<table>
<thead>
<tr>
<th>Topic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cancer Genomics</strong></td>
<td>Brief review of cancer fundamentals, What is cancer and what causes it?</td>
</tr>
<tr>
<td><strong>Mining Cancer Genomic Data</strong></td>
<td><strong>Hands-on analysis</strong> to identify genomic changes in different cancers and identify new targets for therapy</td>
</tr>
<tr>
<td><strong>Towards personalized cancer treatments</strong></td>
<td>Recap on how the immune system normally detects cancer cells and how we can predict mutations that can be recognized by T cells</td>
</tr>
<tr>
<td><strong>Cancer Immunoinformatics</strong></td>
<td><strong>Hands-on analysis</strong> to design personalized cancer vaccines</td>
</tr>
</tbody>
</table>
What is Cancer?

“Cancer is a name given to a collection of related diseases, where some of the body’s cells begin to divide without stopping and spread into surrounding tissue”

Source: https://www.cancer.gov
What is Cancer?

“Cancer is a name given to a collection of related diseases, where some of the body’s cells begin to divide without stopping and spread into surrounding tissue”

Source: https://www.cancer.gov
Cancer is a disease of the Genome

- Caused by changes to genes that control the way our cells function, especially how they **grow and divide**.

- A major challenge in treating cancer is that every tumor is different: Each person’s cancer has a unique combination of genetic changes (both “driver” & “passenger”).

- As the cancer continues to grow, additional changes will occur.
Goals of Cancer Genome Research

- Identify changes in the genomes of tumors that drive cancer progression
- Identify new targets for therapy
- Select drugs based on the genomics of the tumor
- Provide early cancer detection and treatment response monitoring
- Utilize cancer specific mutations to derive neoantigen immunotherapy approaches
Finding Cancer Associated Mutations

Normal cell

Cancer cell

Sequencing machines

AATGCCA
TCATGTC
GGTATCG
CAGC ...

Somatic mutations

ACTGCCA
TCAGGTC
GGTATAG
TAGC ...

Identify all mutations specific to tumor cells

Filter out silent mutations
Mutations detected: Point mutations

Original (Tyrosine)

Silent (Tyrosine)

Missense (Cystine)

Nonsense (STOP)
## Mutations detected:

**Indels**

<table>
<thead>
<tr>
<th>Reference Sequence</th>
<th>CTGGTGACTAGTT</th>
</tr>
</thead>
</table>

**Indels** detected in the reference sequence.
Mutations detected:

Indels

Tumor Sequence 1

- - - C T G G T G A T T - - -

Deletion

CTAG deleted

Reference Sequence

- - - C T G G T G A C T A G T T - - -
Mutations detected:

**Indels**

- **Tumor Sequence 1**: 
  - Reference Sequence: - - - C T G G T G A - - -
  - Tumor Sequence: - - - C T G G T G A T T - - -
  - CTAG deleted

- **Reference Sequence**: 
  - Reference Sequence: - - - C T G G T G A C T - - -
  - CTAG inserted

- **Tumor Sequence 2**: 
  - Reference Sequence: - - - C T G G T G A C T T - - -
  - Tumor Sequence: - - - C T G G T A T C A G A C T T - - -
  - ATCA inserted

- **Deletion**
  - CTAG deleted

- **Insertion**
  - ATCA inserted
Mutations detected: Translocations
What can go wrong in cancer genomes?

<table>
<thead>
<tr>
<th>Type of change</th>
<th>Some common technology to study changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA mutations</td>
<td>WGS, WXS</td>
</tr>
<tr>
<td>DNA structural variations</td>
<td>WGS</td>
</tr>
<tr>
<td>Copy number variation (CNV)</td>
<td>CGH array, SNP array, WGS</td>
</tr>
<tr>
<td>DNA methylation</td>
<td>Methylation array, RRBS, WGBS</td>
</tr>
<tr>
<td>mRNA expression changes</td>
<td>mRNA expression array, RNA-seq</td>
</tr>
<tr>
<td>miRNA expression changes</td>
<td>miRNA expression array, miRNA-seq</td>
</tr>
<tr>
<td>Protein expression</td>
<td>Protein arrays, mass spectrometry</td>
</tr>
</tbody>
</table>

*WGS = whole genome sequencing, WX = whole exome sequencing, RRBS = reduced representation bisulfite sequencing, WGBS = whole genome bisulfite sequencing*
Genomics allows us to answer the question:

How many mutations are there in typical cancers?
Cancers in adults have more mutations than those in children.

