PROTEOGENOMICS & METAPROTEOMICS USING THE GALAXY PLATFORM

PRATIK JAGTAP
University of Minnesota
St.Paul / Minneapolis
Minnesota, USA
MINNEAPOLIS TO NEW DELHI
MINNEAPOLIS, MINNESOTA
The Center for Mass Spectrometry and Proteomics (CMSP) facility provides support to the life sciences research community at the University of Minnesota. The CMSP - a subunit of the Department of Biochemistry, Molecular Biology and Biophysics (BMBB) - is based on both campuses. CMSP provides expertise, equipment, training, and support for analyzing all types of biological molecules - with a focus on proteomic and metabolomic applications.

With years of experience serving the entire biological science area of the University, the CMSP staff has a wide range of expertise in analytical methods used to characterize

http://cbs.umn.edu/cmssp/home
PROTEOGENOMICS & METAPROTEOMICS USING THE GALAXY PLATFORM

• PROTEOMICS OVERVIEW

• PROTEOGENOMICS

• CHALLENGES

• GALAXYP

• BIOLOGICAL INSIGHTS

• METAPROTEOMICS

• LINKS AND ACKNOWLEDGMENTS
PROTEOMICS WORKFLOW

Peaklist Generation

Database Search

Peptide Spectral Match: Statistical Validation

Protein Inference

Eng et al 2011
Mol Cell Proteomics. 10(11): R111.009522.
DEFINING PROTEOMICS: LOOKING WITHIN

Mass spectrum

Reference Protein Database from genomic annotation

Peptide Spectral Match
PROTEOMICS WORKFLOW

Eng et al 2011
Mol Cell Proteomics. 10(11): R111.009522.

Peaklist Generation

Search Database

Database Search

Peptide Spectral Match : Statistical Validation

Protein Inference
PROTEOGENOMICS & METAPROTEOMICS USING THE GALAXY PLATFORM

• PROTEOMICS OVERVIEW

• PROTEOGENOMICS

• CHALLENGES

• BIOLOGICAL INSIGHTS

• METAPROTEOMICS

• GALAXYP

• LINKS AND ACKNOWLEDGMENTS
DEFINING PROTEOGENOMICS: LOOKING WITHIN AND WITHOUT

- Mass spectrum
- Reference Protein Database from genomic annotation
- RNASeq data
- Genome six-frame translation
- cDNA three-frame translation
PROTEOGENOMICS

- Genome annotation
- Gene expression regulation
- Protein variants in disease
- Functional outcomes of genome mutation

PROTEOGENOMICS & METAPROTEOMICS USING THE GALAXY PLATFORM

- PROTEOMICS OVERVIEW
- PROTEOGENOMICS
- CHALLENGES
- BIOLOGICAL INSIGHTS
- METAPROTEOMICS
- GALAXYP
- LINKS AND ACKNOWLEDGMENTS
PROTEOGENOMICS : BIOINFORMATIC CHALLENGES

- Large database sizes (*6-frame and 3-frame translation and metagenomic databases*).
- False-positive sources and their elimination.
- Validation of the peptide identification. (*Search using BLAST-P*)
- PSM Evaluation / Targeted proteomics of identified peptides.
- Genomic localization.
- Disparate tools and numerous processing steps.
PROTEOGENOMICS & METAPROTEOMICS USING THE GALAXY PLATFORM

• PROTEOMICS OVERVIEW

• PROTEOGENOMICS

• CHALLENGES

• GALAXYP

• BIOLOGICAL INSIGHTS

• METAPROTEOMICS

• LINKS AND ACKNOWLEDGMENTS
Benefits of Galaxy

- A web-based bioinformatics data analysis platform.
- Software *accessibility* and usability.
- *Share-ability* of tools, workflows and histories.
- *Reproducibility* and ability to test and compare results after using multiple parameters.
- *Software tools* can be used in a sequential manner to generate *analytical workflows* that can be reused, shared and creatively modified for multiple studies.

Galaxy-P: A new community-based informatics paradigm for MS-based proteomics

- Funded via the NSF Advances in Biological Informatics program
- 3 years of funding; effective July 15, 2012-June 30, 2015

Grant objective in a nutshell:
We propose to extend the Galaxy framework for genomics by deploying and integrating a series of key software programs for MS-based proteomics data analysis, thus creating Galaxy Tool Modules for Proteomics which we refer to as Galaxy-P

Project-based strategy for Galaxy-P development: Collaborate with biological researchers with “real” projects to guide developments.
GALAXY-P: IMPLEMENTATION OF PROTEOMICS TOOLS WITHIN GALAXY ENVIRONMENT.

