HIV infection

HIV Progressor vs Controllers

HLA-HIV associations

Human MHC molecules
Human MHC molecules

Best determinant of HIV-1 disease progression: HLA-B

- >300 significant SNPs within the MHC region
- Specific amino acids in the HLA-B binding groove best determinants of HIV-1 control.
- Independent HLA-C effect

The international HIV-1 Controllers Study Science Express, 2010

Influence of HLA on disease progression

- Rate of HIV disease progression strongly associated with expression of particular HLA-B alleles.

**low VL**: HLA-B*5703, B*8101, B*5801, B*4201, B*0702, B*1518, B*1517, B*1302, B*1301, B*8901 and B*9002

**high VL**: HLA-B*4501, B*5802, B*1801, B*1503 and B*5301

CD8+ T-cell responses to different HIV proteins have discordant associations with viral load

Kapata et al., Nature. 2004 Dec 9;432(7018):769-75
Are B57 and B18 presenting epitopes from different HIV-1 proteins??

A computational approach!
How do we know which peptides bind to an MHC molecule?

- Via peptide elution from MHC molecules
- Mass spec to identify the peptide
- False positive rate: 5%
- Only few large studies available:

How do we know which peptides bind to an MHC molecule? **Via Quantitative ELISA**

- Obtain purified HLA
- Or recombinant heavy chain & β2m
- Obtain indicator peptide
- Perform dose titration of any inhibitory peptide
- Separate free from bound peptide
- Calculate binding and IC50
- Lower IC50, better binding

**Indicator peptide**

<table>
<thead>
<tr>
<th>Good binder</th>
<th>Bad binder</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**IEDB provides quantitative data and results of elution assays**

**Using the available data we can develop methods to predict MHC binding**

- **Simple/Extended Motifs**
  - Allowed/non allowed amino acids

- **Matrix models**
  - Peptide statistics from known peptides

- **Neural networks**
  - A machine learning system that can capture non-linear relations between residues in a peptide
Do the peptide positions contribute independently to binding?

Two adjacent amino acids may compete for the space in a pocket in the MHC molecule.

How does the network learn?

- Initially all weights are random
- Learning from data: Data is crucial!
  - MHC affinities: \textit{in vitro} measurements, not exact binding affinities, but relative measures with respect to a good binder. >1000 measurements for common MHC molecules

How good are the predictions?

\[ r^2 = 0.65 \]

\[ y = 0.92 x + 0.09 \]

Peters et al, PLOS Comp. Biol. 2006
How many MHC molecules do we find in the human population?

Seemingly small differences may have a big impact on the peptide binding motif!

A*6801 and A*6802 have very different peptide binding motifs.

Worldwide HLA class I diversity: Balancing selection

A great variety of binding motifs...
How does the network learn?
When data is not enough we extrapolate from “neighbor molecules”

HLA-B5701
YYAMGEMASTYENIAIVYDSYTAVLAYLWY

HLA-B2705
YHTEYREICAKTDEDTLYLNYHDYTAVLAYEWY

HLA-B3501
YYATYRNIFTNTYESNLYIRYDSYTAVLAYLWY

From proteins to T cell epitopes
20% processed 0.5-5% bind MHC 50% CTL response

=> 1/2000 peptide are immunogenic

Which peptides are the “binders” or predicted epitopes?
The output of prediction methods are IC50 values (a reflection of competition experiments in vitro). HOW do we then define possible epitopes?????

- Use a fix threshold, say 500 nM, and decide that all the peptides binding better than this threshold are likely epitopes.
- Take top 1-5% of all possible 9mers as possible epitopes.
- When comparing peptide binding to different HLA molecules: fix thresholds do NOT work! Use ranks instead.

pMHC predictions: NetMHC

cbs.dtu.dk/services/NetMHC