Naturally occurring dominant drug resistance mutations occur infrequently in the setting of recently acquired hepatitis C

Silvana Gaudieri, Tanya Applegate, Anne Plauzolles, Abha Chopra, Jason Grebely, Michaela Lucas, Margaret Hellard, Fabio Luciani, Greg Dore, Gail Matthews and the ATAHC cohort study group

Direct-acting antiviral (DAA) drugs for HCV infection

Pre-existing drug resistant variants

- Rapid replication and low fidelity of RNA-dependent polymerase results in HCV quasispecies (1 mutation per $10^3$-$10^5$ bases per replication cycle)
- Frequency of strains change over time within host due to selective pressures: replication efficiency; immune response (HLA, KIR, IFN); drugs
- Pre-existing DAA resistance associated variations (RAVs) identified in treatment naïve chronic-infected subjects (sanger-based technology) but not in the context of recently acquired hepatitis C infection
- Use of next-generation sequencing technology to determine frequency of RAVs in ATAHC cohort: circulating viruses in high-risk exposure populations, compensatory mutations, influence of non-drug selection pressures (immune response early in infection)

Pre-existing drug resistant variants (sanger)

- **NS3** protease (wt/RAV)
  - Chronic (n=205)
  - Acute (n=67)
  - Chronic (n=54)
  - Acute (n=3)

<table>
<thead>
<tr>
<th>Amino acid position</th>
<th>wt</th>
<th>RAV</th>
<th>Chronic (n=205)</th>
<th>Acute (n=67)</th>
<th>Chronic (n=54)</th>
<th>Acute (n=3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V36A/M</td>
<td>V/M</td>
<td>M/V</td>
<td>1.8/1.8</td>
<td>V/M</td>
<td>V/M</td>
<td></td>
</tr>
<tr>
<td>Q80K/R</td>
<td>Q/R</td>
<td>Q/K</td>
<td>0.8/6.9</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
</tr>
<tr>
<td>F43C/S</td>
<td>F/S</td>
<td>S/F</td>
<td></td>
<td></td>
<td>F/S</td>
<td>S/F</td>
</tr>
<tr>
<td>T54A/S</td>
<td>T/S</td>
<td>S/T</td>
<td>4.4/0</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
</tr>
<tr>
<td>Q41R/H</td>
<td>Q/H</td>
<td>Q/L</td>
<td>0.8/6.9</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
</tr>
<tr>
<td>V0/L</td>
<td>V/L</td>
<td>V/I</td>
<td></td>
<td></td>
<td>V/L</td>
<td>V/I</td>
</tr>
<tr>
<td>L155K/Q/T</td>
<td>L/K/Q/T</td>
<td></td>
<td>0.6/0</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
</tr>
<tr>
<td>A156S/T/V</td>
<td>A/S/T/V</td>
<td></td>
<td>5.8/13.8</td>
<td>0/0</td>
<td>0/0</td>
<td>0/0</td>
</tr>
</tbody>
</table>

Australian subjects Q80K

- Chronic 9.1% K (n=77)
- ATAHC 5.6% K (n=53)

Subject characteristics

Australian Trial in Acute Hepatitis C (ATAHC, 2004-2007)

| Genotype 1a | 50 (76) |
| Genotype 3a | 14 (21) |
| Genotype 1b | 2 |

HIV+ stored screening or pre-treatment sample available, with an HCV RNA> 1000 IU/ml were included.

- Enrolment
  - Acute asymptomatic (Am with seroconversion or 10x ALT) 24 mth
  - Acute (Ab+ 12 mth Enrolment + OR 6 mth 24 mth)
  - Chronic (n=205)
  - Acute (n=67)
  - Chronic (n=54)
  - Acute (n=3)
Method: FLX 454 Analysis (NGS)

- Raw sequence data collected in the lab or from clients containing the fasta sequences and quality scores using Roche 454 image and signal processing software
- Image data kept for 3 months
- .fna and .qual files kept for 3 months unless special arrangements made
- .fna and .qual files are provided to all clients with the reports as requested

Step 1
- Sort reads by MID (barcode)
- Use "sequence classifier by Id" tool (In house tool)

Step 2
- Trim Primer Sequences
- Use "Stripper" tool (In house tool)

Step 3
- Align Sequences using Reference sequence
- Use the tool "Sequence aligner". (In house tool)

