Can data-independent acquisition mass spectrometry be used confidently for metaproteomics?

Pratik Dilip Jagtap

Annaliese C.S. Meyer
Andrew T Rajczewski
Alfredo Blakeley-Ruiz
Matthew McIlvin
Tim Van Den Bossche
Timothy J Griffin
Brian C. Searle
Makoto Saito
Manuel Kleiner

Woods Hole Oceanographic Institution, Woods Hole, MA
University of Minnesota, Minneapolis, MN
North Carolina State University, Raleigh, NC
Ghent University, Ghent, Belgium
The Ohio State University College of Medicine, Columbus, OH
Can data-independent acquisition mass spectrometry be used confidently for metaproteomics?

- MICROBIOME RESEARCH – Clinical and Environmental Research.

- METAPROTEOMICS – Mass Spectrometry-based Microbiome Analysis.
  – Detects functions along with taxonomy within the microbiome.

- QUANTITATIVE METAPROTEOMICS – The need to quantitate functions expressed along with the taxonomy.

https://www.creative-proteomics.com/
MICROBIOME MEASUREMENTS

• Microbiome measurements can be carried out using metagenomics (taxonomy), metatranscriptomics (taxonomy and RNA expression) and metaproteomics (taxonomy and protein expression).

• For quantitative metaproteomics, various methods have been used:
  - Label-Free
  - Isobaric-labeled
  - DDA
  - Spectral counts
  - DIA
  - Precursor-intensity

Challenges in Quantitative Metaproteomics

• Need for quantitation of taxonomy and functions expressed by microbiomes.
• Shared proteins amongst microbiome taxonomic groups.
• While many taxonomic members of microbiomes are known, there are some taxa that have not been detected, yet.
• Functional unknowns: Almost 40% of genes are without a match in functional databases.
DIA-MS has better reproducibility across replicates.

- **57.7% overlap**
- **96.7% overlap**

DIA-MS offers deeper taxonomy coverage.

DIA-MS offers deeper functional coverage.

What level of accuracy does DIA-MS offer?

Susan T. Weintraub

Brook L. Nunn
**METAPROTEOMICS DATASETS**

Pure cultures of 1 Archaeon, 1 Eukaryote, 5 bacteriophages, 3 gr+ Bacteria and 22 gr− Bacteria

Mix 3 community types with 28 to 32 of the species/strains

4 biological replicates for each community type

**Table:**

<table>
<thead>
<tr>
<th>Cell number</th>
<th>Protein amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>1x</td>
<td>31x</td>
</tr>
<tr>
<td>354x</td>
<td>229x</td>
</tr>
</tbody>
</table>

Maximum fold differences between individual bacterial community members

Kleiner et al., Nature Communications Vol. 8, 1558 (2017)
METAPROTEOMICS QUANTITATION WORKFLOW USING DDA & DIA

Protein digestion of microbial-mix sample (EQUAL CELL & EQUAL PROTEIN & UNEVEN)

NC STATE UNIVERSITY

4X

Proteins

Peptides

PSMs

DDA-MS

DIA-MS

WHOI WORKFLOW

Scaffold

Fusion

Scaffold-DIA

DDA-MS

DIA-MS

UMN WORKFLOW

FragPipe

EncyclopeDIA

DDA-MS

DIA-MS

Matthew McIlvin

Andrew Rajczewski
PEPTIDES DETECTION

EQUAL PROTEIN | DDA | UMN | PEPTIDES

- 4673 (34.7%)

EQUAL PROTEIN | DIA | UMN | PEPTIDES

- 10859 (92.53%)

EQUAL PROTEIN | DDA | WHOI | PEPTIDES

- 9203 (50.66%)

EQUAL PROTEIN | DIA | WHOI | PEPTIDES

- 22692 (90.42%)
PROTEINS DETECTION

EQUAL PROTEIN | DDA | UMN | PEPTIDES

4673 (34.7%)

EQUAL PROTEIN | DIA | UMN | PEPTIDES

10859 (92.53%)

EQUAL PROTEIN | DDA | UMN | PROTEINS

2527 (52.98%)

EQUAL PROTEIN | DIA | UMN | PROTEINS

3942 (93.88%)

WORKFLOW
Most of the taxonomic members are detected by both UMN and WHOI workflows.

DIA detected previously undetected viral proteins (such as phage ES18).
Proteins Assigned to Taxa

DIA Protein Groups

DDA Protein Groups

Taxon

Protein Groups/Taxon

WORKFLOW
Protein Measurements From Low Abundant Phages

**PHAGE DIA INTENSITIES**

- Phage P22
- Phage F0
- Phage F2
- Phage ES18

**PHAGE DDA SPECTRAL COUNTS**

- Phage P22
- Phage F0
- Phage F2
- Phage ES18

**PHAGE DDA PRECURSOR INTENSITY**

- Phage P22
- Phage F0
- Phage F2
- Phage ES18
**Coefficient of Variation**

- **Cupriavius metallireducens** (High abundance taxon)
- **Bacillus subtilis** (Medium abundance taxon)
- **Staphylococcus aureus** (Low abundance taxon)

**WHOI WORKFLOW**

**UMN WORKFLOW**
SUMMARY & FUTURE PLANS

• DIA-MS analysis using the ground-truth dataset, offers reproducible data as compared to DDA-MS datasets at both peptide and protein levels.
• Taxonomy analysis using the ground-truth data showed that DIA-MS offers deeper and reproducible proteome coverage for taxa at high, intermediate and low-levels. Low abundant phage proteins were detected by DIA-MS.
• DIA-MS measured the taxonomic groups at a higher accuracy than DDA-MS; especially when compared with intensity measurements.
• Less than 20% CV was observed for DIA measurements of selected taxa (at various levels of abundance).
• DIA-MS offers a reproducible, accurate method for both taxonomy and functional measurements in metaproteomics research.
• We plan to acquire MS data from instruments with higher sensitivity and assess the accuracy, reproducibility and depth of measurements.
ACKNOWLEDGMENTS