Welcome to GalaxyP

GalaxyP is a multi-omics data analysis platform with particular emphasis on mass spectrometry based proteomics. GalaxyP is developed at the University of Minnesota, deployed at the Minnesota Supercomputing Institute, and is an extension of the popular Galaxy project. The GalaxyP project is supported by a grant from NSF.

This public Galaxy instance is meant for testing with small-scale data sets, and sharing workflows and tools. Users with larger data analysis needs are encouraged to install a local instance of Galaxy and access GalaxyP tools via the Tool Shed.

Updates

Sept. 11, 2015
ProteinPilot is available.

February 12, 2015
All Galaxy Tools running normally

Links

Large-scale multi-omics data integration and analysis: challenges and opportunities

Metaproteomics: an opportunity-rich complement to metagenomics
Software tools can be used in a sequential manner to generate analytical workflows that can be reused, shared and creatively modified for multiple studies.

For example, Protein Database Downloader downloads UniProt protein FASTA databases of various organisms.
GALAXY-P: IMPLEMENTATION OF PROTEOMICS TOOLS WITHIN GALAXY ENVIRONMENT.
PROTEOGENOMICS: STEPS INVOLVED

~ 2 million proteins
- Database Generation
- Peaklist generation
- Database search
- First-step
- Two-step
- Identifying peptides from translated nucleotide db
- Automated BLAST-P search

~ 10,000 proteins
- Peptide-Spectral-Match Evaluation

~ 5,000 proteins
- Genomic context analysis

~ 1,000 peptides
- 100 peptides
- 50 peptides
“Using Galaxy-P to leverage RNA-Seq for the discovery of novel protein variations.”
Galaxy-P provides an integrated platform for every step of proteogenomic analysis.

- Build target database – download and translate EST databases or perform gene prediction with Augustus.
- Numerous tools for identification and text manipulation.
- Workflow utilizing BLAST to identify novel peptides.
- Tool to assess peptide-spectrum matches and visualize spectra.
- Visualize identified peptides on the genome.
- 140 steps: Seamless, integrated proteogenomic workflow.

Flexible and accessible workflows for improved proteogenomic analysis using Galaxy framework.

J. Proteome Res. (2014) DOI: 10.1021/pr500812t
Link: z.umn.edu/pgfirstlook
## PSM EVALUATION

<table>
<thead>
<tr>
<th>MascotIdentities threshold</th>
<th>MascotScore</th>
<th>ScaffoldPeptide Probability</th>
<th>acquisitionNum</th>
<th>msLevel</th>
<th>polarity</th>
<th>peaksCount</th>
<th>sequence</th>
<th>precursorMZ</th>
<th>precursorCharge</th>
<th>retentionTime</th>
</tr>
</thead>
<tbody>
<tr>
<td>39.325017</td>
<td>33.32</td>
<td>0.95</td>
<td>26957</td>
<td>2</td>
<td></td>
<td>300</td>
<td>AQFEGIVTDLLRR</td>
<td>506.723600001939905</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>40.910688</td>
<td>41.64</td>
<td>0.95</td>
<td>87885</td>
<td>2</td>
<td></td>
<td>300</td>
<td>SQYFSTAAADGTFUVEIK</td>
<td>904.7714000290986</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>40.918777</td>
<td>29.78</td>
<td>0.8760495</td>
<td>32450</td>
<td>2</td>
<td></td>
<td>300</td>
<td>MKETAENLYGHTAK</td>
<td>531.4867000193991</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>41.335624</td>
<td>89.47</td>
<td>0.95</td>
<td>86558</td>
<td>2</td>
<td></td>
<td>300</td>
<td>STNGDTFLGGEFDQALLR</td>
<td>1028.3580000290983</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>41.48911</td>
<td>70.94</td>
<td>0.95</td>
<td>63125</td>
<td>2</td>
<td></td>
<td>300</td>
<td>AQFEGIVTDLLRR</td>
<td>682.0720000290983</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>41.729988</td>
<td>105.71</td>
<td>0.95</td>
<td>93899</td>
<td>2</td>
<td></td>
<td>300</td>
<td>STNGDTFLGGEFDQALLR</td>
<td>1028.8940000290984</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>41.81472</td>
<td>29.33</td>
<td>0.6915607</td>
<td>11016</td>
<td>2</td>
<td></td>
<td>300</td>
<td>LVQMPAKR</td>
<td>436.2794000290985</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>41.81279</td>
<td>27.29</td>
<td>0.5512762</td>
<td>11006</td>
<td>2</td>
<td></td>
<td>300</td>
<td>LVQMPAKR</td>
<td>436.2738000290985</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Page 6 of 247 Showing records of 12313 total records.