~50-100 mutations

~4 or 5 mutations

Vogelstein et al.
Science (2013)
Cancers in adults have more mutations than those in children

~50-100 mutations

~4 or 5 mutations

Most of these mutations are likely “passenger” mutations

Vogelstein et al. Science (2013)
DNA damage from smoking and sun exposure...
Genomic approaches can identify the genes most commonly mutated in cancer.

Arrange all genes in a matrix, ordered by chromosomes.
Identifying genes most commonly mutated in cancer

Add all data together to see which genes are most often mutated
Identifying genes most commonly mutated in cancer

Are any of these known cancer genes?

Add all data together to see which genes are most often mutated
Identifying genes most commonly mutated in cancer

Many are famous porto-oncogenes, many others are new cancer genes!
Three Main Types of Cancer Genes:

- **Oncogenes**, such as **Ras**, normally function to accelerate cell division and growth. They can be mutated to act like stuck gas pedals.

- **Tumor suppressor genes**, such as **p53** normal act like breaks. Mutations can cause these breaks to fail.

- **DNA repair genes**, such as **BRCA1 & 2**, normally function to fix minor damage to DNA when it replicates. When these genes are mutated, DNA damage can accumulate and lead to cancer.
Cell growth and survival genes

Many participate in signaling pathways that promote cell proliferation (E.G. EGFR, Ras, BRAF, MEK etc.)
Cell growth and survival genes

Many participate in signaling pathways that promote cell proliferation (E.G. EGFR, Ras, BRAF, MEK etc.)
Regulators of Cell Cycle and Cell Death

Some stimulate the cell cycle:
- Cyclin D1
- CDK4

Some inhibit the cell cycle:
- P53
- RB

Oncogenes:
- Cyclin D1
- CDK4

Suppressor genes:
- P53
- RB
p53 Regulates Cell Division

Probably the most famous cancer gene that is mutated in about half of all tumors. Often called the ‘guardian of the genome’

- p53 normally shuts down cell division when a cell is stressed (e.g. by DNA damage)
- When DNA is damaged, p53 activates genes that stop cell growth or trigger the cell to die.
- Thus, p53 guards against changes to cells that might lead to tumor formation.
- It appears necessary to inactivate p53 to develop many forms of cancer.
Hands-on time!

https://bioboot.github.io/bgggn213_W19/lectures/#17

Part 1 Only Please
Representative H&E micrographs of rectus abdominis biopsies are displayed for two patients without cancer (left) and four patients with pancreatic cancer (right).
# Today’s Menu

<table>
<thead>
<tr>
<th>Cancer Genomics</th>
<th>Brief review of cancer fundamentals, What is cancer and what causes it?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mining Cancer Genomic Data</td>
<td><strong>Hands-on analysis</strong> to identify genomic changes in different cancers and identify new targets for therapy</td>
</tr>
<tr>
<td><strong>Towards personalized cancer treatments</strong></td>
<td>Recap on how the immune system normally detects cancer cells and how we can predict mutations that can be recognized by T cells</td>
</tr>
<tr>
<td><strong>Cancer Immunoinformatics</strong></td>
<td><strong>Hands-on analysis</strong> to design personalized cancer vaccines</td>
</tr>
</tbody>
</table>
Next Up:
Cancer Immunotherapy
- 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
- **Neoepitopes** are presumably recognized by tumor-infiltrating lymphocytes (TILs)
- **Neoepitopes** are highly tumor-specific!

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

**Q.** How can such a vaccine be designed?
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

Slide from: Bjoern Peters (LIAI)
HLA Typing: Targeted sequencing of HLA locus

DNA Isolation

PCR Primary Amplification (exons 1-5)

PCR Primary Amplification
Product Purification

Sequencing Reactions (forward & reverse orientations)

Sequencing Reaction Precipitation

Utilization of 96 sample sequencing instrument

Sequencing Analysis


Slide from: Bjoern Peters (LIAI)
TRADITIONAL CANCER THERAPIES

Kills Cancerous Cells
Kills Healthy Cells

DRUGS OR RADIATION

CANCER IMMUNOTHERAPIES

Selectively Kills Cancerous Cells

Unleash Patient’s Immune System

IMMUNOTHERAPY

Healthy Cells
Hands-on time!

https://bioboot.github.io/bimm143_F18/lectures/#17

Part 2: Designing a personalized cancer vaccine
Workflow:

- **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 to select peptides that can be presented to T cells in this patient.

