MS/MS to Targeted Proteomics (MRM)

How it worked on the Human Lens Proteome

Jayson Falkner PhD
jay@singleorganism.com
“With only a few exceptions, what the genomics companies are doing right now is recreational genomics,” Dr. Goldstein said in an interview. “The information has little or in many cases no clinical relevance.”

The undiscovered share of genetic risk for common diseases, he said, probably lies not with rare variants, as suggested by Dr. Goldstein, but in unexpected biological mechanisms.

Nicholas Wade, New York Times, April 15th 2009
Genomics has already arrived

**Commercial Products**
- **Affymetrix**: Genome Wide Human SNP Array 6.0 (906,600 SNPs)
- **Illumina**: human1m-duo bead chip (1.1 million evenly distributed loci)

*Table 1: GWAS published per year as listed in the NIGRI/NIH www.genome.gov resource.*

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of GWAS Studies</strong></td>
<td>14</td>
<td>441</td>
<td>901</td>
<td>137 (as of March)</td>
</tr>
</tbody>
</table>

**Commercial Services**
- **deCODEme**: [http://www.decodeme.com/](http://www.decodeme.com/)
  - “Complete Scan”, “Cardio Scan”, “Cancer Scan”
- **23andMe**: [https://www.23andme.com/](https://www.23andme.com/)
  - “Time Magazine’s 2008 Invention of the Year, now $399.”
Proteomics is arriving

Tandem Mass Spectrometry (MS/MS)
Discovery of proteins and post-translational modifications in a sample.

Targeted Proteomics (MRM aka SRM)
Rapid and sensitive monitoring of proteins and post-translational modifications in a sample.
Proteomics Timeline

Peptide Ionization: ESI and MALDI

MS/MS: LCQ and Sequest

On-line Separation: MudPIT
Washburn et al. (2001) "Large-scale analysis of the yeast proteome by multidimensional protein identification technology" Nat Biotech Vol 19

MRM: MIDAS and QTRAP
Unwin et al (2005) “Multiple Reaction Monitoring to Identify Sites of Protein Phosphorylation with High Sensitivity” MCP 4.8
MudPIT: More MS/MS Peptide Identifications

Challenges in MS/MS

- Speed and Sensitivity
  - Keshishian et al, MCP 2007

- Measuring Changes
  - ICAT: Gygi et al, Nat. Biotech 1999
  - SILAC: Ong et al, MCP 2002
  - ITRAQ: Ross et al, MCP 2004
  - AQUA: Kirkpatrick, Methods 2005

- Post-translational Modifications
  - Bonanza: Falkner et al, JPR 2008
The Promise of MRM

- Works only on known targets
- More sensitive and faster than MS/MS
  - Better lower level of detection
  - Can monitor low m/z ions
  - Minimal fractionation required (10 min gradients!?)
- 5-10% CV and absolute quant
  - Heavy labeled peptides required
MRM: Triple Quadrupoles

- Better by design vs 3D trap
  - Holds more ions, 2x injection, 8x ion ejection
  - No low mass loss (aka 1/3\(^{rd}\) m/z issue)
- EPI mode filters **only** the ion of interest

