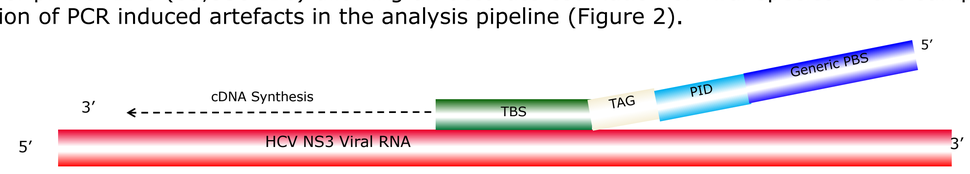

Figure 1 Reverse transcription with Primer ID. NS3 target binding Site (TBS), TAG 3bp Subject tag, PID=N8 → 65,536 unique primers, Generic PBS 2 x 17 bp primer binding sites

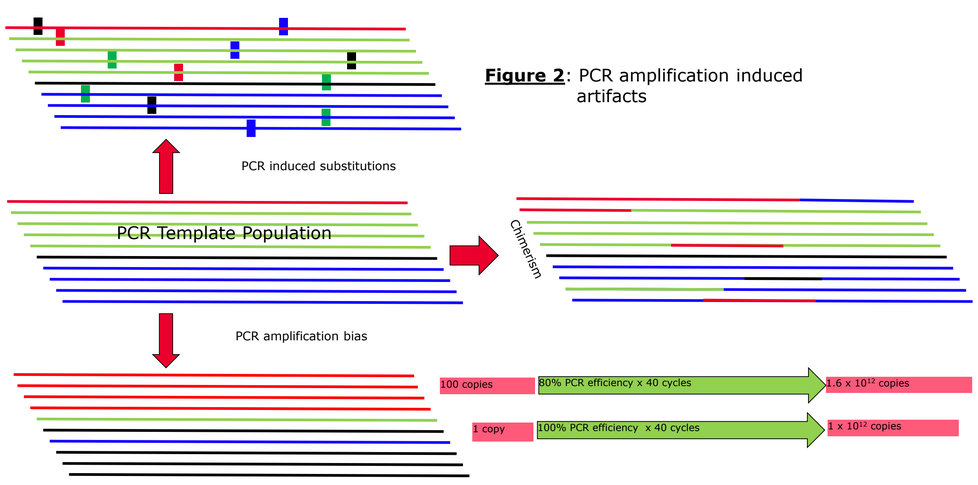

Figure 2: PCR amplification induced artifacts

Workflow

The cDNA template was amplified by nested PCR of the NS3 protease region. Amplicons were sequenced using both Sanger population sequencing and Next generation sequencing on the Roche 454 GS FLX+ Platform with XLR70 Titanium Chemistry (Figure 3).

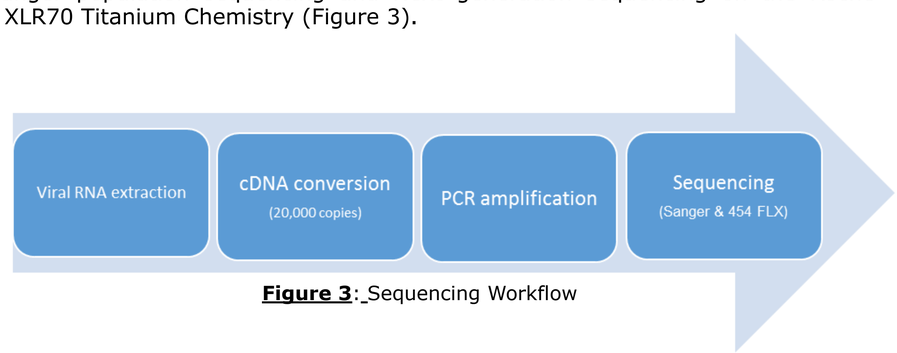

Figure 3: Sequencing Workflow

Eleven sites known to be associated with boceprevir drug resistance were screened: V36A/M, R155K/Q/T, Q41R, F43S, T54A/S, V55A, Q80K/R, S138T, A156S/T/V, V158I, D168A/T/V and V170A/T. Each site was studied in order to evaluate the frequency of RAVs within the viral quasispecies.

