Estimating HSV-2 superinfection using a novel custom genotyping platform

Why is HSV-2 superinfection relevant?

HSV-2 prophylactic vaccines are needed to control the HSV-2 epidemic

Does the natural immune response to HSV-2 protect against infection with another strain?

Will we need to create a vaccine that elicits a better or different response than natural infection?

Aims and Hypotheses

Aim:
Determine the prevalence of and risk factors for HSV-2 superinfection.

Hypotheses:
Prevalence of superinfection:
• Higher in women vs. men
• Higher in those with >10 sex partners vs. <10 sex partners
• Higher in HIV-infected vs. HIV-uninfected

Approach
Phase I: Next generation sequencing to identify population prevalent SNP
Phase II: Create genotyping platform and genotype paired samples

Disclosures

• U.S. National Institutes of Health: Grant recipient
• AiCuris GmbH, Sanofi: Principal Investigator
• Agenus, Genocea, Vical: Co-investigator

HSV-2 superinfection: What is known?

No standard method for HSV-2 genotyping

HSV-2 infection with more than one strain (“superinfection”) has been reported in small studies
PCRs based assays of variable HSV DNA repeats
HIV seronegative: 1/8 (12.5%)
HIV seropositive: 11/11 (100%)

Superinfection has been reported for other herpesviruses
CMV: 29% of pregnant women had 2 or more strains
Roest JID 2006
Ross et al JID 2010

HSV-2 Genomic variation

- Illumina sequencing
- Genital swab samples from 39 people
  - USA, Peru, Africa
- 2481 SNPs
- 456 prevalent SNPs evaluated for genotyping
Most informative SNPs for genotyping ranked using FastTagger

With 96 SNPs:
Able to determine whether samples match or do not match with >90% probability

96 SNPs best able to distinguish between specimens chosen for GoldenGate

GoldenGate Workflow

- Bead array based platform developed for human high throughput genotyping
- Biotin label DNA hybridize to allele specific SNP oligos
  - A "call" for each of 96 SNPs is generated
  - All samples with a call rate ≥ 90% SNPs considered valid
  - Comparison to deep sequencing: 8 samples matched at all sites

Methods

- Compare SNP results from paired samples
- Define superinfection: 25 SNPs different between pairs
- Confirm mismatched pairs are from the same person
  - Deletion/insertion polymorphisms on human DNA used for forensic analysis

Results

- 1152 samples
  - 59 negative controls (4-6 per plate)
    - Median call rate: 29% (IQR: 12-45%)
  - 1093 experimental samples
    - 1004 (92%) had call rate ≥ 90%
    - 960 paired samples (480 pairs)
      - 17 pairs without matching HSV sequences pending confirmation that they are from the same person
      - Excluded from this preliminary analysis
  - 11 SNPs did not perform well and were excluded

Demographics

<table>
<thead>
<tr>
<th></th>
<th>463 Pairs</th>
<th>1152 samples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Male</strong></td>
<td>215 (46%)</td>
<td>60 (59%)</td>
</tr>
<tr>
<td><strong>Median Age (IQR)</strong></td>
<td>34 (27, 44)</td>
<td>12 (3, 34)</td>
</tr>
<tr>
<td><strong>Continent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America (US)</td>
<td>274 (59%)</td>
<td>35 (28%)</td>
</tr>
<tr>
<td>South America (Peru)</td>
<td>60 (51%)</td>
<td></td>
</tr>
<tr>
<td>Africa (*)</td>
<td>133 (29%)</td>
<td></td>
</tr>
<tr>
<td><strong>Lifetime number of sexual partners median (IQR)</strong></td>
<td>11 (3, 34)</td>
<td></td>
</tr>
<tr>
<td><strong>HIV seropositive</strong></td>
<td>233 (29%)</td>
<td></td>
</tr>
<tr>
<td><strong>Median months between samples (IQR)</strong></td>
<td>5 (2, 11)</td>
<td></td>
</tr>
</tbody>
</table>

