Sequence profiling mucosal T-cell repertoires

Lisa Raeburn
Genome Sciences Centre

International Human Microbiome Congress
Vancouver, B.C., Canada
March 9 –11 2011
Mucosal Immune Repertoire

- Major undertaking in exploration of gut microbiome diversity
- Interaction between the host adaptive immune system to the microbiome less understood
- Profiling the T-cell immune repertoire
Activation of a T-cell in Gut

Gut Lumen

Epithelial Cell

T-cell

Lamina Propria

Commensal Microbe
Pathogenic Microbe

M-cell
Activation of a T-cell in Gut

- Gut Lumen
- Epithelial Cell
- Lamina Propria
- Commensal Microbe
- Pathogenic Microbe

pMHC: TCR Interaction

APC

MHC TCR

Infected Cell Killed
pMHC : TCR interaction

TCR alpha subunit
TCR beta subunit
CDR1
CDR2
CDR3
Peptide antigen
MHC
B2 Microglobulin
Generation TCRβ CDR3 diversity

TCRβ Locus

Somatic Recombination

TCRβ chain mRNA
Generation TCRβ CDR3 diversity

TCRβ Locus

Somatic Recombination

TCRβ chain mRNA
T-cell Repertoire Profiling Approach

Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing


Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes

T-cell Repertoire Library Construction

1. Obtain mRNA from PBMC prep (Peripheral Blood Mononuclear Cells, which include T-cells)

2. PCR amplify the TCR using 5' RACE

3. Shear off uninformative region of V region (by sonication) and sequence

Shear

Sequence from both ends

~150 bp
Handling Sequencing Error

- Real clonotypes can differ by a single nucleotide

- **Aggressive filtering** of raw data
  - Double stranded coverage
  - Q30 assignment to each base
  - No high quality discrepancies
J-gene as Benchmark for Error Handling
J-gene as Benchmark for Error Handling

Filtered Data

J gene sequences
J-gene as Benchmark for Error Handling

Filtered Data  J gene sequences  Apply D96 cutoff

Very rare clonotypes cannot be distinguished from sequence errors
## TCRβ CDR3 Diversity

<table>
<thead>
<tr>
<th>Subject</th>
<th>Blood Draw</th>
<th>Raw reads Sequenced</th>
<th>Total TCRβ sequences</th>
<th>Distinct TCRβ sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male 1</td>
<td>1,2</td>
<td>1,564,567,374</td>
<td>209,671,125</td>
<td>846,935</td>
</tr>
<tr>
<td>Male 2</td>
<td>1,2</td>
<td>307,058,456</td>
<td>6,219,383</td>
<td>193,551</td>
</tr>
<tr>
<td>Female</td>
<td>1,2</td>
<td>91,110,650</td>
<td>1,069,612</td>
<td>93,990</td>
</tr>
</tbody>
</table>
TCRβ diversity in peripheral blood

Donor 1:

Multiple libraries are required to capture full diversity of a blood sample
TCRβ diversity in peripheral blood

Donor 1:

Multiple libraries are required to capture full diversity of a blood sample

Repertoire diversity is substantially greater than the diversity captured in a single blood sample
Sharing of TCRβ repertoires

Female shared 1.1% nucleotide CDR3β sequences with male 1. Female shared 14.2% amino acid CDR3β sequences with male 1.
Sharing of TCRβ repertoires

Female shared 1.1% nucleotide CDR3β sequences with male 1.
Female shared 14.2% amino acid CDR3β sequences with male 1.

Male 2 shared 0.7% nucleotide CDR3β sequences with male 1.
Male 2 shared 11.7% amino acid CDR3β sequences with male 1.
Sharing of TCRβ repertoires

Female shared 1.1% nucleotide CDR3β sequences with male 1. Female shared 14.2% amino acid CDR3β sequences with male 1.

Male 2 shared 0.7% nucleotide CDR3β sequences with male 1. Male 2 shared 11.7% amino acid CDR3β sequences with male 1.

