**INTRODUCTION**

- Integration of different ‘omics’ data reveals novel discoveries into biological systems.
- However, the need for multiple, disparate software makes the integration of multi-omic data a serious challenge.

**METHODS & DATABASES**

RAW files from multiple datasets (see below) were generated from Orbitrap Velos instrument. The processed peak lists were searched using ProteinPilot® v4.5 (AB Sciex) within Galaxy-P (described in Section 2.3). After optimization and testing, workflows were used in a sequential manner to generate workflows that can be reused, shared, and creatively modified for multiple studies.

**METAPROTEOMICS**

**SECC and Salivary datasets**
- **RESULTS SUMMARY**
  - **Dataset**
    - Human Salivary
    - SECC with sucrose
    - SECC without sucrose
  - **Number of spectra**
    - HSA/Salivary: 985,974
    - SECC sucrose: 137,970
    - SECC no sucrose: 22,269
  - **Distinct peptides**
    - HSA/Salivary: 1926
    - SECC sucrose: 26,126
    - SECC no sucrose: 5
  - **Peptide sets**
    - HSA/Salivary: 12
    - SECC sucrose: 65
    - SECC no sucrose: 133
  - **Species**
    - HSA/Salivary: 65
    - SECC sucrose: 125
    - SECC no sucrose: 125
  - **Notes**
    - HSA/Salivary data was analyzed using MEGAN.

**PROTEOGENOMICS**

**COPD dataset**
- **RESULTS SUMMARY**
  - **Dataset**
    - COPD
  - **Number of spectra**
    - COPD: 985,974
  - **Novel protein peptides after PSI-M0 evaluation**
    - COPD: 254
  - **Number of distinct peptides after PSI-M0 evaluation and for genomic context**
    - COPD: 52
  - **Dataset**
    - COPD: 985,974
  - **Notes**
    - Representation of genomic organization of identified novel protein-specific peptides from PRB1 and PRB2 genes on chromosome 12.
    - Human chromosome 12

**CONCLUDING REMARKS**

- **Saliary ProteoGenomics**: 52 novel protein peptides identified (Notably - PRB1 and PRB2) for COPD, PRB1 and PRB2 peptides.
- SECC Metaproteomics: Analysis of outputs from Galaxy-P workflows and MEGAN analysis is currently ongoing for three replicates.
  - Quantitative proteomics: Workflows have been used on multiple replicates. Reproducibility analysis and RNASeq data integration in works.
- **IMMEDIATE PLANS**
  - Working along with genomics research community and Genomics Center to develop on integration of RNASeq derived workflows for database generation.
  - Working closely with metagenomics / microbiome research community to develop functional pathway analytical workflows using the metaproteomics data.
  - Working on correlating RNASeq quantitative information with quantitative iTRAQ proteomic information for both model and non-model organisms.
- **FUTURE PLANS**
  - Installation and testing of open-source tools (such as MS-GF+ and PeptideShaker). The installation and testing is being carried out through international collaboration between developers and users.
  - Improving on current metaproteomic and proteogenomic workflows.
  - Acknowledgements: This work was funded by NSF 1147907. Also many thanks to John Chilton (Penn State) for his work in the first year of the project.

**WORKFLOW AND TUTORIAL**

**PROTEOGENOMICS**

**Salivary and OPML datasets**
- **RESULTS SUMMARY**
  - **Dataset**
    - Salivary and OPML
  - **Number of spectra**
    - Salivary: 157,299
    - OPML: 988,974
  - **Novel protein peptides**
    - Salivary: 156,405
    - OPML: 2,528
  - **Number of distinct peptides after PSI-M0 evaluation and for genomic context**
    - Salivary: 21
    - OPML: 52
  - **Dataset**
    - Salivary: 157,299
    - OPML: 988,974
  - **Notes**
    - Representation of genomic organization of identified novel protein-specific peptides from PRB1 and PRB2 genes on chromosome 12.
    - Human chromosome 12

**Qualitative Proteomics**

- **Chronic Obstructive Pulmonary Disease (COPD)**: Linked lung cancer tissue samples were collected and subjected to iTRAQ labeling and 2D IEF-MS. The dataset was searched against Human Uniprot database.

**Quantitative Proteomics**

- **Distinct peptides**
  - **Dataset**
    - COPD: 985,974
  - **Novel protein peptides**
    - COPD: 156,405
  - **Number of distinct peptides after PSI-M0 evaluation and for genomic context**
    - COPD: 21
- **Dataset**
  - **Notes**
    - Representation of genomic organization of identified novel protein-specific peptides from PRB1 and PRB2 genes on chromosome 12.