## MRM: Known Targets

<table>
<thead>
<tr>
<th>Protein</th>
<th>Source</th>
<th>MW (kDa)</th>
<th>Signature Peptide</th>
<th>MH+ (mono)</th>
<th>z (Q1)</th>
<th>MRM Transitions&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Q1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Q3</td>
</tr>
<tr>
<td>Aprotinin</td>
<td>bovine lung</td>
<td>6.4</td>
<td>AGLC&lt;sub&gt;amc&lt;/sub&gt;QTFVYG&lt;sub&gt;GCamcR&lt;/sub&gt;&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1488.7</td>
<td>2</td>
<td>744.8, 858.3, <strong>959.4</strong>, 1087.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AGLC&lt;sub&gt;amc&lt;/sub&gt;QTF&lt;sup&gt;[13C&lt;sub&gt;5&lt;/sub&gt;]&lt;/sup&gt;VYG&lt;sub&gt;G&lt;/sub&gt;&lt;sub&gt;G&lt;/sub&gt;&lt;sub&gt;CamcR&lt;/sub&gt;&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1493.7</td>
<td>2</td>
<td>747.3, 863.4, <strong>964.4</strong>, 1092.5</td>
</tr>
<tr>
<td>Leptin</td>
<td>mouse</td>
<td>10</td>
<td>INDISHTQSVSAK</td>
<td>1399.7</td>
<td>3</td>
<td>467.2, 543.2, 586.8, <strong>643.8</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>INDISHTQ&lt;sup&gt;[13C&lt;sub&gt;5&lt;/sub&gt;]&lt;/sup&gt;VSAK</td>
<td>1404.7</td>
<td>3</td>
<td>468.9, 543.2, 589.3, <strong>646.3</strong></td>
</tr>
<tr>
<td>Myoglobin</td>
<td>Horse heart</td>
<td>17</td>
<td>LFTGHPETLEK</td>
<td>1271.7</td>
<td>3</td>
<td>424.6, 506.2, <strong>579.7</strong>, 716.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LFTGHPET&lt;sup&gt;[13C&lt;sub&gt;6&lt;/sub&gt;]&lt;/sup&gt;LEK</td>
<td>1277.7</td>
<td>3</td>
<td>426.6, 509.2, <strong>582.8</strong>, 722.4</td>
</tr>
<tr>
<td>Myelin Basic</td>
<td>bovine</td>
<td>18</td>
<td>HGFLPR</td>
<td>726.4</td>
<td>2</td>
<td>363.7, 385.2, 532.3, <strong>589.3</strong></td>
</tr>
<tr>
<td>Protein</td>
<td></td>
<td></td>
<td>HGF&lt;sup&gt;[13C&lt;sub&gt;6&lt;/sub&gt;]&lt;/sup&gt;LPR</td>
<td>732.4</td>
<td>2</td>
<td>366.7, 391.2, 538.3, <strong>595.3</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>YLASASTMDHAR</td>
<td>1322.6</td>
<td>3</td>
<td>441.5, 488.2, <strong>523.7</strong>, 817.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Y&lt;sup&gt;[13C&lt;sub&gt;6&lt;/sub&gt;]&lt;/sup&gt;LASASTMDHAR</td>
<td>1328.6</td>
<td>3</td>
<td>443.5, 488.2, <strong>523.7</strong>, 817.3</td>
</tr>
<tr>
<td>Prostate Specific</td>
<td>human</td>
<td>30</td>
<td>IVGGWEC&lt;sub&gt;amc&lt;/sub&gt;E&lt;sub&gt;K&lt;/sub&gt;&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1077.5</td>
<td>2</td>
<td>539.3, 808.3, <strong>865.3</strong>, 964.4</td>
</tr>
<tr>
<td>Antigen</td>
<td></td>
<td></td>
<td>I&lt;sup&gt;[13C&lt;sub&gt;5&lt;/sub&gt;]&lt;/sup&gt;VGGWEC&lt;sub&gt;amc&lt;/sub&gt;E&lt;sub&gt;K&lt;/sub&gt;&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1082.5</td>
<td>2</td>
<td>541.7, 808.3, <strong>865.3</strong>, 969.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LSEPAELTDAVK</td>
<td>1272.7</td>
<td>2</td>
<td>636.7, 775.4, 846.4, <strong>943.4</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LSEPAE&lt;sup&gt;[13C&lt;sub&gt;6&lt;/sub&gt;]&lt;/sup&gt;LTDAVK</td>
<td>1278.7</td>
<td>2</td>
<td>639.7, 781.3, 852.5, <strong>949.4</strong></td>
</tr>
<tr>
<td>Peroxidase</td>
<td>Horse radish</td>
<td>48</td>
<td>DTIVNELR</td>
<td>959.5</td>
<td>2</td>
<td>480.3, 531.2, <strong>630.3</strong>, 743.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DTIVNE&lt;sup&gt;[13C&lt;sub&gt;6&lt;/sub&gt;]&lt;/sup&gt;LR</td>
<td>965.5</td>
<td>2</td>
<td>483.3, 537.3, <strong>636.3</strong>, 749.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SSDLVALSGGHTFGK</td>
<td>1475.7</td>
<td>3</td>
<td>492.6, 703.3, <strong>790.3</strong>, 974.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SSDLVA&lt;sup&gt;[13C&lt;sub&gt;6&lt;/sub&gt;]&lt;/sup&gt;LSGGHTFGK</td>
<td>1481.8</td>
<td>3</td>
<td>494.6, 703.3, <strong>790.3</strong>, 980.5</td>
</tr>
</tbody>
</table>

Keshishian et al MCP 6.12 2007
MRM: Known Targets

Keshishian et al MCP 6.12 2007
MRM: Absolute Quant

- Plasma = most complex background
  - 2.5ng/ml at least for quant
  - CVs less than 15% and ~5% on average
- 1000-fold LOQ increase from IgY-12
Human Lens Proteome

Cataract development in lens may be caused by destabilizing post-translational modifications of crystallins. This live repository analyzes datasets and summarizes such modifications in healthy and cataractous human lens proteins.