**STNGDTFLGGEFDQALLR 85558**

**Peptide Spectral Match Evaluation**

**Spectral Visualization**

**Filtering of Peptide Spectral Matching Metrics**
GENOME VISUALIZATION USING IGV BROWSER

3-frame translated peptide sequences
Reference peptide sequence
Identified novel proteoform peptide sequence

General Transfer Format file for Human genome

Peptide to GTF conversion
Visualization in genomic context

Peptide Summary of new proteoform peptides with quality peptide spectral matching characteristics.
PROTEOGENOMICS & METAPROTEOMICS USING THE GALAXY PLATFORM

• PROTEOMICS OVERVIEW

• PROTEOGENOMICS

• CHALLENGES

• GALAXYP

• BIOLOGICAL INSIGHTS

• METAPROTEOMICS

• LINKS AND ACKNOWLEDGMENTS
The dataset was searched against FASTA database with human proteins, contaminant proteins, 3-frame translated cDNA database from EnSEMBL and Human Oral Microbiome database (HOMD).

Table 1. Summary of Genomic Organization of Peptides Corresponding to Novel Proteoforms

<table>
<thead>
<tr>
<th>genomic rearrangements</th>
<th>peptides</th>
<th>chromosome location(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>alternate frame</td>
<td>26</td>
<td>1,3,5,7, 8, 9, 11, 12, 14, 16, and 19</td>
</tr>
<tr>
<td>untranslated region</td>
<td>15</td>
<td>2, 4, 6, 7, 8, 11, 12, 13, 14, and 19</td>
</tr>
<tr>
<td>pseudogenes</td>
<td>6</td>
<td>1, 3, 6, 14, 19, and X</td>
</tr>
<tr>
<td>intronic region</td>
<td>2</td>
<td>12 and 16</td>
</tr>
<tr>
<td>novel exon junctions</td>
<td>2</td>
<td>15 and 17</td>
</tr>
<tr>
<td>antisense</td>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

*J Proteome Res. (2014) 13(12):5898-908. doi: 10.1021/pr500812t*
Figure 4. Representation of organization of identified peptides corresponding to a novel proteoform from PRB1 and PRB2 genes on chromosome 12. View is a zoomed-in screenshot of chromosome 12, which shows the orientation of expression, amino acid sequences within three frames of translation, reference files in the tracks, and amino acid sequence of the identified peptide corresponding to a novel proteoform. The red arrows indicate the direction and amino acid sequence (from amino-terminal to carboxy-terminal) of the identified peptides. A red asterisk indicates a stop codon in the normal coding frame. Block arrows in red indicate multiple distinct peptides identified during the proteogenomic analysis.
Galaxy-P provides an integrated platform for every step of proteogenomic analysis.

- Build target database – download and translate EST databases.
- Numerous tools for identification and text manipulation.
- Workflow utilizing BLAST to identify novel peptides.
- Tool to assess peptide-spectrum matches and visualize spectra.
- Visualize identified peptides on the genome.
- 140 steps: Seamless, integrated proteogenomic workflow.

Flexible and accessible workflows for improved proteogenomic analysis using Galaxy framework.

J. Proteome Res. (2014) DOI: 10.1021/pr500812t
Link: z.umn.edu/pgfirstlook
Tracing of core body temperature ($T_b$, black line) from a single animal measured by a surgically implanted transmitter, along with the controlled ambient temperature (blue line) over the course of the hibernation season.

*TOR (Torpor), J-IBA (January IBA), M-IBA (March IBA)*
The datasets were run in triplicates and were searched against proteomic dataset from RNASeq data.

Differentially expressed genes from RNASeq data and differentially expressed proteins from iTRAQ data were compared.

Functional analysis of differentially expressed proteins revealed that:
- Protein expression in hibernation relative to AUGUST highlights fatty acid metabolism and altered calcium handling and contractile function in the heart.