Results

- **PrimerID removes the effect of primer bias and different PCR efficiencies (Fig 4 and 5)**

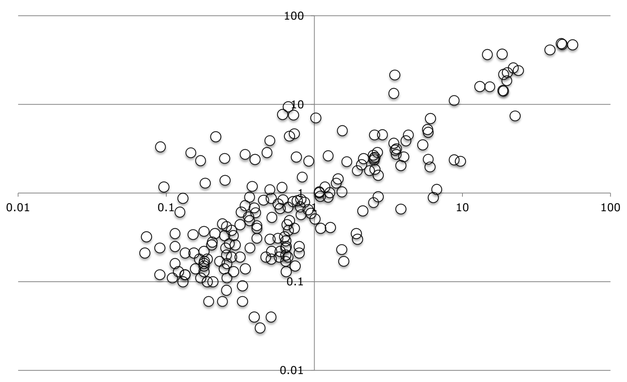

Figure 4: Correlation between variation observed from PrimerID analysis (X-axis) and normal variant analysis without bar-coding (Y-axis). $R^2 = 0.8975$

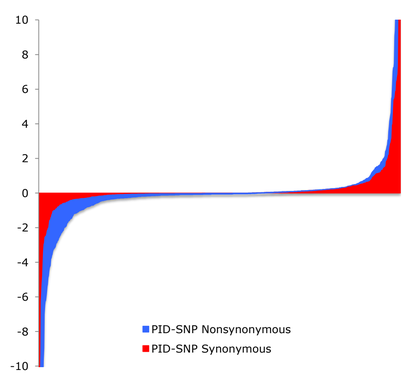

Figure 5: Percentage difference between variant detected by bar-coding PrimerID method and without for synonymous and nonsynonymous sites

- ***454-Primer ID 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*.**

			Baseline			Virologic Failure			Follow-up			
			Population sequencing	454-PrimerID Sequencing		Population sequencing	454-PrimerID Sequencing		Population sequencin	454-PrimerID Sequencing		
Subject #	RAVs detected by population sequencing	RAVs detected by 454/PrimerID sequencing		# Unique PID at RAV position	%RAV		# Unique PID at RAV position	%RAV		# Unique PID at RAV position	%RAV	Days since Virologic Failure
1	V36L	V36L	+	251	93.6	+	33	93.9	+	253	98.4	894
	V55A	V55A	+	253	93.3	+	34	91.2	+	269	77.3	
		V55G	-	253	0.0	-	34	0.0	-	269	21.2	
		R155K	-	253	0.0	-	34	8.8	-	272	0.0	
		V158I	-	253	0.0	-	34	94.1	-	272	0.0	
2	V36M	V36M	-	463	0.0	+	359	42.9	-	743	0.0	982
		V36L	-	463	0.0	-	359	1.1	-	743	0.0	
		V36A	-	463	0.0	-	359	11.7	-	743	0.0	
		T54A	-	463	0.0	-	388	2.6	-	746	0.0	
		T54S	-	463	0.0	-	388	10.6	-	746	0.0	
	V55A	V55A	-	464	0.0	+	388	1.5	-	746	0.0	
	R155K	R155K	-	464	0.0	+	391	53.5	-	746	0.0	
3	V36M	V36M	-	695	0.0	+	763	100	-	696	0.0	674
		V55A	-	695	0.0		705	0.0	-	1038	0.1	
	R155K	R155K	-	705	0.0	+	772	99.5	-	1052	0.0	
4		V36I	-	528	0.0	-	471	0.2	-	809	0.0	528
		V36M	-	528	0.0	-	471	19.7	-	809	0.0	
	V55A	V55A	+	526	97.9	+	484	98.6	+	809	99.9	
		V55P	-	526	1.9	-	484	0.0	-	809	0.0	
	R155K	R155K	-	528	0.0	+	484	97.1	+	809	79.9	
		D168N	-	528	0.0		484	0.2	+	809	79.9	
5	V36M	V36M	-	123	0.0	+	40	100	-	392	0.0	890
		T54I	-	149	0.0	-	49	0.0	-	392	0.3	
	R155K	R155K	-	156	0.0	+	51	98.0	-	392	0.3	
		R155T	-	156	0.0	-	51	2.0	-	392	0.0	
		D168N	-	156	0.0	-	52	1.9	-	392	0.0	
6		V36M	-	1392	0.0	-	377	5.8	-	555	1.1	555
		T54S	-	1369	0.0	-	389	5.4	-	556	0.0	
		V55A	-	1357	0.1	-	390	0.5	-	556	0.0	
		R155K	-	1398	0.0	-	389	14.7	-	556	0.0	
		R155M	-	1398	0.0	-	389	0.3	-	556	0.0	
	R155T	R155T	-	1398	0.0	+	389	81.0	-	556	0.2	
	D168N	D168N	-	1397	0.0	+	389	81.2	-	556	0.2	
7		V36L	-	533	0.0	-	1092	0.1	-	299	0.0	895
		V36M	-	533	0.0	-	1092	2.5	-	299	0.0	
	T54S	T54S	-	547	1.1	+	1101	97.6	+	305	99.7	
	R155K	R155K	-	550	1.3	+	1105	100	+	309	99.4	
		D168N	-	550	0.0	-	1105	0.1	-	309	0.0	
8		V36M	-	204	0.0	-	394	2.5	-	889	0.0	847
	R155K	R155K	-	213	0.0	+	395	100	-	908	0.0	
	V158I	V158I	-	213	0.0	+	395	97.5	-	908	0.0	
9	V36M	V36M	-	675	0.0	+	492	93.9	-	223	0.0	961
		V36L	-	675	0.0	-	492	0.2	-	223	0.0	
		T54A	-	679	0.0	-	493	0.4	-	231	0.0	
		T54S	-	679	0.0	-	493	3.0	-	231	0.0	
		V55A	-	680	0.1	-	492	0.2	-	231	0.0	
	R155K	R155K	-	683	0.0	+	493	99.2	-	232	0.0	
		R155T	-	683	0.0	-	493	0.2	-	232	0.0	
10	V36M	V36M	-	581	0.0	+	264	81.4	-	349	0.0	688
		T54A	-	584	0.2	-	271	0.0	-	351	0.0	
		T54S	-	584	0.0	-	271	14.8	-	351	0.0	
		R155I	-	586	0.0	-	274	0.4	-	352	0.0	
	R155K	R155K	-	586	0.3	+	274	96.0	-	352	0.0	
		A156V	-	586	0.0	-	274	0.4	-	352	0.0	
11	V36M	V36M	-	359	0.0	+	259	76.1	-	69	0.0	651
	R155K	R155K	-	374	0.0	+	397	99.2	-	78	1.3	
12		V36M	-	193	0.0	-	463	1.1	-	387	0.0	814
		T54A	-	196	0.0	-	492	0.8	-	417	0.0	
	R155K	R155K	-	196	2.6	+	495	81.0	-	417	4.3	

