Immunoinformatics resources for the understanding of immunological information

A case study in personalized cancer immunotherapy

Bjoern Peters
La Jolla Institute for Allergy and Immunology

Overview

• Part I - Lecture: Biological Background
  – T cell immune responses target non-self entities
  – Cancer cells bear somatic mutations
  – Cancer immunotherapy aims to target immune responses to cancer cells

• Part II – Lecture: Bioinformatic guided approaches
  – Sequencing approaches identify tumor specific somatic mutations
  – HLA binding predictions can identify which of these will be immunogenic

• Part III – Hands on session: Design a personalized cancer vaccine

La Jolla Institute for Allergy and Immunology (LIAI)

HLA molecules as sensors of non-self

HLA = Human MHC molecules
Mouse Virus

CD8⁺ T cell epitopes in viral infection

APC

MHC-I

Virus

• How do peptides get loaded on MHC molecules?
• How do T cells distinguish self- from non-self peptides?

MHC-I - Antigen processing and presentation pathway

Proteasome

Protein

Golgi

TAP

ER

MHC-I

T Cell Receptor


MHC:peptide binding mode

X-Ray Structure: Madden, Cell 1993.
Viewer: Beaver and Ponomarenko, Immune Research, 2007

• Each human has 6 types of MHC molecules (alleles)
• >3000 alleles are known
• Distinct binding specificities → individual epitope repertoire
What is cancer?

- All cancers derive from single cells that have acquired the characteristics of continually dividing in an unrestrained manner and invading surrounding tissues.
- Cancer cells behave in this abnormal manner because of changes in the DNA sequence of key genes, which are known as cancer genes. Therefore all cancers are genetic diseases.

What is a mutation?

- **Germline mutation**
  - A change in the DNA sequence that can be inherited from either parent
- **Somatic mutation**
  - A change in the DNA sequence in cells other than sperm or egg
  - The mutation is present in the cancer cell and its offspring, but not in the patient’s healthy cells
Mutations & cancer genes

- Cancer genes are causally implicated in *oncogenesis*.
- Mutations in cancer genes can occur somatically or can be inherited.
- Mutations in some cancer genes can be inherited from parents, in which case they are present in every cell of the body. Such people are at a higher risk of developing cancer.
- Somatic mutations can occur in any of the cells of the body except the germ cells (sperm and egg) and therefore are not passed on to children.

Importance of somatic DNA changes in human cancer

Only 5–10% of cancer cases have a clear hereditary component, e.g., *BRCA1* and *BRCA2* in breast cancer.

Even in those cases where susceptibility is clearly inherited, somatic changes are required for cancer to develop.

Examples of mutations

<table>
<thead>
<tr>
<th>Sequence 1</th>
<th>Sequence 2</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTCGTTAGGCA</td>
<td>ACTCGTTAGGCA</td>
<td>Substitution</td>
</tr>
<tr>
<td>ACTCGTTAGGCA</td>
<td>ACTCGGCA</td>
<td>Deletion</td>
</tr>
<tr>
<td>ACTCGTTAGGCA</td>
<td>ACTCGTTATCAGGCA</td>
<td>Insertion</td>
</tr>
<tr>
<td>ACTCGTTAGGCA</td>
<td>ACTTTGCAGGCA</td>
<td>Inversion</td>
</tr>
<tr>
<td>ACTCGTTAGGCA</td>
<td>ACTCGTAGTTAGGCA</td>
<td>Duplication</td>
</tr>
</tbody>
</table>

Cancer progression

Mutations in multiple cancer genes are required for the development and progression of a single cancer.

- Benign Tumour
  - *In situ* cancer
  - Invasive cancer
  - Metastatic cancer
Neopitopes (Neoantigens)

- Cancers genomes accumulate mutations
- Mutations in coding regions are translated in mutated protein sequences
- Mutated peptides can be presented as epitopes on MHC to T cells

**Neopitopes** are presumably recognized by tumor-infiltrating lymphocytes (TILs)

**Neopitopes** are highly tumor-specific!

Schumacher & Schreiber, Science. 2015 Apr 3;348(6230):69-74

Cancer Immunotherapy

- **Vaccination:** Introduce or boost an immune response against a specific target (antigen)
- Cancer cells contain non-self antigens that *could* be recognized by T cells, but presence of cancer means this mechanism has failed, typically by the tumor suppressing immune responses
- **Checkpoint blockade treatments:** Block immune suppressive mechanisms to boost T cell immune responses against cancer cells.
- Problem: Checkpoint blockade is unspecific, and will also boost unwanted autoimmune responses
- **Personalized Cancer Immunotherapy:** Boost anti-tumor response with vaccine containing peptides corresponding to cancer mutations that can be recognized by T cells.

→ How can such a vaccine be designed?

