Integrating Bioinformatics with Transplant Research at Comprehensive Transplant Centre
Bioinformatics at CTC

Bioinformatics core as a research support, primarily for CTC members

Provides state-of-the-art bioinformatics training and analytic support services to investigators interested in basic, pre-clinical, clinical and epidemiological transplant studies, within and outside Northwestern University

Works as a platform such that CTC and other NU researchers could do collaborative research works involving data from:
- Genomics and Transcriptomics studies
- Proteomics and Metabolomics experiments
- Public Databases
Bioinformatics Core at CTC

Bioinformatics core is focused to analyze high throughput and/or high-dimensional biological data arising from diverse technology platforms, including but not limited to,

- **Arrays:** Gene expression, transcript/isoform, SNP, Exon, methylation
- **Next Generation Sequencing:** DNA-seq, RNA-seq, WES, ChIP-seq, etc.

Core utilizes NU’s High Performance Computing Cluster, QUEST

Supports data storage

(Currently up to 5 Tb)

**QUEST6**

- Number of Nodes: 184 nodes and 5152 cores total, 28 cores/node
- Processor: Intel Xeon E5-2680, v4 14C 2.4 GHz
- Memory: Per node (Per Core) 128 GB
- Type: DDR4 1866 MHz
- Number of GPU nodes: 15
Core provides various analysis services, including but not limited to,

- Sequencing data processing, Differential gene/isoform analysis, Gene Set Enrichment Analysis, Pathway Analyses, Class discovery, pattern recognition, Genome Wide Association Studies, and Transcription factor binding site analysis
  - Micro-array
  - RNA-Seq
  - ImmunoSeq
  - ChIP-Seq
  - Single-cell Seq
  - Metabolomics
  - Proteomics
Micro-Arrays

- feature extraction
- quality control
- Normalization / batch effect correction
- Differential expression analysis / feature Selection
- Clustering/Classification
- biological interpretation of the results
Kidney Transplant + arrays + machine learning


<table>
<thead>
<tr>
<th>Data Set</th>
<th>Paired Samples, n</th>
<th>TX: subAR (% subAR prevalence)</th>
<th>Probability Threshold</th>
<th>% Negative (spared biopsy)</th>
<th>NPV</th>
<th>True Negative</th>
<th>False Negative</th>
<th>% Positive (pickup subAR)</th>
<th>PPV</th>
<th>True Positive</th>
<th>False Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery Set</td>
<td>530</td>
<td>400:130 (24.5)</td>
<td>0.375</td>
<td>74.7</td>
<td>88</td>
<td>349</td>
<td>42</td>
<td>25.3</td>
<td>61</td>
<td>83</td>
<td>51</td>
</tr>
<tr>
<td>Validation set 1</td>
<td>138</td>
<td>96:42 (30.4)</td>
<td>0.375</td>
<td>71.7</td>
<td>78</td>
<td>77</td>
<td>22</td>
<td>28.3</td>
<td>51</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Validation set 2</td>
<td>129/138</td>
<td>93:36 (27.9)</td>
<td>0.375</td>
<td>72.1</td>
<td>80</td>
<td>74</td>
<td>19</td>
<td>27.9</td>
<td>47</td>
<td>17</td>
<td>19</td>
</tr>
</tbody>
</table>
IPA to show biologic relevance of the differentially expressed genes that were used to populate the biomarker model and confirm shared pathways between the discovery and validation cohorts.

Gene Set Enrichment Analysis

IPA

GSEA
Liver Transplant + arrays + machine learning

Model 1 (2019)

Normalization: RMA
Discovery data: NU (46 AR, 45 TX)
Algorithm: Random forest
Features: 36 gene probe
AUC: 0.92
Validation Data: CTOT-14 (14 AR, 50 TX)

The probability score line slopes were positive preceding AR, and negative preceding TX and non-AR (TX + ADNR) \((P \leq .001)\) and following AR treatment.
Liver Transplant + arrays + machine learning

Model 2 (2020)
Normalization - fRMA
Data - 61 AR and 162 non-AR
Discovery and Validation data ratio - 70:30
Algorithm - Random forest
Features - 59 probes
AUC- 0.83

The probability score line slopes are positive preceding AR, and negative preceding nonAR (TX + ADNR) and following AR treatment.
Arrays to study MCMV reactivation

- Treatment with IS of latently infected mice alone does not induce viral reactivation, but transplant of latently infected allogeneic kidneys combined with IS facilitates MCMV reactivation in the graft and dissemination to other organs.
- The IS regimen effectively dampens allo-immune inflammatory pathways and depletes recipient anti-MCMV but does not affect ischemia-reperfusion injury pathways.
- MCMV reactivation similar to that seen in allogeneic transplants combined with also occurs after syngeneic transplants.

RNA-seq

- Gene expression and differential expression
- Alternative expression analysis
- Transcript discovery and annotation
- Allele specific expression
  - Relating to SNPs or mutations
- Mutation discovery
- Fusion detection
- RNA editing
Glutamine deprivation enhances directed differentiation of renal organoids

- Metabolomic and transcriptomic analysis during directed differentiation reveals key pathways that play a significant role in differentiation of renal organoids, especially the alanine, aspartate and glutamate pathway.