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)**

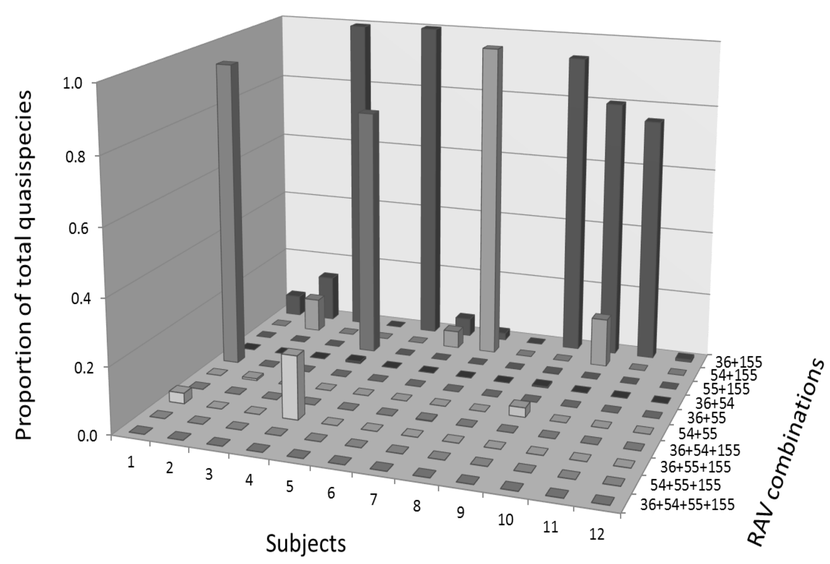

Figure 6: Multiple RAV combinations on single strains at VF.

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 of template molecules in the PCR process and yielded data to accurately quantify the number of unique viral sequences being studied in subjects that have failed treatment with one of the DAAs, boceprevir. Specifically, 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[4]. 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-sequence studies.

Conclusions

1) Ultra-deep sequencing identifies clinically relevant low frequency RAVs in circulating quasispecies at levels not detectable using Sanger sequencing
2) PCR amplification induced artefacts are eliminated using the primer ID methodology and analysis pipeline
3) Specifically, the study shows 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 prevalence 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

1. Jensen DM. 2011. A new era of hepatitis C therapy begins. The New England journal of medicine 364(13): 1272-1274.
2. Applegate TL, Gaudieri S, Plauzolles A, Chopra A, Grebely J, Lucas M, Hellard M, Luciani F, Dore GJ, Matthews GV. 2014. Naturally occurring dominant drug resistance mutations occur infrequently in the setting of recently acquired hepatitis C. Antiviral therapy.
3. Jabara CB, Jones CD, Roach J, Anderson JA, Swanstrom R. 2011. Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. Proceedings of the National Academy of Sciences of the United States of America 108(50): 20166-20171.
4. Le Pogam S, Seshaadri A, Kosaka A, Chiu S, Kang H, Hu S, Rajyaguru S, Symons J, Cammack N, Najera I. 2008. Existence of hepatitis C virus NS5B variants naturally resistant to non-nucleoside, but not to nucleoside, polymerase inhibitors among untreated patients. The Journal of antimicrobial chemotherapy 61(6): 1205-1216.

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

Http://iiid.Murdoch.edu.au