- **Step 3:** Determine if the identified peptides are specific for the tumor (i.e. are they found anywhere else?)

- **Final question:** Which peptide would you choose?
Depictions of the peptide bound MHC and T-cell receptor

Note:
- Anchor residues in the peptide bind to the allele-specific pockets of the MHC molecule.
- Certain MHC molecules (alleles) preferentially bind peptides with specific anchor residues in the 8- or 9-amino-acid peptide sequence.
- We want our tumor specific residues to be within 8 to 9-mer sequences bound by a patient HLA alleles!

Bonus Slides
(For Reference)
Measuring and predicting MHC:peptide binding

**Experimental Basis: MHC Binding Assay**

- List of peptides with allele specific binding affinity

<table>
<thead>
<tr>
<th>Sequence</th>
<th>IC&lt;sub&gt;50&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>QIVTMFEAL</td>
<td>3.6</td>
</tr>
<tr>
<td>LKGPDYKG</td>
<td>308</td>
</tr>
<tr>
<td>NFCNLTSAF</td>
<td>50,000</td>
</tr>
<tr>
<td>AQSCRTFR</td>
<td>38,000</td>
</tr>
<tr>
<td>CTFAPGFGM</td>
<td>143</td>
</tr>
<tr>
<td>CFGNTAVAK</td>
<td>50,000</td>
</tr>
</tbody>
</table>

- log(IC<sub>50</sub>) ~ Binding free Energy

- low IC<sub>50</sub> → high affinity

**Impossible to measure all peptides**

→ Predict binding peptides using machine learning

Find function F<sub>i</sub> in F<sub>1</sub>, F<sub>2</sub>, F<sub>3</sub>, ...

F<sub>i</sub> (Sequence) ≈ Affinity

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

**T cell epitope mapping**

<table>
<thead>
<tr>
<th>ORF 1</th>
<th>MGQIVTMFEALPHIDEVINIVLIVITGIKAVYN...</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF 2</td>
<td>MGLKGPDIYKVQKFSVEFDMSHLNLTMPNACSANN...</td>
</tr>
<tr>
<td>ORF 3</td>
<td>MHNFCLTSAFNKKTDFDHLMSTIVSSLHLSIDGNNSY...</td>
</tr>
<tr>
<td>ORF 4</td>
<td>MSASQSCRTFRGRVLDMFRTAFGGKYMRSGGWGTGSD...</td>
</tr>
<tr>
<td>ORF 5</td>
<td>MHCTYAGPFMSRIILSSQEKTKFFTRRLAGTFTWTLS...</td>
</tr>
<tr>
<td>ORF 6</td>
<td>MKCFGNTAVAKCNYNHDAAECFDMLRLIDYNKAAALSFK...</td>
</tr>
<tr>
<td>ORF 7</td>
<td>MLMRNHLLYDLMGVFYCNYSKFWYLEHAKTGETSVPKC...</td>
</tr>
</tbody>
</table>
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 (IC50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FQPQNGSFTI</td>
<td>0.50</td>
</tr>
<tr>
<td>ISVANKIYLM</td>
<td>0.72</td>
</tr>
<tr>
<td>RVYEAALYYV</td>
<td>2.37</td>
</tr>
<tr>
<td>FQPQSGQFI</td>
<td>3.42</td>
</tr>
<tr>
<td>LYEKVKSQML</td>
<td>3.46</td>
</tr>
<tr>
<td>FKSVEFDMSSL</td>
<td>4.07</td>
</tr>
<tr>
<td>FQPQNGQFHF</td>
<td>4.18</td>
</tr>
<tr>
<td>VLMLPQWFL</td>
<td>4.24</td>
</tr>
<tr>
<td>YMTLGQVVF</td>
<td>4.39</td>
</tr>
<tr>
<td>EDVKNAVGV</td>
<td>4.40</td>
</tr>
<tr>
<td>VFYEQMKRF</td>
<td>4.90</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Offset: 4.3
Genetic and genomic approaches can identify a cancers molecular signature to usefully stratify tumors for treatment
Stratify tumors based on molecular patterns