- Glutamine deprivation enhances differentiation with more robust acquisition of the PAX8 marker compared to control.
Metabolomic and RNA-seq data analysis

PCA analysis reveals distinct metabolic profiles from hESCs (D0) to NPCs (D9).

Venn Diagram

Impact analysis
Pathway impact analysis of metabolites identifies the alanine, aspartate and glutamate pathway as having the highest impact.

Volcano Plot
The 10 most highly expressed genes that differ between NPCs in complete media versus in glutamine deprivation conditions.
ChIP-seq identifies the binding sites of DNA-associated proteins and can be used to map global binding sites for a given protein.

<table>
<thead>
<tr>
<th>ChIP-Seq mark</th>
<th>Functional association</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3K4me3</td>
<td>Active promoters</td>
</tr>
<tr>
<td>H3K4me1</td>
<td>Active enhancers</td>
</tr>
<tr>
<td>H3K27ac</td>
<td>Active promoters and enhancers</td>
</tr>
<tr>
<td>H3K27me3</td>
<td>Inactive chromatin</td>
</tr>
<tr>
<td>RNA Pol II</td>
<td>Transcription</td>
</tr>
</tbody>
</table>

**Modification**

<table>
<thead>
<tr>
<th>Modification</th>
<th>Transcription</th>
<th>Histone-modified sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylation</td>
<td>Activation</td>
<td>H3 (K9,K14,K18,K56)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H4 (K5,K8,K12,K16)</td>
</tr>
<tr>
<td>Methylation</td>
<td>Activation</td>
<td>H3 (K4,K36,K79)</td>
</tr>
<tr>
<td>Repression</td>
<td>H3 (K9,K27)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>H4 (K20)</td>
<td></td>
</tr>
<tr>
<td>Phosphorylation</td>
<td>Activation</td>
<td>H3 (S10)</td>
</tr>
<tr>
<td>Ubiquitylation</td>
<td>Activation</td>
<td>H2B (K 1 2 3)</td>
</tr>
<tr>
<td>Repression</td>
<td>H2A (K 1 1 9)</td>
<td></td>
</tr>
<tr>
<td>S-nitrosylation</td>
<td>Repression</td>
<td>H3 (7)</td>
</tr>
<tr>
<td></td>
<td>H4 (K5,K8,K12,K16)</td>
<td></td>
</tr>
</tbody>
</table>
Study of host and viral genes interactions
RNA-seq and ChipSeq results

HCMV infection induces global loss of Pol II and H3K27Ac occupancy to cellular promoters

Mock-infected vs infected Kasumi-3 at 1dpi

loss of de novo transcription in response to HCMV infection

ImmunoSeq

highly optimized multiplex PCR-based assay that exclusively targets rearranged T-cell and B-cell receptor genes.

Project investigates the presence of donor reactive T cells (DRTCs) in blood and urine samples in the identification of post-transplant kidney rejection

DRTCs are identified by mixing irradiated donor PBMCs with recipient PBMCs followed by sorting for CD4 and CD8 cells and sequencing
scRNA-seq

Molecular Cell doi: 10.1016/j.molcel.2017.01.023
Analyze the transcriptomics of mouse lung cells with and without latent infection of MCMV before and after transplantation

<table>
<thead>
<tr>
<th>IPA CPTP vs CPTN</th>
<th>-log(B-H p-value)</th>
<th>Ratio</th>
<th>z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Th1 and Th2 Activation Pathway</td>
<td>11.1</td>
<td>0.327</td>
<td>#NUM!</td>
</tr>
<tr>
<td>Altered T Cell and B Cell Signaling in Rheumatoid Arthritis</td>
<td>10.4</td>
<td>0.411</td>
<td>#NUM!</td>
</tr>
<tr>
<td>T Helper Cell Differentiation</td>
<td>9.85</td>
<td>0.438</td>
<td>#NUM!</td>
</tr>
<tr>
<td>Th1 Pathway</td>
<td>9.39</td>
<td>0.347</td>
<td>4.218</td>
</tr>
<tr>
<td>Phospholipase C Signaling</td>
<td>8.96</td>
<td>0.256</td>
<td>3.571</td>
</tr>
<tr>
<td>STAT3 Pathway</td>
<td>8.96</td>
<td>0.326</td>
<td>0</td>
</tr>
<tr>
<td>Th2 Pathway</td>
<td>8.96</td>
<td>0.324</td>
<td>1.915</td>
</tr>
<tr>
<td>Role of NFAT in Regulation of the Immune Response</td>
<td>8.65</td>
<td>0.287</td>
<td>2.949</td>
</tr>
<tr>
<td>iCOS-iCOSL Signaling in T Helper Cells</td>
<td>7.3</td>
<td>0.324</td>
<td>4.382</td>
</tr>
<tr>
<td>PD-1, PD-L1 cancer immunotherapy pathway</td>
<td>7.3</td>
<td>0.33</td>
<td>-3.182</td>
</tr>
</tbody>
</table>
Thank You.

kandpal@northwestern.edu