**Overview**

**Purpose**: Demonstrate a generic, integrated workflow for untargeted metabolomics study using a UHPLC/benchtop Thermo Scientific™ Orbitrap™ mass spectrometer and informatics software.

**Introduction**

Metabolomics is a rapidly growing field of post-genomic biology, aiming to comprehensively characterize the small molecules in biological systems. Nonbiological systematic biases from instrument calibration or the order of sample injection account for the most significant errors in LC/TDF-MS data [1]. Here we present a workflow using a UHPLC/benchtop quadrupole Orbitrap platform and automated data analysis software for untargeted metabolic profiling of plasma samples for biomarker discovery from the Zucker diabetes fatty (ZDF) rat model. The optimal conditions for sample preparation, liquid chromatography (LC), column, mass spectrometry (MS), and data processing parameters are explored.

**Methods**

**Sample Preparation**

Plasma samples were deproteinized with organic solvent. Four extraction solvent systems including methanol (MeOH), acetonitrile (ACN), acetone, and 1:1:1 of MeOH/ACN/acetone were reconstituted in methano-water (19:1) containing isotopically labeled internal standard (IS), d5-hippuric acid for LC-MS analysis. Solvent blank, pooled QC, and biological samples were analyzed in a randomized injection order.

**Liquid Chromatography (or more generally Separations)**

UHPLC separation was implemented on a Thermo Scientific™ Dionex™ UltiMate™ 3000 HPG (high-pressure gradient) pump using Thermo Scientific™ HyperSil GOLD RP C18 column at 450 μL/min column temperature at 55 °C. LC solvents were 0.1% FA (A) and 0.1% FA in MeOH (B). Apply linear gradient from 0.5–50% B for 5.5 min, followed by increasing to 98% at 11 min, hold 98% B for 6 min, then decrease to 0.5% at 13 min, then equilibrate for another 2 min.

**Mass Spectrometry**

The Thermo Scientific™ Q Exactive™ mass spectrometer was operated under electrospray ionization (ESI) positive, and polarity switching modes. Full scan (m/z 67–980) used resolution 70,000 with automatic gain control (AGC) target of 1×106 ions and a maximum ion injection time (IT) of 35 ms. Data-dependent MS/MS were acquired on a “TopN” data-dependent mode using the following parameters: resolution 17,500, AGC 1×105 ions; maximum IT 80 ms; 2.0 amu isolation window; normalized collision energy 35% ± 50% underfill into 1.2% ASAP trigger 2–4 s, and dynamic exclusion 6 s. Source ionization parameters were: spray voltage, 3.8 kV; capillary temperature, 300 °C; and S-Lens level, 45.

**Data Analysis**

Differential analyses of the obese versus lean plasma were performed using Thermo Scientific™ SIEVE™ 2.1 software which executes background subtraction, component detection, peak alignment, differential analysis (Figure 1). Statistical results, putative IDs, and pathways were generated after searching ChemSpider and KEGG™. Metabolites of interest were identified via MS/MS mass spectral database matching. The raw files were converted to mzCloud format using ProteoWizard and also analyzed by XCMS Online [2] to compare the results.

**FIGURE 1. Untargeted metabolomics workflow**

**Results**

**Challenges in Untargeted Metabolomics Study**

- Complexity of biological samples
- Diversity of small molecule metabolites: polar and non-polar analytes
- Ionization requires both positive and negative ion
- Wide range of concentrations
- No universal method for chromatographic separation
- Multiple sources of variability
- Structure elucidation of unknowns is expensive: lack of synthetic standards

**Preparing for the UHPLC-MS Data Acquisition**

Prior to the real samples analysis, a solvent blank with internal standard (IS) is injected at the beginning to check the solvent and the LC-MS status. The injections of the real samples should be randomized in order to eliminate systematic bias. Tripleptide injections of the pooled plasma are intermittently repeated throughout the whole batch to validate consistent performance of the overall system.

**FIGURE 2. UHPLC/MS Experimental Design and Run Sequence.** Left, schematic showing the vials and sample names. Right, detailed content and overall time for each step.

**UHPLC provides fast chromatography for high throughput analysis, the typical peak width is 4–6 seconds.** Our method can baseline resolve Isoleucine and Leucine, generating peak width 1.2 s at FWHM. Refer to Figure 4.

**TABLE 2. Representative metabolites that are significantly changed**

<table>
<thead>
<tr>
<th>Compound</th>
<th>Formula</th>
<th>m/z</th>
<th>Concn. (ppm)</th>
<th>P-value</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normetanephrine</td>
<td>C9H13NO3</td>
<td>170.081</td>
<td>2.03</td>
<td>3.44E-04</td>
<td>0.0</td>
</tr>
<tr>
<td>Methylnoradrenaline</td>
<td>C8H11NO3</td>
<td>170.081</td>
<td>2.03</td>
<td>3.44E-04</td>
<td>0.0</td>
</tr>
<tr>
<td>LysoPC(20:3)</td>
<td>C28H52NO7P</td>
<td>546.3525</td>
<td>5.84</td>
<td>9.23E-03</td>
<td>4.0</td>
</tr>
<tr>
<td>LacCer(d18:1/14:0)</td>
<td>C44H83NO13</td>
<td>834.5936</td>
<td>8.10</td>
<td>7.13E-04</td>
<td>0.0</td>
</tr>
<tr>
<td>PC35:4</td>
<td>C43H78NO8P</td>
<td>768.5498</td>
<td>8.01</td>
<td>3.17E-04</td>
<td>4.0</td>
</tr>
<tr>
<td>PC 38:3</td>
<td>C46H86NO8P</td>
<td>812.6095</td>
<td>8.58</td>
<td>3.08E-02</td>
<td>4.0</td>
</tr>
<tr>
<td>PC38:2</td>
<td>C46H88NO8P</td>
<td>836.6133</td>
<td>8.55</td>
<td>2.22E-02</td>
<td>0.0</td>
</tr>
<tr>
<td>PC44:10</td>
<td>C50H80NO8P</td>
<td>854.5657</td>
<td>7.25</td>
<td>3.65E-02</td>
<td>4.0</td>
</tr>
</tbody>
</table>

**Conclusion**

An efficient and robust workflow for untargeted metabolomics is presented here. The reliable high-resolution, accurate-mass (HR/AM) performance of the Q Exactive LC-MS system eliminates the need for technical replicates on biological samples. The superior S/N in Orbitrap data allows efficient data reduction in SIEVE 2.1 software, resulting in much reduced data analysis effort. KEGG pathway visualization allows quick access to biological pathway mapping. The MS+ spectral library mzCloud facilitates accurate compound identification.

**References (if necessary)**