- Good prognosis
  - Favorable response

- Bad prognosis
  - Unfavorable response

- Increased toxicity
Stratify tumors based on molecular patterns

Series of tumors

Microarray analysis

Hierarchical clustering

Correlation with outcome and clinicopathological features

Survival over months
TCGA Pan-Cancer project

12 tumor types
- Leukemia (LAML)
- Lung adenocarcinoma (LUAD)
- Lung squamous (LUSC)
- Kidney (KIRC)
- Bladder (BLCA)
- Endometrial (UCEC)
- Glioblastoma (GBM)
- Head and neck (HNSC)
- Breast (BRCA)
- Ovarian (OV)
- Colon (COAD)
- Rectum (READ)

Oomics characterizations
- Mutation
- Copy number
- Gene expression
- DNA methylation
- MicroRNA
- RPPA
- Clinical data

Thematic pathways
Samples
Genes/loci
Platforms
For example, breast cancer may be classified into various types based upon which proteins are expressed on the surface of the tumor cells. Breast tumors that express human epidermal growth factor 2 (HER2), estrogen receptor (ER), and progesterone receptor (PR), or are triple negative (do not express HER2, ER, or PR) behave differently and have different prognoses. Tumors that are HER2 positive are treated with medications that bind to HER2 (e.g. trastuzumab, lapatinib) and inhibit its activity. ER and PR are hormone receptors, and ER/PR positive tumors are treated with antihormonal therapies (e.g. tamoxifen and aromatase inhibitors). Triple negative tumors have the poorest prognosis and are unlikely to respond to HER2-targeted therapies or antihormonal therapies. Such cancers are usually treated very aggressively with chemotherapy.

As more has been learned about the molecular signature of various cancer subtypes, therapies that are specifically targeted to those signatures have been developed. Conventional chemotherapy acts on all rapidly dividing cells and does not distinguish between cancer cells and normal cells.
Classification of Breast Cancer

- **Infiltrating breast cancer**
  - ER IHC
    - ER-:
      - HER2 (+):
        - HER2+ classic
      - HER2 (-):
        - CK5/6 EGFR IHC
          - ER-, HER2- CK5/6+ and/or EGFR+ → Basal-like phenotype
          - ER-, HER2- CK5/6- and EGFR- → Non-basal, triple negative phenotype
    - HER2 IHC:
      - ER+ luminal B HER2 variant
      - ER+ luminal A
  - ER+:

- (trastuzumab, lapatinib)
- (tamoxifen and aromatase inhibitors)
- (aggressive chemotherapy)
Readings to find out more…

The Genetic Basis for Cancer Treatment Decisions

Janet E. Dancey, Philippe L. Bedard, Nicole Onetto, and Thomas J. Hudson

Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada
NCIC-Clinical Trials Group, Queen’s University, Kingston, ON K7L 3N6, Canada
Princess Margaret Hospital, Division of Medical Oncology and Hematology, University Health Network
Department of Medicine
Department of Medical Biophysics
Department of Molecular Genetics
University of Toronto, Toronto, ON M5S 1A1, Canada

Correspondence: tom.hudson@oicr.on.ca
DOI 10.1016/j.cell.2012.01.014

Personalized cancer medicine is based on increased knowledge of the cancer mutation repertoire and availability of agents that target altered genes or pathways. Given advances in cancer genetics, technology, and therapeutics development, the timing is right to develop a clinical trial and research framework to move future clinical decisions from heuristic to evidence-based decisions. Although the challenges of integrating genomic testing into cancer treatment decision making are wide-ranging and complex, there is a scientific and ethical imperative to realize the benefits of personalized cancer medicine, given the overwhelming burden of cancer and the unprecedented opportunities for advancements in outcomes for patients.
Cancer immune surveillance and escape

- Mutations in cells occur frequently
- The immune system has the capacity to detect and eliminate such mutated cells, and will do so on a regular basis
- Only when mutated cells find ways to hide from- or suppress an attack by the immune system, they can grow and spread unhindered leading to clinically apparent tumors

Can the immune response against an ‘escaped’ cancer be re-activated?

