Proteomics
Objectives

• Proteomics
• Sample preparation and analysis
• What is integrated approach
• Lipid rafts
• Lipid rafts isolation
• Integrative analysis
Reference literature

• Advances in Experimental Medicine and Biology, Volume 919, Modern Proteomics – Sample Preparation, Analysis and Practical Applications by Hamid Mirzaei and Martin Carrasco (can be access through U of M libraries)

• Integrative Analysis of Subcellular Quantitative Proteomics Studies Reveals Functional Cytoskeleton Membrane–Lipid Raft Interactions in Cancer: Anup D. Shah et al, DOI: 10.1021/acs.jproteome.5b01035 J. Proteome Res. 2016, 15, 3451–3462
“Proteome” is the entire set of proteins, produced or modified by an organism or biological system.

"proteomics" is a large-scale comprehensive study of a specific proteome, including information on protein abundances, variations and modifications, along with their interacting partners and networks, in order to understand cellular processes.

“Clinical proteomics” is a sub-discipline of proteomics that involves the application of proteomic technologies on clinical specimens such as blood.
Protein synthesis

Post translational modifications (PTMs)

- human genome ≈ 25,000 genes,
- human proteome ≈ over 1 million proteins
- proteomic diversity.
- single genes encode multiple proteins
- Further complexity to proteome is due to PTMs

Types of PTMs

- Phosphorylation
- Methylation
- Acetylation
- Sulfation
- Amidation
- Hydroxylation
- Sumoylation
- Nitration
- Formylation
- Palmitoylation
- Glycosylation
- Ubiquitination

Molecular cell biology, seventh edition, by Lodish et al, chapter2, figure 2.15, page-36
Roles of PTM and associated diseases

- regulate activity, localization and interaction with proteins, nucleic acids, lipids, and cofactors
- PTMs are most often mediated by enzymes such as kinases, phosphatases, transferases and ligases, which add or remove functional groups, proteins, lipids or sugars to or from amino acid side chains, and proteases,
Sample preparations and experimental flow
Cells

• **Cell culture:**
  • Types of cells: Primary, secondary and cell lines
  • ATCC Cell Biology Collection: https://www.atcc.org
  • Easy accessibility, rapid growth rate and ease of manipulation (both genetically and biochemically) make cell lines an attractive model in research.

*To understand....*

• Differences between normal and aberrant cells
• To understand the signal transduction of molecular pathways, post-translational modifications (PTMs) such as phosphorylation
• to test the effect of various chemical compounds (for example inhibitors or activators) or pharmacological drugs on different cellular systems
Sample preparation

• **Tissue culture:**
  • is the growth of animal tissue outside (in a culture medium) of the organism,
  • in vitro model of animal tissue
  • mixed cell populations like *in-vivo*
  • major downside: of performing proteomic investigations on cultured cells is that they cannot provide accurate insight into disease progression in vivo

• **Organs:**

• **Exosomes:**
  • cell-derived vesicles size 30 and 100 nm
  • Secreted proteins can be found in both biological fluids and conditioned media from cell cultures

• **Urine:**
Sample preparation

• **Saliva**
  • Contains proteins of clinical relevance
  • 30% of blood proteins are also present in saliva,
  • an important tool for clinical application
Experimental flow

Sample Preparation → 2-D Electrophoresis → Spot Detection & Image Analysis

(I)  (II)  (III)

Enzymatic Digestion → Peptide-Mass Fingerprinting → Protein Identification

(IV)  (V)  (VI)

Peptide Sequencing via MS → Database Search

Experimental flow

Lysate Preparation
- Lysis
- Fractionation
- Depletion
- Enrichment
- Protein Assay

In-solution w/precipitation

Fractionation & Clean up
- Detergent Removal
- Enrichment
- Fractionation
- Desalting
- Peptide Assay

