INTRODUCTION TO PROTEOMICS II
<table>
<thead>
<tr>
<th>Outline</th>
<th>Terminology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview</td>
<td>Differential expression</td>
</tr>
<tr>
<td>Workflow</td>
<td>Targeted vs. Discovery</td>
</tr>
<tr>
<td>Challenges</td>
<td>Dynamic range</td>
</tr>
<tr>
<td>Experimental design</td>
<td>Inference</td>
</tr>
<tr>
<td>Tradeoffs</td>
<td>Bias</td>
</tr>
<tr>
<td>Statistics</td>
<td>Variability</td>
</tr>
<tr>
<td>Case studies</td>
<td>Normalization</td>
</tr>
<tr>
<td></td>
<td>Blocking</td>
</tr>
<tr>
<td></td>
<td>Randomization</td>
</tr>
</tbody>
</table>
‘Omics Technologies

Mass Spectrometry-Based ‘Omics
Some types of proteomics studies

- **Gene annotation**
  - Isoform 1: RSPIA
  - Isoform 2: RSPGH
  - Identification of splicing variants

- **Differential expression**
  - B/A fold change: 3.50, 1.01, 0.55
  - Assessing molecular differences between cell types, such as ESCs and iPSCs

- **Absolute abundance**
  - Copies per cell: $3 \times 10^9$, $4 \times 10^6$, $6 \times 10^2$
  - Investigation of relationships between transcription and translation

- **Temporal dynamics**
  - Proteome dynamics of fate change in ESCs

- **Spatial localization**
  - Localization: Cytoplasm, Cell surface, Nucleus
  - Defining the protein composition of mitotic chromosomes

---


© 2014 Regents of the University of Minnesota. All rights reserved.
Typical Proteomics Pipeline

Sample Prep
- Proteolytic Digest
- LC-ESI
- MS
- MS/MS
- Data Analysis

Mass Spec

Data Analysis
- Cellular Composition
- Organellar Proteome
- Interaction Protein
- Protein ID
- PTM Analysis

Adapted from Walther T, Mann M. JCB 2010;190:491-500
After a Decade - Potential Unmet

- Biological
  - Diverse
  - Ever changing
  - Dynamic Range

- Complex Workflow
  - Sample Preparation
  - Mass Spectrometry
  - Data Analysis

Bantscheff et al., 2011
Mass Spec + Inference

• The “Inference Problem”
• If we don’t find protein, it doesn’t mean it’s not there
• Identification and quantification requires statistics – which requires experimental design
Experimental Design

• Experimental design is the part of statistics that you carry out before you can carry out an experiment
• Proper planning can save many headaches
• You should design your experiments with a particular biological question in mind
Typical Proteomics Pipeline

Sample Prep
- Proteolytic Digest
- LC-ESI
- MS
- MS/MS
- Data Analysis

Experimental Design
- Cellular Composition
- Organellar Proteome
- Interaction Protein
- PTM Analysis
- Protein ID

Adapted from Walther T, Mann M. JCB 2010;190:491-500
Design Choices

• Discovery vs. Targeted
• Workflow Technology
• Tradeoffs
• Statistical Rigor
Number Analytes vs. Number Samples

Typical Discovery Experiment

- **Patient Group A**
- **Patient Group B**

**Prepare Samples**
- Peptides

**HPLC-MS/MS**
- MS\(^1\) & MS\(^2\) Spectra

**Preprocess**
- Quantify
- Identify

**Analyze Statistically**

**Differentially Abundant Peptides/Proteins**

**Sources of Bias and Variability**

**Population**

**Biological**

**Instrument**

**Sample Handling**
Why Variance Matters

**3 Replicates of Analyte 3**

**High Variance**

- Sample A
- Sample C

**Mean Fold Change 2.09**

**t-test p-value 0.07**

**3 Replicates of Analyte 3**

**Low Variance**

- Sample A
- Sample C

**Mean Fold Change 1.89**

**t-test p-value 0.0001**
Avoiding Bias & Variability in Results

• Appropriate number of biological samples
• Appropriate number of technical Replicates
• Normalization
• Blocking
• Randomization

Choices depend on workflow / technology choices
Normalization

Before normalization

![Box plots showing intensity before normalization: Array 1, Array 2, Array 3, Array 4.]

After normalization

![Box plots showing intensity after normalization: Array 1, Array 2, Array 3, Array 4.]

Block Design (Example Diagram)

Randomization

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tip 3</td>
<td>Tip 3</td>
<td>Tip 2</td>
<td>Tip 1</td>
<td>Tip 3</td>
</tr>
<tr>
<td>Tip 1</td>
<td>Tip 4</td>
<td>Tip 1</td>
<td>Tip 4</td>
<td>Tip 2</td>
</tr>
<tr>
<td>Tip 4</td>
<td>Tip 2</td>
<td>Tip 3</td>
<td>Tip 2</td>
<td>Tip 3</td>
</tr>
<tr>
<td>Tip 2</td>
<td>Tip 1</td>
<td>Tip 4</td>
<td>Tip 3</td>
<td>Tip 2</td>
</tr>
</tbody>
</table>
Before you start ...

Consult experts & statisticians

Plan, Plan, Plan!!!!