Overview

- Part I - Lecture: Biological Background
  - T cell immune responses target non-self entities
  - Cancer cells bear somatic mutations
  - Cancer immunotherapy aims to target immune responses to cancer cells
- Part II – Lecture: Bioinformatic guided approaches
  - Sequencing approaches identify tumor specific somatic mutations
  - HLA binding predictions can identify which of these will be immunogenic
- Part III – Hands on session: Design a personalized cancer vaccine

DNA and RNA sequencing identifies tumor specific somatic mutations

Which mutations can be recognized by the patient's T cells?

→ Resulting peptides have to bind HLA molecules of the patient
HLA Typing: Targeted sequencing of HLA locus

**DNA isolation**

**PCR Primary Amplification** (exons 1-13)

**PCR Primary Amplification**

**Product Purification**

**Sequencing Reactions** (forward & reverse orientations)

**Sequencing Reaction Precipitation**

**Utilization of 96 sample sequencing instrument**

**Sequencing Analysis**

**Sequence-Based Typing**

### Calculate scoring matrix from affinities

Machine learning PSSM = Minimize the difference between predicted and measured binding affinities by varying the matrix values

<table>
<thead>
<tr>
<th>Peptide</th>
<th>log(IC_{50})</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOPOPQSPI</td>
<td>0.50</td>
</tr>
<tr>
<td>TSVANKYNN</td>
<td>0.72</td>
</tr>
<tr>
<td>KVEALYYV</td>
<td>2.37</td>
</tr>
<tr>
<td>FOPOPQSPI</td>
<td>3.42</td>
</tr>
<tr>
<td>LYKVKVSQL</td>
<td>3.46</td>
</tr>
<tr>
<td>FKSVEFDM</td>
<td>4.07</td>
</tr>
<tr>
<td>FOPOPQSPI</td>
<td>4.18</td>
</tr>
<tr>
<td>VLKLFWFSL</td>
<td>4.24</td>
</tr>
<tr>
<td>YMTLGGVVP</td>
<td>4.39</td>
</tr>
<tr>
<td>EDVKNAGGG</td>
<td>4.40</td>
</tr>
<tr>
<td>VYFLQMKRF</td>
<td>4.90</td>
</tr>
</tbody>
</table>

### Predictions available as webserver

- Immune Epitope Database (IEDB) Analysis resource


---

**Experimental Basis: MHC Binding Assay**

**List of peptides with allele specific binding affinity**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>IC_{50}</th>
</tr>
</thead>
<tbody>
<tr>
<td>QTVTFEAL</td>
<td>3.6</td>
</tr>
<tr>
<td>LKPDITYK</td>
<td>308</td>
</tr>
<tr>
<td>NFXMLNAS</td>
<td>50,000</td>
</tr>
<tr>
<td>AQSSCRTFR</td>
<td>38,000</td>
</tr>
<tr>
<td>CTYAGPPG</td>
<td>143</td>
</tr>
<tr>
<td>CFONHAYAK</td>
<td>50,000</td>
</tr>
</tbody>
</table>

log(IC_{50}) ~ Binding free Energy

Low IC_{50} → high affinity

**Impossible to measure all peptides**

→ Predict binding peptides using machine learning

Find function \( F_i \) in \( \{ F_1, F_2, F_3, \ldots \} \)

Many different approaches (ANN, SVM, HMM, LP, ... )

**T cell epitope mapping**

**HLA A*0201**

**Analysis**

...
Evaluating binding predictions

- Percentile rank < 0.5% = high affinity binder
- Percentile rank 0.5%-1% = intermediate binder
- Percentile rank 1% - 2% = low affinity binder
- Percentile rank 2% - 5% = borderline
- Percentile rank >5% is a non-binder
Overview

• Part I - Lecture: Biological Background
  – T cell immune responses target non-self entities
  – Cancer cells bear somatic mutations
  – Cancer immunotherapy aims to target immune responses to cancer cells

• Part II – Lecture: Bioinformatic guided approaches
  – Sequencing approaches identify tumor specific somatic mutations
  – HLA binding predictions can identify which of these will be immunogenic

• Part III – Hands on session: Design a personalized cancer vaccine

Steps

• Step 1: Identify sequence regions that contain all 9-mer peptides that are only found in the tumor
• Step 2: Run HLA binding prediction to identify 9-mer peptides in the sequence regions unique to the tumor that can be presented to T cells
• Step 3: Select the top peptide for each HLA allele
• Step 4: What is the un-mutated form of the chosen peptides in the patient? What is their MHC binding affinity?
• Step 5: Are the peptides really specific for the tumor? Examine this using NCBI BLAST
• Step 6: Decide: Which peptide would you choose?
Cancer genes

• There are two types of cancer genes:
  – Tumour suppressor genes
  – Oncogenes

• To date, we know of approximately 400 somatic “cancer genes” * but there are almost certainly more to be found

• COSMIC is a catalogue of somatic mutations found in cancer genes in human tumours and is available at: http://www.sanger.ac.uk/genetics/CGP/cosmic/

*TUMOUR SUPPRESSOR GENE

These genes normally function to PREVENT cell growth/division

*COSMIC v47 release. July 2010*

ONCOGENE

Genes which normally function to PROMOTE cell growth/division in a controlled manner