Tracing the origin of the Indigenous Australian HBV subgenotype C4

Lilly Yuen, Margaret Littlejohn, Rosalind Edwards, Sarah Bukulatjpi, Paula Binks, Kathy Jackson, Jane Davies, Joshua S Davis, Steven Tong, Stephen Locarnini
Background

• CHARM
  • 126 of >180 cases genotyped
  • 100% infected with HBV/C4
  • Exclusively found among Aboriginal Australians
  • Dispersed over a vast area of the NT (1.4 million km²)

• Aboriginal Australians have the oldest continuous human culture outside of Africa
Methods

• Data: HBV genome sequences
  • Test sequences:
    • CHARM: HBV-C4 sampled from Indigenous Australians (n = 59)
  • Reference sequences from public database:
    • GenBank (n = 216): HBV-C subgenotypes, HBV-J, Southeast Asian Primate HBV (Gibbons and Orangutans)

• Recombination analysis (3Seq, Simplot, Simmonics)

• Phylogenetic analysis (MEGA v6)

• Time divergence analysis (BEAST)
HBV/C4 is a recombinant virus

Backbone: HBV genotype C
Surface gene: HBV genotype J

(Littlejohn et al, JMV 2014 86:695)
Surface protein of C4 and J are very similar

<table>
<thead>
<tr>
<th>'a' determinant domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>HBV-J</td>
</tr>
<tr>
<td>HBV-C4</td>
</tr>
<tr>
<td>Other HBV-C subgenotypes</td>
</tr>
</tbody>
</table>
Questions:

- How can HBV/C4 have a surface gene that is so similar to HBV/J?
  - J is rare (only one case worldwide, in Borneo)
  - C4 is exclusively found among Aboriginal Australians
  - C4 is present in such a vast area of NT, separated from Borneo by the Wallace Line

- Where did the recombination event occur?

- How did HBV/C4 get into Australia and when did this occur?

- Why is the C-part of HBV/C4 so different to the other HBV/C subgenotypes?

Modified from http://www.abc.net.au/science/indepth/img/Wallace/wallaceline.jpg
HBV/C4 clusters by NT regions
Biogeography clustering and language correlation

Two geographically distinct groups of C4

- Languages can be broadly classified into Pama-Nyungan (PN) and Non-Pama-Nyungan (NPN)
- HBV-C4a predominantly identified from NPN speaking individuals (grey region)
- HBV-C4b predominantly identified from PN speaking individuals (brown region)
Time divergence analysis - Method

- HBV does not have a molecular clock
- Time divergence analysis for HBV can be done by calibrating tree nodes of a phylogenetic tree with known dates, to infer the age of other nodes
- HBV has co-specified with modern humans since leaving Africa (Magnius & Norder 1995, Paraskevis et al 2013)
- We can equate:
  - oldest fossil or genetic-based age known for a particular population, who are infected with a unique HBV subgenotype, as
  - age of the most recent common ancestor (tMRCA) for that HBV subgenotype
  - calibration points to infer tMRCA of other HBV subgenotypes

<table>
<thead>
<tr>
<th>HBV Subgenotype</th>
<th>Population</th>
<th>Evidence</th>
<th>Prior (Age, kyo)</th>
<th>Reference</th>
</tr>
</thead>
</table>
| HBV-C3          | Melanesian   | Oldest AMH Fossil | 40 ± 5          | • Summerhayes 2010 Science 330:78
|                 |              |                 |                  | • O’Connell and Allen 2004 J Arch Sci 31:835
|                 |              |                 |                  | • Groube et al 1986 Nature 324:453            |
| HBV-C2          | North Asians | Oldest AMH Fossil | 40+              | • Fu 2014 Nature 514:445                       |
Time divergence analysis - Results

Colour Codes

- HBV-C4
- HBV-J
- Other C subgentotypes (C1-3, C5-16)
- SEA primate HBV

Surface gene J - component
Non-overlapping core gene C - component
### Inferred Most Recent Common Ancestor (MRCA) Ages

<table>
<thead>
<tr>
<th>MRCA</th>
<th>C4</th>
<th>C4 + J</th>
<th>All C + J</th>
</tr>
</thead>
<tbody>
<tr>
<td>J part</td>
<td>52.08</td>
<td>69.25</td>
<td>na</td>
</tr>
<tr>
<td>(surface gene)</td>
<td>[95% HPD: 29 - 77]</td>
<td>[95% HPD: 40 - 105]</td>
<td></td>
</tr>
<tr>
<td>C part</td>
<td>43.82</td>
<td>na</td>
<td>72.54</td>
</tr>
<tr>
<td>(non-overlapping core)</td>
<td>[95% HPD: 22 - 73]</td>
<td>[95% HPD: 39 - 116]</td>
<td></td>
</tr>
</tbody>
</table>