Nobel Prize 2018 in Medicine

- Identification of the molecules PD-1 and CTLA-4 that function as ‘T cell brakes’ (immune checkpoints)
- Blockade of PD-1 and CTLA-4 results in activation of T cells which has "fundamentally changed the outcome for certain groups of patients with advanced cancer"
- "Similar to other cancer therapies, adverse side effects are seen, which can be serious and even life threatening. They are caused by an overactive immune response leading to autoimmune reactions [...]"

"for their discovery of cancer therapy by inhibition of negative immune regulation"

Several trials for personalized cancer vaccines are currently ongoing.
Personalized Cancer Immunotherapy

After sequencing Tamara’s tumor and normal tissue, the team identified mutations expressed solely by cancer cells in her body. Schoenberger and LJI’s Bjoern Peters, PhD, developed a novel algorithm to select mutations that are recognized by the immune system. This algorithm was deployed to recognize the neoantigens that generated the strongest T cell response from Tamara’s tissue samples. These neoantigens were then presented to Tamara’s own T cells and cultured over a two-week period using 50 milliliters of her blood to develop a personalized vaccine.
Your Turn

Read and share your thoughts on the following class Readings

- Calling cancer's bluff with neoantigen vaccines
- Can genomics help detect early cancer and monitor treatment effectiveness?
- The increasing cost of cancer therapies

https://bioboot.github.io/bimm194_W18/readings/
1. Predict consequences of mutations

Map mutations into genome annotations to predict its possible effect

Tools to annotate consequences of mutations

- ANNOVAR
- VAGrENT
- Ensembl VEP
- ASOoVIR
- snpEff
- annTools
2. Assess the functional impact of nsSNVs

nsSNVs = non-synonymos Single Nucleotide Variant (missense)

ATC  GAA  GCA  CGT
Met  Glu  Ala  Gly

ATC  GAC  GCA  CGT
Met  Asp  Ala  Gly

Computational methods to assess the functional impact of nsSNVs

MutationTaster  LogRe  Condel  MutPred  SNPs&GO
CanPredict  PolyPhen2  CHASM  PMut  SNPefffect
SIFT  MutationAssessor  transFIC
3. Identify cancer drivers from somatic mutations

Patient cohort

Normal cell

Cancer cell

Sequencing machines

AATGCCA
TCATGTC
GGTATCG
CAGC ...

ACTGCCA
TCAGGTC
GGTATAG
TAGC ...

Somatic mutations

Which mutations are cancer drivers?

Find signals of selection across tumors
Cancer is an evolutionary process

Yates and Campbell et al, Nat Rev Genet 2012
How to differentiate drivers from passengers?

ACTG\text{CCTACGTCTCAACCGTGACTTCAAATCGCTTTAACCCTGACTCCCATGCTACTGC}
ATCTCGGGTTAACTCGACGTTTTT\text{CATGCATGTGTGCACCCCCAATATATATATGCA}\text{ACTT}
TTGTGCACCTCTGTACGCAGGGACGTGGCATGTGCAGCTGCTACGTCTCACGGTCGACTTCAAATCG
\text{TTAACCCGTACTCCCATGCTACTGCATCTCGGGTTAACTCGACGTTTT}
G\text{CATGCATGTGTGCACCCCCAATATATATGCAATCTTTTGTGCACCTCTGTCACGCGCGAGTTGGCACTGTCGCCCCTGTGTGCATGTGCACCTGCTACGTCTC
\text{TGAGTTTTG\text{CATGCATGTGTGCACCTGTCACCTGTCACCTGTCACCTG}}
How to differentiate drivers from passengers?

Find signals of positive selection across tumour re-sequenced genomes
Signals of positive selection

Recurrence

MuSiC-SMG / MutSigCV

Identify genes mutated more frequently than background mutation rate

Mutation clustering

OncodriveCLUST
PIK3CA is recurrently mutated in the same residue in breast tumours
IntOGen Mutations Analysis

To interpret catalogs of cancer somatic mutations.

**Cohort analysis**

Use this if you have a list of somatic mutations for a cohort of tumors and want to identify driver mutations, genes and pathways.

[View an example] [Analyse your data]

**Single tumor analysis**

Use this if you have a list of somatic mutations for a single tumor and want to rank them based on their implication in cancer development.

[View an example] [Analyse your data]

Gonzalez-Perez et al, Nature Methods 2013