<table>
<thead>
<tr>
<th></th>
<th>CDR3 (aa)</th>
<th>CDR3 (nt)</th>
<th>CDR3(aa) / CDR3(nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>♀ total</strong></td>
<td>86,255</td>
<td>89,663</td>
<td>0.96</td>
</tr>
<tr>
<td><strong>♀ unique</strong></td>
<td>73,006</td>
<td>73,947</td>
<td>0.99</td>
</tr>
<tr>
<td><strong>♀ shared with ♂</strong></td>
<td>13,249</td>
<td>15,716</td>
<td><strong>0.84</strong></td>
</tr>
<tr>
<td><strong>♂ total</strong></td>
<td>165,931</td>
<td>177,763</td>
<td>0.93</td>
</tr>
<tr>
<td><strong>♂ unique</strong></td>
<td>144,781</td>
<td>150,992</td>
<td>0.96</td>
</tr>
<tr>
<td><strong>♂ shared with ♀</strong></td>
<td>21,150</td>
<td>26,771</td>
<td><strong>0.79</strong></td>
</tr>
</tbody>
</table>
Profiling the mucosal T-cell repertoire

• Normal tissue samples from 43 people

• Use same approach for profiling
  – Add barcoding step
# TCRβ CDR3 Diversity in colon/rectum

## Sequencing and Assembly Statistics

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of samples</td>
<td>43</td>
</tr>
<tr>
<td>Total read pairs</td>
<td>25,954,853</td>
</tr>
<tr>
<td>Total CDR3β sequences assembled*</td>
<td>60,234</td>
</tr>
<tr>
<td>Distinct ntCDR3β sequences assembled*</td>
<td>16,179</td>
</tr>
<tr>
<td>Distinct aaCDR3β sequences assembled*</td>
<td>12,439</td>
</tr>
</tbody>
</table>

*Values reported are in correct reading frame
<table>
<thead>
<tr>
<th>Sequencing and Assembly Statistics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of samples</td>
<td>43</td>
</tr>
<tr>
<td>Total read pairs</td>
<td>25,954,853</td>
</tr>
<tr>
<td>Total CDR3β sequences assembled*</td>
<td>60,234</td>
</tr>
<tr>
<td>Distinct ntCDR3β sequences assembled*</td>
<td>16,179</td>
</tr>
<tr>
<td>Distinct aaCDR3β sequences assembled*</td>
<td>12,439</td>
</tr>
</tbody>
</table>

*Values reported are in correct reading frame

Convergence
TCRβ Diversity Across Samples
TCRβ Diversity Across Samples

Distinct TCRβ (nt) sequences

Total Depth

Sample ID

Unique TCRβ rearrangements

Distinct TCRβ (nt) sequences

Total Depth

Sample ID

Distinct TCRβ (nt) sequences

Total Depth

Sample ID
TCRβ Diversity Across Samples

Distinct TCRβ (nt) sequences

Total Depth

Sample ID

7 31
TCRβ Diversity Across Samples
TCRβ Diversity Across Samples

[Graphs showing distribution of TCRβ diversity across different samples]
TCRβ aaCDR3 sequence sharing
TCRβ V-J Gene Usage

Mucosal samples

Public Samples
(Freeman et al. 2009)

Plots generated using Circos software (Krzywinski et al. 2009)
TCRβ V-J Gene Usage

Mucosal samples

Public Samples
(Freeman et al. 2009)

Plots generated using Circos software (Krzywinski et al. 2009)
TCRβ V-J Gene Usage

Mucosal samples

Public Samples
(Freeman et al. 2009)

Plots generated using Circos software (Krzywinski et al. 2009)
Conclusion

• Approach can be used to characterize mucosal immune response to microbes at sequence level resolution

• Approach can be applied in the investigation both normal and abnormal mucosal immune repertoires
Acknowledgements

Genome Sciences Centre
RL. Warren
JD. Freeman
RA. Holt
Holt Laboratory members
M. Krzywinski

Illumina Sequencing & Bioinformatics Groups (BCCA-GSC)
R. Moore

Tumour Tissue Repository (BCCA-DRC)
R. Barnes
P. Watson

Genome Canada, Genome BC, MSFHR, CIHR