About
- NEI Grant# 5R01EY007755-16
- SOSI / Cluster / Mr. M

Publications
- Wilmarth et al. 2006

Reference
- Protein List
- Peptide/Modification List

Key Figures (Supporting Data)

http://lens.singleorganism.com
Revising Analysis Strategies

• Initial Plan
  – MS/MS MudPIT of data
  – MS/MS quantification (O18 labeling)
  – Spectral count to compare proteins and mods

• Revised Plan
  – Reanalyze MS/MS survey data
  – MRM relative quant w/exogenous standard
  – Compare proteins and mods
MS/MS Data to MRM

- 28 samples of ~40 fractions = 1,000+ LC-MS/MS runs
- 330,000 spectra for cluster/p-mod analysis
  - ~4,000 unique peak lists
  - ~1,300 identified unique peak lists
- 3x transition per peptide; i.e. ~12,000 trans
  - Picked 400 targets to “scout” (1,200 trans)
    - 4 MRM methods w/100 targets each (300 trans per run)
  - Final MRM method with 300 targets (900 trans)
MS/MS Data to MRM

Cluster

330,000 MS/MS spectra w/identifications
QTOF + Sequest Analysis

4,000 Unique MS/MS spectra

2,700 Unknown
1,300 Identified

Mr. M

Hand picked 400 targets (1,200 transitions)

List A
List B
List C
List D

4 lists of 300 transitions each (unscheduled)

QTRAP transitions per block limit

QTRAP 4000

Scheduled MRM ~300 targets (900 trans)
~45 min per sample per LC-MRM
MS/MS Similarity Grouping

Falkner et al, J. Proteome Res., 2008, 7 (11)
Example “cluster”

![Diagram](image.png)

<table>
<thead>
<tr>
<th>ID</th>
<th>Peptide</th>
<th>Score</th>
<th>m/z</th>
<th>m/z Shift</th>
<th># m/z Shifts</th>
<th>RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>35115</td>
<td>IPADVDPLTITSSDSSDGVLTVNGPR</td>
<td>1.00</td>
<td>875.45</td>
<td>0.00</td>
<td>867</td>
<td>99.84</td>
</tr>
<tr>
<td>35218</td>
<td>YRIPADVDPLTITSSSDDGVLTVNGPR</td>
<td>0.52</td>
<td>981.84</td>
<td>106.39</td>
<td>748</td>
<td>93.73</td>
</tr>
<tr>
<td>35437</td>
<td>IPADVDPLTITSSSDDGVLTVNGPR</td>
<td>0.25</td>
<td>882.78</td>
<td>7.33</td>
<td>196</td>
<td>96.15</td>
</tr>
<tr>
<td>34420</td>
<td>IPADVDPLTITSSSDDGVLTVN(1)GPR</td>
<td>0.55</td>
<td>875.79</td>
<td>0.33</td>
<td>146</td>
<td>25.75</td>
</tr>
<tr>
<td>35867</td>
<td>IPADVDPLTITSSS(28)DGVLTVNGPR</td>
<td>0.71</td>
<td>884.79</td>
<td>9.34</td>
<td>97</td>
<td>104.60</td>
</tr>
<tr>
<td>34188</td>
<td>IPADVDPLTITSSSDDGVLTVN(1)GPRK</td>
<td>0.24</td>
<td>918.52</td>
<td>43.06</td>
<td>77</td>
<td>40.73</td>
</tr>
<tr>
<td>35583</td>
<td>IPADVDPLTITSSSDDGVLTV(-17)VNGPR</td>
<td>0.36</td>
<td>869.81</td>
<td>-5.64</td>
<td>36</td>
<td>30.00</td>
</tr>
</tbody>
</table>
Example “cluster”
Identified Peak Lists Before and After Cluster Analysis

Sequest + Decoy

Sequest + Decoy + Cluster
MRM AUC Sensitivity

Transitions To Review

Transition List Summary: 608 validated, 156 scheduled, 1074 candidate but not scheduled.
MRM Quant Variance

- 10-30% CV prior to correction
- 5-15% CV w/exogenous protein standard
MRM of Tricky P-mods

- +1 Da; too small to target
- Software can target using RP shift
Conclusion

- MS/MS works well for discovery
  - MudPIT takes a long time (weeks)
  - Build a library for MRM use
- MRM works well for monitoring
  - Monitor ~300 targets in <1 hour
  - Great sensitivity
  - Minimal sample variance
  - Minimal sample handling