Peptide Preparation
- Buffer exchange
- Reduction
- Alkylation
- Digestion

SDS-PAGE
## In-silico

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Published</th>
<th>Free</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mascot [65]</td>
<td>1999</td>
<td></td>
</tr>
<tr>
<td>OMSSA [10]</td>
<td>2004</td>
<td>✓</td>
</tr>
<tr>
<td>InSpect [66]</td>
<td>2005</td>
<td>✓</td>
</tr>
<tr>
<td>Crux [67]</td>
<td>2008</td>
<td>✓</td>
</tr>
<tr>
<td>MS-GF+ [9]</td>
<td>2010</td>
<td>✓</td>
</tr>
<tr>
<td>Tide [12]</td>
<td>2011</td>
<td>✓</td>
</tr>
<tr>
<td>MassWiz [68]</td>
<td>2011</td>
<td>✓</td>
</tr>
<tr>
<td>Andromeda [69]</td>
<td>2011</td>
<td>✓</td>
</tr>
<tr>
<td>Byonic [70]</td>
<td>2012</td>
<td></td>
</tr>
<tr>
<td>Peaks DB [71]</td>
<td>2012</td>
<td></td>
</tr>
<tr>
<td>Morpheus [72]</td>
<td>2013</td>
<td>✓</td>
</tr>
<tr>
<td>MS Amanda [8]</td>
<td>2014</td>
<td>✓</td>
</tr>
</tbody>
</table>
Mass spectrometry based proteomics
Why the trypsin is the best protease for MS?

• Trypsin is the best choice as protease for mass spectrometry.

• Trypsin cleaves at the carboxylic side of lysine and arginine residues.

• Trypsin generates peptides of optimal sizes for mass spec analysis.

• Tryptic peptides have a strong C-terminal charge, and therefore they can be efficiently ionized.
Removal of detergents

• Why? Because of Ion suppression and interference
Bottom up proteomics

• To identify as many of the protein components of a biological sample as possible
Bottom-up proteomics

**CELL CULTURE OR TISSUE**
Proteins for bottom-up analyses come from a variety of sources.

**EXTRACTION OF PROTEINS**
Proteins are extracted and, in some cases, fractionated to reduce complexity.

**GENERATION OF PEPTIDES**
Proteins are denatured, reduced, alkylated, and digested into peptides. Peptides are, in some cases, fractionated to reduce complexity.

**AUTOMATED DATA ANALYSIS**
Peptides are identified using Proteome Discoverer software, an automated program capable of analyzing CID, HCD, EThcD, and ETD spectra.

**ANALYSIS BY LC-MS/MS**
Peptides are analyzed by LC-MS/MS on Orbitrap-based mass spectrometers.
Top-down proteomics

- Many problems can be addressed by studying the cell signaling pathways, disease state characterization, and biomarker discovery.

- Top-down proteomics: has the ability to detect degradation products, sequence variants, and combinations of post-translational modifications.

http://planetorbitrap.com/top-down-proteomics#tab:overview
CELL CULTURE OR TISSUE
Proteins for top-down analyses come from a variety of sources.

EXTRACTION OF PROTEINS
Proteins are extracted and denatured.

SEPARATION OF PROTEINS
Proteins are separated, most often by molecular weight, to reduce sample complexity and ensure maximal identification of intact proteins.

AUTOMATED DATA ANALYSIS
Intact proteins are identified in an automated fashion using ProSightPC software, including characterization of post-translational modifications, sequence polymorphisms, and cleavage sites.

ANALYSIS BY LC-MS/MS
Intact proteins are analyzed by LC-MS/MS on Orbitrap-based mass spectrometers.

http://planetorbitrap.com/top-down-proteomics#tab:overview
Outcome of the Evaluation-3:

- There was wash of sodium hydroxide + twice wash of Tris-buffer:

<table>
<thead>
<tr>
<th>Name of proteins</th>
<th>Abbreviation</th>
<th>Counts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma membrane</td>
<td>PM</td>
<td>7</td>
</tr>
<tr>
<td>Plasma membrane and other locations</td>
<td>PMOT</td>
<td>5</td>
</tr>
<tr>
<td>Membrane part</td>
<td>MP</td>
<td>6</td>
</tr>
<tr>
<td>Membrane part and other location</td>
<td>MPOT</td>
<td>1</td>
</tr>
<tr>
<td>Mitochondria</td>
<td>MITO</td>
<td>19</td>
</tr>
<tr>
<td>Cytoplasmic</td>
<td>CYTO</td>
<td>6</td>
</tr>
<tr>
<td>Endoplasmic reticulum</td>
<td>ER</td>
<td>5</td>
</tr>
<tr>
<td>No information available</td>
<td>NOINFO</td>
<td>10</td>
</tr>
<tr>
<td>Nucleus</td>
<td>NEU</td>
<td>19</td>
</tr>
<tr>
<td>Golgi apparatus</td>
<td>GOLGI</td>
<td>1</td>
</tr>
<tr>
<td>Intracellular</td>
<td>INCE</td>
<td>2</td>
</tr>
<tr>
<td>Extracellular</td>
<td>EXCE</td>
<td>0</td>
</tr>
<tr>
<td>Multiple regions</td>
<td>MULTI</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>85</td>
</tr>
</tbody>
</table>