Step 4
- If required separate by Amplicons
- Use "Select reads covering regions" (In house tool)

Step 5
- Re- do the alignments for each amplicon if required
- Use the tool "Sequence aligner". (In house tool)

Step 6
- Generate a SNP report
- Use "Codon usage" or "SNP analyser" tool (In house tool)

Results: Protease gene (n = 50 GT1a)

Coverage 3918-5239 reads

- % individuals with RAVs present (n=50)
- No association found between RAVs (frequency or number within subject) and HIV status or duration

Results: Polymerase gene (n = 50)

Coverage 2012-6722 reads

- Limited evidence for compensatory mutations – NS3

For most DAA resistance associated sites no evidence of co-variation in more than one subject

Results: NS5A (n=28)

- Limited evidence for compensatory mutations – NS3

For most DAA resistance associated sites no evidence of co-variation in more than one subject
Overlap between drug and immune pressure

Limited evidence of effect of immune pressure on frequency of RAVs – NS3

- HLA-A2-restricted epitope CINGVCWTV includes T54 and V55
  - 3/8 HLA-A2 positive >1% RAV at V55 and 2/17 HLA-A2- have RAV >1% at V55

- HLA-A24-restricted epitope MYTNVDQDL includes Q80.
  - 1/4 HLA-A24+ dominant K and 2/13 HLA-A24- have different dominant amino acid

- HLA-A2-restricted epitope HAVGIFRAA includes 155 and 156
  - 1/8 with HLA-A2 RAV 14.7% at 155. No change >1% within HLA-A2-

Summary

- Next generation sequencing identifies low frequency RAVs in most individuals but typically <1%
  - Relevance of low frequency variants in DAA treated subjects unknown
  - Presence of compensatory mutations will be investigated within a longitudinal cohort + boceprevir

- No obvious association between RAV frequency or number with HIV status, duration of infection or adaptive immune response

- Future use of primer ID adaptations/3rd gen sequencing technologies can eliminate amplification bias

Acknowledgements

Kirby Institute
Gail Matthews, Tanya Applegate, Greg Dore, Jason Grebely, Pip Marks
(ATAHC cohort)

NID
Simon Mallal, Michaela Lucas, Anne Plauzolles, Ian James, Linda Choo, Susan Herrmann, Abha Chopra, Don Cooper, Mark Watson

Funding from the National Health and Medical Research Council
### Molecular tagging of viral cDNA using a PID

**Unique tagging during RT reaction**

- **Viral RNA**
- **5'**
- **3'**

- **RT reaction**
- **Primer Id**
  - NNNNNNNN
  - (65K diff molecules)
  - Generic primer
  - 2 x 17bp sites
  - 3' **cDNA**
  - Primer
  - **TAG**
  - **PID**

- **TARGET BS**
- **3'**

- **Sample ID**
- **3 base tag (MID)**

- **Primer ID**
  - NNNNNNNN
  - (65K diff molecules)

- **8 random bases gives 4^8=65,536 unique combinations of PID**

- **Manufacture N8 critical**

- **Critical that the ratio of template to PID is kept low**

### PCR artefacts and errors

**PCR artefacts and errors**

- **Misincorporation**
- **Differential Amplification ("PCR bias")**
- **Recombination** (especially at high Template and # cycles)

### Considerations

**RT reaction**

- **PCR artefacts and errors**
  - **Misincorporation**
  - **Differential Amplification ("PCR bias")**
  - **Recombination**

**Sequencing**

- **Sequencing errors**

**DNA Library Prep**

- **NG Library Prep**

**Sample Pooling and NG Sequencing**

- **Sequencing**
  - **Sequencing errors**

**Data Analysis**

- **Data Analysis**
  - **Bin tags**

**Extract reads for each PID and error correct**

**Count unique PIDs and RAV's for each sample**

---


**We used the Primer ID method to identify RAVs in a diverse population of H77 variants**

**The Primer ID method was first used in HIV variant sequencing and relies on a special primer in the RT step**

**Image courtesy of Dr Cass Jabara UNC**

**PCR artefacts and errors**

**PCR artefacts and errors**

- **Misincorporation**
- **Differential Amplification ("PCR bias")**
- **Recombination**

**PCR artefacts and errors**

- **Misincorporation**
- **Differential Amplification ("PCR bias")**
- **Recombination**

---