162 novel peptide sequences were identified in all three replicates.
PROTEOGENOMICS & METAPROTEOMICS
USING THE GALAXY PLATFORM

• PROTEOMICS OVERVIEW

• PROTEOGENOMICS

• CHALLENGES

• GALAXYP

• BIOLOGICAL INSIGHTS

• METAPROTEOMICS

• LINKS AND ACKNOWLEDGMENTS
“The large-scale characterization of the entire protein complement of environmental microbiota at a given point in time”


“Through the application of metaproteomics to different microbial consortia over the past decade, we have learnt much about key functional traits in the various environmental settings where they occur.”

DEFINING METAPROTEOMICS:
LOOKING WITHIN AND WITHOUT

Mass spectrum

Reference Protein Database
from genomic annotation

RNASeq data

Genome six-frame translation

cDNA three-frame translation

Metagenomic sequences
DEFINING METAPROTEOMICS

Mass spectrum

Database generation

Peaklist generation

Database search

First-step

Two-step

Identifying peptides from translated nucleotide db

Automated BLAST-P search

metagenomic sequences
MATCHED PAIR OF CONTROL VERSUS LESION

Oral premalignant lesion (OPML) versus control

Digested O/N with trypsin

SCX Fractionation and and LC-MS (7 Fractions)

Thermofinnigan Orbitrap (Orbi MS, MS/MS LTQ)

7 RAW Files each

The dataset was searched against FASTA database with human proteins, contaminant proteins, 3-frame translated cDNA database from EnSEMBL and Human Oral Microbiome database (HOMD).

Figure 4. Comparison of genera identified and SEED functional groups within control (magenta) and OPML (tan) samples by using MEGAN5 analysis. (A) Radial plot showing the relative distribution of genus-level peptides. The width of the segment assigned to each genus is proportional to the relative abundance of peptides for that genus, and the heights of the magenta and tan spokes within each segment are proportional to the relative abundance of peptides for that genus assigned to the Control and OPML groups. (B) Radial plot showing the relative distribution of peptides into SEED functional roles. The width of the segment assigned to each functional role is proportional to the relative number of identified peptides for that category, and the heights of the magenta and tan spokes within each segment are proportional to the relative abundance of peptides assigned to the Control and OPML groups.
METAPROTEOMICS OF CHILDHOOD CARIES

- *In vitro* investigation of sucrose-induced changes in the metaproteomes of children with caries.

- Major shifts in taxonomy and function in paired microcosm oral biofilms grown without and with sucrose respectively.

  - Twelve replicates have been analyzed.

- SEED analysis of Oral microcosm biofilms showed characteristic NS and WS patterns of protein expression that were highly conserved across taxonomically diverse communities.

- Targeted proteomic approaches then can be used to determine whether those proteins are also expressed when plaque is exposed to sucrose in the mouth.
PROTEOGENOMICS & METAPROTEOMICS USING THE GALAXY PLATFORM

• PROTEOMICS OVERVIEW

• PROTEOGENOMICS

• CHALLENGES

• GALAXYP

• BIOLOGICAL INSIGHTS

• METAPROTEOMICS

• LINKS AND ACKNOWLEDGMENTS
NSF award 1458524 “A unified Galaxy-based platform for multi-omic data analysis and informatics”. PI: Tim Griffin
Award dates: 9/1/15-8/31/18

• Enhance the Galaxy environment with **new interactive visualization tools** and data exchange functionalities necessary for effective multi-omic data analysis.

• Extend the Galaxy environment to analyze and process diverse **metabolomics** data and support workflows for metabolic activity profiling.

• Extend the Galaxy environment for integrative genomic-proteomic data analysis supporting **proteogenomic and metaproteomic applications**.

• Promote usage of Galaxy for multi-omics by the research community and provide undergraduate training opportunities in computational systems biology via a local area institutional network.
GALAXYP: ONGOING PROJECTS

COMMUNITY BASED SOFTWARE DEVELOPMENT

- Community-based software development model should prove effective for future implementation, testing and continued improvement of command-line driven software tools.

REPERTOIRE OF WORKFLOWS

- Sharing of analytical workflows that can be reused, shared and creatively modified for multiple studies.
- Multiple workflows for metaproteomics, quantitative proteomics, proteogenomics, RNASeq workflows, are being developed, shared and used.