#### Malakunanjja II Rock Shelters
(Arnhem Land)

Zone of first human occupation
(thermoluminescence dating)

- 52 ± (7, 11) K year (top of artefact layer)
- 61 ± (9, 13) K year (bottom of artefact layer)


#### Madjedbebe
(Clarkson et al 2017, Nature 547:306)

Human occupation zone of around 65 K years old
Approximately 70K years ago

Sunda

aJ

Borneo

aC

C4

Wallace Line

Timor

Daly River

East Arnhem

Sahul


~ 70 kyo
Approximately 52K years ago

- Sunda
- Borneo
- Timor
- Daly River
- East Arnhem
- Wallace Line
- Sahul

~52 kyo
Summary

• Genotype C4 is the oldest of the modern non-African human HBVs at 52 kyo
  • significantly pushing back the current age proposed for human HBV [Paraskevis et al 2013]

• Around 70,000 years ago
  • People infected with the ancestral strain of HBV/C were already on Island SEA
  • A subgroup of these people (ancestors of modern Aboriginal Australians) became co-infected with an ancestral strain of HBV/J → recombination event → HBV/C4 subgenotype

• Around 52,000 years ago
  • The ancestors of Aboriginal Australians island hopped to Timor, Ashmore Reef and then into Australia when the sea level was ~62m below today, and remained in long term isolation
  • A group of these people carried HBV/C4, and then dispersed in 2 waves across NT

• Sometime after 52,000 years ago
  • People who were infected with only the ancestral strain of HBV/C and did not migrate to Australia subsequently dispersed throughout Asia and the Pacific
  • The ancestral strain of HBV/C also co-evolved with these people, and diverged into the present day HBV C subgenotypes

• This is the first time viral sequences have been used to track human migration
Remote community participants & contributors

Communities
- Daly River
  - Daly River, Nardidi, Peppiminarti, Wadeye
- Tiwi Island
  - Milikapiti, Nguiu, Wurrumiyanga
- Groote Eylandt
  - Angurugu
- East Arnhem
  - Elcho, Galiwin’ku, Gapuwiyak, Gove, Gunyangara, Milingimbi, Nhulunbuy, Numbulwar, Ramingining, Yirrkala
- West Arnhem
  - Goulburn Island, Gunbalunya, Jabiru, Maningrida, Minjilang, Oenpelli, Warruwi
- Darwin
- Katherine
  - Borroloola, Bulman, Kalkaringi, Katherine, Lajamanu, Ngukurr, Tennant Creek, Top Springs, Urapungra, Victoria River Downs
- Central Australia
  - Ali Curung, Alice Springs, Hermannsburg, Indulkana, Kaltukatjara, Mimili, Indulkana, Oodnadatta, Ti Tree, Wallace Rockhole, Yuendumu

Menzies School of Health Research
- A/Prof Steven Tong
- Dr Josh Davis
- Dr Jane Davies
- Sara Bukulatjpi (Galiwin’ku)
- Paula Binks

RDH Liver Clinic & Outreach Services
- Alice Springs
  - Catherine Marshall
  - Suresh Sharma
  - Rebecca Katiforis
  - Krispin Hajkowicz
  - Sarah Whiting
  - Jon Marrow
- Registrars
  - Dr Sushena Krishnaswamy
  - Dr Saliya Hewagama
Community Consultations & Knowledge Transfer

- **Consultations:**
  - A/Prof James Ward
    - Pitjantjatjara and Nurrunga decent
    - met in Jul 2016
  - ongoing

- **Knowledge transfer back to communities:**
  - Galiwin’ku, East Arnhem
  - Wade Eye, Daly River
Acknowledgements

VIDRL, Australia
- Margaret Littlejohn
- Rosalind Edwards
- Kathy Jackson
- Stephen Locarnini

UNSW, Australia
- Fabio Luciani

Menzies School of Health Research (Royal Darwin Hospital)
- Steven Tong
- Josh Davis
- Jane Davie
- Sarah Bukulatjpi
- Paula Binks
- Suresh Sharma
- Krispin Hajkowicz
- Sarah Whiting
- Cathy Corbett
- Sara Mgaieth
- Cam Jeremia
- Bridget Barber
- Claire Gordon

This research was supported by the Victorian Life Sciences Computation Initiative (VLSCI), an initiative of the Victorian Government, Australia, on its Facility hosted at the University of Melbourne, project number UOM0002.