Results: Plasma membrane proteins: 7 out of 85 ≈ 8.2%

Next focus: To remove the Mitochondrial and Nuclei proteins
<table>
<thead>
<tr>
<th>Protein Ids</th>
<th>Name of proteins</th>
<th>Subcellular location</th>
<th>#TMHs</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1LNY3</td>
<td>Neural cell adhesion molecule 1 OS=Rattus norvegicus GN=Ncam1 PE=1 SV=2</td>
<td>MP</td>
<td>1</td>
</tr>
<tr>
<td>A0A0G2K0M8</td>
<td>Neural cell adhesion molecule 1 OS=Rattus norvegicus GN=Ncam1 PE=1 SV=1</td>
<td>PM</td>
<td>2</td>
</tr>
<tr>
<td>F1LUV9</td>
<td>Neural cell adhesion molecule 1 OS=Rattus norvegicus GN=Ncam1 PE=1 SV=3</td>
<td>MP</td>
<td>1</td>
</tr>
<tr>
<td>A0A0G2JSK5</td>
<td>Integrin beta OS=Rattus norvegicus GN=Itgb1 PE=1 SV=1</td>
<td>PM</td>
<td>1</td>
</tr>
<tr>
<td>P48675</td>
<td>Desmin OS=Rattus norvegicus GN=Des PE=1 SV=2</td>
<td>PM</td>
<td>0</td>
</tr>
<tr>
<td>Q6RU5V</td>
<td>Ras-related C3 botulinum toxin substrate 1 OS=Rattus norvegicus GN=Rac1 PE=1 SV=1</td>
<td>PMOT</td>
<td>0</td>
</tr>
<tr>
<td>A0A0G2KB63</td>
<td>Prohibitin-2 OS=Rattus norvegicus GN=Phb2 PE=1 SV=1</td>
<td>MP</td>
<td>0</td>
</tr>
<tr>
<td>A0A096PL6B</td>
<td>Fibronectin OS=Rattus norvegicus GN=Fn1 PE=1 SV=2</td>
<td>PM</td>
<td>0</td>
</tr>
<tr>
<td>Q63016</td>
<td>Large neutral amino acids transporter small subunit 1 OS=Rattus norvegicus GN=Slc7a5 PE=1 SV=2</td>
<td>PM</td>
<td>9</td>
</tr>
<tr>
<td>A0A0G2K695</td>
<td>Myoferlin OS=Rattus norvegicus GN=Myof PE=1 SV=1</td>
<td>PMOT</td>
<td>1</td>
</tr>
<tr>
<td>A0A0G2JTH4</td>
<td>Leukocyte surface antigen CD47 OS=Rattus norvegicus GN=Cd47 PE=1 SV=1</td>
<td>PM</td>
<td>5</td>
</tr>
<tr>
<td>Q6P742</td>
<td>Proteolipid protein 2 OS=Rattus norvegicus GN=Plp2 PE=2 SV=1</td>
<td>MP</td>
<td>4</td>
</tr>
<tr>
<td>A0A0H2UHQ0</td>
<td>4F2 cell-surface antigen heavy chain OS=Rattus norvegicus GN=Slc3a2 PE=1 SV=1</td>
<td>PM</td>
<td>1</td>
</tr>
<tr>
<td>O35112</td>
<td>CD166 antigen OS=Rattus norvegicus GN=Alcam PE=1 SV=1</td>
<td>PM</td>
<td>1</td>
</tr>
<tr>
<td>B0BN20</td>
<td>Tetraspanin OS=Rattus norvegicus GN=Tspan6 PE=1 SV=1</td>
<td>MP</td>
<td>4</td>
</tr>
<tr>
<td>B0BNK1</td>
<td>RAB5C, member RAS oncogene family OS=Rattus norvegicus GN=Rab5c PE=1 SV=1</td>
<td>PMOT</td>
<td>0</td>
</tr>
<tr>
<td>B0BNG3</td>
<td>Lectin, mannose-binding 2 OS=Rattus norvegicus GN=Lman2 PE=1 SV=1</td>
<td>PMOT</td>
<td>1</td>
</tr>
<tr>
<td>P14562</td>
<td>Lysosome-associated membrane glycoprotein 1 OS=Rattus norvegicus GN=Lamp1 PE=1 SV=1</td>
<td>PMOT</td>
<td>1</td>
</tr>
<tr>
<td>Q9WUW2</td>
<td>Vesicle associated membrane protein 2B OS=Rattus norvegicus GN=Vamp2 PE=2 SV=1</td>
<td>MP</td>
<td>1</td>
</tr>
<tr>
<td>A0JN29</td>
<td>Limb and neural patterns OS=Rattus norvegicus GN=Lnpk PE=1 SV=1</td>
<td>MPOT</td>
<td>2</td>
</tr>
</tbody>
</table>
Application to clinical proteomics: Phases