<table>
<thead>
<tr>
<th>WORKFLOW</th>
<th>INPUT</th>
<th>TOOLS</th>
<th>OUTPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Peaklist generation</td>
<td>RAW File, cDNA database, Protein FASTA files.</td>
<td>mmass, MFQ, Pismo</td>
<td>Merged Protein FASTA file</td>
</tr>
<tr>
<td>2 Database Generation</td>
<td>MGF Files, Search database.</td>
<td>group file, peptide summary and PSIPRED</td>
<td></td>
</tr>
<tr>
<td>3 Database Search by Two-Step Method</td>
<td></td>
<td>FDR report</td>
<td></td>
</tr>
<tr>
<td>4 Peptide list with sequence</td>
<td>Peptide Summary</td>
<td>PSM Evaluation file</td>
<td></td>
</tr>
<tr>
<td>5 BLAST-P Analysis</td>
<td>Peptide list with accession numbers within</td>
<td>PSM Evaluation metric and HTML Links.</td>
<td></td>
</tr>
<tr>
<td>6 Peptide Spectral Match</td>
<td>Peptide Summary, mrm files</td>
<td>PSM Evaluation score and HTML Links.</td>
<td></td>
</tr>
<tr>
<td>7 Peptide to GTF conversion</td>
<td>Peptide Summary, cDNA database, GTF file.</td>
<td>Peptides to GTF</td>
<td></td>
</tr>
</tbody>
</table>
COMMUNITY-BASED SOFTWARE DEVELOPMENT

Software Developers

SearchGUI / PeptideShaker

Galaxy Wrapper

Improvements to the software tool

Software tool deposited in Galaxy Toolshed

GalaxyP

USER FORUM / GITHUB

Users test the tools and provide feedback to developers.

Software tool installed in GalaxyP

Users test the tools and provide feedback to developers.
PROTEOMICS TOOLS ACCESSIBLE VIA THE TOOL SHED

https://toolshed.g2.bx.psu.edu/

**Repositories in Category Proteomics**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Type</th>
<th>Metadata Revisions</th>
<th>Tools or Package Verified</th>
<th>Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>appendfor</td>
<td>Add false discovery rate to tabular data.</td>
<td>Unrestricted</td>
<td>0 (2013-05-10)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>blast plus remote blastp</td>
<td>NCBI BLAST+ remote blastp</td>
<td>Unrestricted</td>
<td>4 (2015-05-04)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>blastxml to tabular selectable</td>
<td>Converts blast xml file to a tabular with options for unmatched queries, and number of hits to convert</td>
<td>Unrestricted</td>
<td>1 (2014-10-08)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>deteXt</td>
<td>This tool allows users to download protein databases from common sources.</td>
<td>Unrestricted</td>
<td>4 (2014-09-26)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>decoyfasta</td>
<td>Galaxy tool wrapper for the transproteomic pipeline decoyFASTA tool.</td>
<td>Unrestricted</td>
<td>6 (2014-10-08)</td>
<td>n/a</td>
<td>galaxy</td>
</tr>
<tr>
<td>directx Laptagacon</td>
<td>Bumbershoot DirectX and Tagacon</td>
<td>Unrestricted</td>
<td>0 (2014-09-26)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>fasta merge files and filter unique sequences</td>
<td>Merge FASTA files, keeping only unique sequences</td>
<td>Unrestricted</td>
<td>0 (2014-09-26)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>feature alignment</td>
<td>Feature Alignment of peakgroups below a FDR</td>
<td>Unrestricted</td>
<td>n/a</td>
<td>n/a</td>
<td>galaxy</td>
</tr>
<tr>
<td>filter by fasta ids</td>
<td>Extract sequences from a FASTA file based on a list of IDs</td>
<td>Unrestricted</td>
<td>0 (2014-09-26)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>gams msms analysis</td>
<td>GCMS and LCMS workflows</td>
<td>Unrestricted</td>
<td>0 (2015-05-15)</td>
<td>n/a</td>
<td>proteomic</td>
</tr>
<tr>
<td>ideconvor</td>
<td>Bumbershoot idbConvor, a part of Bumbershoot IDPikker.</td>
<td>Unrestricted</td>
<td>2 (2014-09-30)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>iQuant dj</td>
<td>iQuant is a tool that performs tags based isotopic quantification</td>
<td>Unrestricted</td>
<td>0 (2014-09-26)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>make protein decy</td>
<td>Generate a decoy database from an input set of protein sequences</td>
<td>Unrestricted</td>
<td>1 (2015-03-26)</td>
<td>n/a</td>
<td>kascepte</td>
</tr>
<tr>
<td>mascot</td>
<td>Mascot MS/MS Search</td>
<td>Unrestricted</td>
<td>9 (2015-03-29)</td>
<td>no</td>
<td>kaspecte</td>
</tr>
<tr>
<td>mof formattor</td>
<td>This repository contains a tool wrapper for the TINT MGF formattor.</td>
<td>Unrestricted</td>
<td>1 (2014-09-26)</td>
<td>no</td>
<td>kaspecte</td>
</tr>
<tr>
<td>mowconvert</td>
<td>Tool wrappers for the mowconvert application distributed as part of Proteowizard.</td>
<td>Unrestricted</td>
<td>8 (2014-09-26)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>ms data convertor</td>
<td>AB SCIEX MS Data Converter</td>
<td>Unrestricted</td>
<td>1 (2015-03-11)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>msgfplus</td>
<td>MS/MS+ Galaxy Wrapper</td>
<td>Unrestricted</td>
<td>15 (2015-03-26)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>ms riff loader</td>
<td>Loads AB Sciex riff files from URLs</td>
<td>Unrestricted</td>
<td>0 (2015-03-10)</td>
<td>no</td>
<td>galaxy</td>
</tr>
<tr>
<td>myriMatch</td>
<td>Bumbershot MyriMatch</td>
<td>Unrestricted</td>
<td>0 (2014-09-26)</td>
<td>no</td>
<td>galaxy</td>
</tr>
</tbody>
</table>