*Botswana, Cameroon, Kenya, South Africa, Tanzania, Uganda, Zambia, Zimbabwe
Prevalence of superinfection

<table>
<thead>
<tr>
<th>Number of mismatches</th>
<th>Related pairs N=463</th>
<th>Unrelated pairs N=1920</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>418 (90.3%)</td>
<td>4 (0.2%)</td>
</tr>
<tr>
<td>1-4</td>
<td>23 (4.9%)</td>
<td>11 (0.6%)</td>
</tr>
<tr>
<td>≥5</td>
<td>22 (4.8%)</td>
<td>1905 (99.2%)</td>
</tr>
</tbody>
</table>

Prevalence of superinfection (≥5 mismatches): 22/463 (4.8%), (95% CI: 2.8%, 6.7%)

Number of mismatches between paired specimens

Superinfection
N=22

Median Mismatches:
Superinfected pairs: 18 (IQR: 11-22)
Unrelated pairs: 23 (IQR: 19-27)

Risk factors for superinfection

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Univariate analysis</th>
<th>Multivariate analysis*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td></td>
<td>RR (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>Male</td>
<td>0.7 (0.3, 1.5)</td>
<td>0.337</td>
</tr>
<tr>
<td>Age in decades</td>
<td>0.8 (0.5, 1.2)</td>
<td>0.254</td>
</tr>
<tr>
<td>Continent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North America (US)</td>
<td>Ref</td>
<td>Ref</td>
</tr>
<tr>
<td>South America (Peru)</td>
<td>Ref</td>
<td>Ref</td>
</tr>
<tr>
<td>Africa (*)</td>
<td>Ref</td>
<td>Ref</td>
</tr>
<tr>
<td></td>
<td>3.6 (1.0, 13.3)</td>
<td>0.0491</td>
</tr>
<tr>
<td></td>
<td>5.5 (2.0, 15.2)</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>4.0 (1.4, 11.5)</td>
<td>0.011</td>
</tr>
<tr>
<td>Lifetime # sex partner (each additional ten)</td>
<td>0.9 (0.8, 1.2)</td>
<td>0.593</td>
</tr>
<tr>
<td>HIV seropositive</td>
<td>5.3 (2.2, 12.8)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Samples ≥3 years apart</td>
<td>4.3 (1.5, 12.3)</td>
<td>0.0077</td>
</tr>
</tbody>
</table>

Poisson regression
*No interaction between HIV status and continent

Modeling the prevalence of superinfection

Only 2 samples tested per person
Incorporate diversity and distribution of viral types
Prevalence superinfection ~20%

Strengths/Limitations

Strengths:
Large well characterized dataset
Novel, robust methodology based on rationally chosen SNPs to differentiate strains

Limitations:
Convenience dataset
Confirmation that additional samples are from the same person pending
Definitions for strains are needed

Conclusions

Naturally induced immunity at ganglion or mucosa is not sufficient to prevent reinfection
Implications for vaccine development.

BUT, prevalence of superinfection is relatively low (~5%)  
-This is likely lowest estimate, given that 2 samples were performed per person, variable follow up

Increased risk of superinfection in persons with HIV infection
-Lack of immunity?

Increased risk of superinfection in Africa
-Given high seroprevalence, increased exposure?
Acknowledgements

- **Sequencing**
  - Anna Rashevsky
  - Stacy Selke
  - Meei-Li Huang
  - Jon Guan
  - Cassie Sather

- **Modeling**
  - Dan Reeves
  - Josh Schiffer

- **Data Analysis**
  - Matt Fitzgibbon
  - Kurt Diem
  - Amalia Magaret
  - Anqi Cheng

- **PIs**
  - Anna Wald
  - David Koelle

- **Participants**
  - Specimens
    - Connie Celum
    - Jairam Lingappa

- **Funders**
  - NIH R21 AI96058
  - NIH P01 AI030731