<table>
<thead>
<tr>
<th>Phase</th>
<th>Samples</th>
<th>Process</th>
<th>Numbers of analytes</th>
<th>Numbers of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery</td>
<td>Proximal fluids, Cell line supernatants, Animal model plasma, ‘Gold standard’ human plasma (reduced biological variation)</td>
<td>Abundant protein depletion, Extensive fractionation, LC-MS/MS (low throughput)</td>
<td>1,000s</td>
<td>10s</td>
</tr>
<tr>
<td>Qualification</td>
<td>‘Gold standard’ human plasma (reduced biological variation)</td>
<td>Abundant protein depletion, Modest fractionation, +/- Immunoaffinity peptide enrichment, SID-MRM-LC-MS/MS (low-moderate throughput; high multiplexing)</td>
<td>30–100</td>
<td>10s</td>
</tr>
<tr>
<td>Verification</td>
<td>Population-derived human plasma (normal biological variation)</td>
<td>Abundant protein depletion, Modest fractionation, +/- Immunoaffinity peptide enrichment, SID-MRM-LC-MS/MS (moderate throughput; high multiplexing)</td>
<td>10s</td>
<td>100s</td>
</tr>
<tr>
<td>Validation and assay</td>
<td>Population-derived human plasma (normal biological variation)</td>
<td>Immunoassay (high throughput; low multiplexing)</td>
<td>4–10</td>
<td>Many 1,000s</td>
</tr>
</tbody>
</table>

(Kate Ri)
Clinical proteomics: Biomarker

- **Number of candidates**: 100-1,000 candidates
  - **Triage**
    - **Detectability in clinical samples**: PRM, AIMS, SWATH, SRM with crude SIS
    - **A assay platform**: PRM, AIMS, SWATH, SRM with crude SIS
    - **A assay characteristics**:
      - Specificity: +++
      - Repeatability: +++
      - Precision: +++
      - LOD/LLOQ: N/A
      - Linear range: N/A
      - Reproducibility: N/A
    - **Degree of assay validation**: Low
- **Quantification**: 50~100
  - **Qualification**
    - **Confirm the relative abundance changes**: SRM with SIS peptides
    - **Eliminate false-positive candidates**: ELISA
    - **A assay characteristics**:
      - Specificity: ++++
      - Repeatability: ++++
      - Precision: ++++
      - LOD/LLOQ: N/A
      - Linear range: N/A
      - Reproducibility: N/A
    - **Degree of assay validation**: Moderate
- **Verification**: <10
  - **Define the sensitivity specificity of biomarker panel**:
    - SRM with "winged" SIS or labeled full-length protein
    - **A assay characteristics**:
      - Specificity: ++++
      - Repeatability: ++++
      - Precision: ++++
      - LOD/LLOQ: +++
      - Linear range: +++
      - Reproducibility: +++
    - **Degree of assay validation**: High
Integrative analysis of subcellular quantitative proteomics studies reveals functional cytoskeleton membrane-Lipid raft interactions in cancer

Contributors:


* The University of Queensland Diamantina Institute, The University of Queensland, Translational Research Institute, Brisbane, Queensland 4102, Australia

# Ovarian Cancer Action Research Centre, Department of Surgery and Cancer, Imperial College London Hammersmith Campus, London W12 0NN, United Kingdom

$ Division of Bioinformatics, The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville Victoria 3052, Australia
Lipid rafts as micro-domain for cancer progression

- Microheterogeneity in plasma membrane
- Micro-domain structure
- Plays role in molecular association and transduction in cell signalling

Ref: 1. F. Julicher and M. Zapotocky, MPI-PKS, Dresden Signal transduction in olfactory cilium: stochastic dynamics of ion channels
What is integrative & network analysis?
Is there any **cross-cancer similarities** in protein profiles of rafts?