https://github.com/galaxyproteomics/tools-galaxyp
Links to workflows, webcast, pages, documentation...

Workflows
Proteogenomic studies: http://z.umn.edu/pg140
Metaproteomic studies: http://z.umn.edu/metaproteomics1

Webcast
Using ProteinPilot within Galaxy-P: z.umn.edu/ppingp

Pages
Proteogenomics page: z.umn.edu/proteinpilotpage
Metaproteomics page: z.umn.edu/metaproteomicspage

Workshop / Tutorial on proteogenomics:
Mass Spectrometry-based Proteomics Data Analysis using Galaxy-P: z.umn.edu/gcc2015gp
Multi-omic data analysis using Galaxy

In this paper, the authors discuss the use of Galaxy-P, a software framework for multi-omic data analysis, in the context of proteogenomics. They present a case study involving the analysis of multi-omic data from a specific biological sample, demonstrating the effectiveness of the Galaxy-P framework in integrating data from various omics platforms.

Flexible and Accessible Workflows for Improved Proteogenomic Analysis Using the Galaxy Framework

The authors describe the development of flexible and accessible workflows for improved proteogenomic analysis using the Galaxy framework. They highlight the importance of integrating multi-omic data for a comprehensive understanding of biological systems and present a case study demonstrating the effectiveness of their approach.

Proteomic Profiles in Acute Respiratory Distress Syndrome Differentiates Survivors from Non-Survivors

This paper reports on the use of proteomics to differentiate between survivors and non-survivors of acute respiratory distress syndrome (ARDS). The authors present a proteomic profile that can be used to predict survival outcomes in ARDS patients.

Using Galaxy-P to leverage RNA-Seq for the discovery of novel protein variations

The authors propose the use of Galaxy-P to enhance the analysis of RNA-Seq data for the discovery of novel protein variations. They demonstrate the effectiveness of their approach by identifying novel protein variations in a specific genetic dataset.

Metaproteomic analysis using the Galaxy framework

This paper describes the development of a metaproteomic analysis framework using the Galaxy platform. The authors present a case study illustrating the effectiveness of their framework in analyzing metaproteomic data from a complex biological sample.
Biochemistry, Molecular Biology & Biophysics

Tim Griffin
Candace Guerrero
Kevin Murray

Department of Medicine
Brian Sandri
Kevin Viken
Maneesh Bhargava
Chris Wendt

Department of Biology (Duluth)
Matt Andrews
Katie Vermillion
Kyle Anderson

School of Dentistry
Joel Rudney

Laboratory Medicine and Pathology
Somi Afiuni
Amy Skubitz

Biochemistry, Molecular Biology and Biophysics
Laurie Parker
Tzu-Yi Yang
Sarah Parker
Sean Seymour
Sricharan Bandhakavi

Community Based Software Development

Ira Cooke
La Trobe University, Melbourne, Australia

Bjoern Gruening
University of Freiburg, Freiburg, Germany

Lennart Martens
VIB Department of Medical Protein Research, Ugent, Belgium

Harald Barsnes and Marc Vaudel
University of Bergen, Bergen, Norway

John Chilton
Galaxy Team
Penn State University

Center for Mass Spectrometry and Proteomics
Ebbing de Jong
LeeAnn Higgins
Todd Markowski

Funding
NSF, NIH
QUESTIONS?

Mail

pjagtap@umn.edu

Follow

twitter.com/usegalaxyp