- Regulate molecular interactions
- Cell signaling in carcinogenesis

Quantitative lipid rafts proteomics

- Raft-cytoskeleton interaction
- OPCML: opioid binding protein cell-adhesion molecule
- Protein-protein interaction & simulation analysis
Materials and methods
Lipid rafts changes during cancer progression

From reported study
- Breast cancer
- Melanoma (skin) cancer

From integrative study
- Renal cell cancer

Overlap analysis:
Up-regulation & Down-regulation

Cytoskeletal protein enrichment analysis

- “Cytoskeleton-associated proteins from gene ontology (GO)
- Supplement table-1

- 48293 Human: Uniprot proteins
- 4393 Cytoskeleton-associated proteins
- 684: Lipid rafts-localized cytoskeleton from Raftprot
Quantitation

1. Plasma:
   - denature, reduce, alkylate, quench, and digest

2. Peptides:
   - stable isotope-labeled standards

3. Peptides + Standards:
   - solid-phase extraction

4. Desalted Peptides + Standards:
   - standard-flow UHPLC-MRM/MS

5. Quantitative Data
Triplex-SILAC

**Light:** Control SKOV3 cells, normal isotopic Lys, and Arg

**Medium:** P95R mutant expressing SKOV3 cells, $^2$H$_4$-Lys, and $^{13}$C$_6$$^{14}$N$_4$-Arg

**Heavy:** WT OPCML expressing SKOV3 cells, $^{13}$C$_6$$^{15}$N$_2$-Lys, and $^{13}$C$_6$$^{15}$N$_4$-Arg
Isolation of lipid raft fraction: Detergent resistant membrane (DRM)

DRM
DRM method
Preparation of DRM

- Cholesterol-disruption [methyl-β-cyclodextrin, filipin, nystatin]
- Cells lysis for 1 h at 4°C in 1% Triton X-100 (in 25 mM Mes, pH 6.5)
- Centrifugation for 10 min at 12,000 g (remove nuclear fraction)
- Lysates were combined and mixed with an equal volume of 90% sucrose (final sucrose concentration of 45%)
- Additional layers of 35% and 5% sucrose in MBS were gently placed on top
- Centrifugation at 166,000 g for 18 h at 4°C
- low-density light-scattering band (18% sucrose) was extracted,
- Diluted 4X in MBS, and centrifuged an additional 2 h @ 166,000 g, 4°C to pellet the detergent-resistant material (DRM)
LC-MS/MS, Database searching, SILAC quantitation, and statistical analysis

nLC-QTOF MSMS
- MS/MS mode
- 8MS
- 4 MS/MS per second

SwissProt database:
- Carbamido-methylation
- SILAC ratios: L/H and L/M

Statistics:
- Quantitative Proteomics p-value calculator
Results
Differential expression of lipid raft proteins during cancer progression

Unique expression; 23 altered (19 Upregulated and 4 down regulated)

Link for Supplement-2
Upregulation of cytoskeleton protein

19 up-regulated
4 down-regulated

12 out of 23: cytoskeleton

Enrichment in raft protein in malignant tumors

Breast cancer
- 36 % of cytoskeleton prot.

RCC
- 45% of cytoskeleton prot.

Melanoma
- 31% of cytoskeleton prot.
Tumor suppressor: OPCML

- Opioid binding protein cell-adhesion molecule (OPCML)
- OPCML is a GPI-anchored protein localized to membrane rafts
- 118 epithelial ovarian cancer (EOC) cases → inactivation of OPCML at 11q25$
- 92% of 489 high-grade serous ovarian cancers reported loss of OPCML expression#

WT and P95R OPCML alters the rafts proteome of SKOV3 cells

**Integrated analysis**

- **Enrichment of cytoskeleton assembly at lipid raft membranes as a common feature in cancer development and progression**

**Then..........opposite should be:**

- **Suppressor should reverse the above association**
OPCML regulates lipid raft proteome

(a) Triplex SILAC experiment workflow
(b) Domain architecture of OPCML
(c) replacement of proline to arginine
(d) magnitude and significance of the lipid raft protein level changes
(e) magnitude and significance of the lipid raft protein level changes
(f) Overlap between WT OPCML and P95R OPCML-regulated lipid raft proteins
Further reading for preparations

• Why the trypsin is being used as most common enzyme for mass spectrometry based proteomics
• What is meant by SILAC? How does it work?
• Proteomics, and significance (e.g., in target discoveries in various diseased conditions; list at least two application)
• General protocol for proteomics: Top-down and bottom up
• Why the trypsin is the best choice for MS based proteomics
• How can you remove the detergent after protein digestion? Why it is so important prior to MS analysis?
Thank you!