

# Integrating Protein Prospector into the Trans-Proteomic Pipeline

Peter R. Baker<sup>1</sup>, Giselle M. Knudsen<sup>1</sup>, Alex S. Yamana<sup>1</sup>, Luis Mendoza<sup>2</sup>, Robert L. Moritz<sup>2</sup> and Robert J. Chalkley<sup>1</sup>

<sup>1</sup>Mass Spectrometry Facility, Dept. of Pharmaceutical Chemistry, University of California, San Francisco, USA

<sup>2</sup>Institute for Systems Biology, Seattle, WA



## Introduction

**Protein Prospector** is a set of tools for the analysis of LC/MS/MS proteomics data (identification and quantification).

•Batch-Tag search engine pioneered many advanced options:

- Cross-linked peptide data analysis
- Glycopeptide data analysis
- Mass modification searching
- ETD/ETHcD data analysis.

The **Trans-Proteomic Pipeline (TPP)**<sup>1</sup> has some overlapping, but also several unique functionalities:

- Peptide and Protein Prophet: Statistical validation of search results
- iProphet: Merging results from multiple search engines
- Also includes tools for quantification analysis
- Recent support for cross-linking analysis

**ProHits**<sup>2</sup> is software for analysis of Affinity Purification mass spectrometry data. It uses TPP output files as input

**This poster described the integration of Protein Prospector with TPP and subsequent analyzing of Protein Prospector results using ProHits**

## Methods

•Protein Prospector uses two modules for batch processing of LC/MS/MS data:

- Batch-Tag scores the best results for each spectrum.
- Search Compare combines Batch-Tag results into proteins and filters results by FDR.
- Both programs can be run from the command line.

•Search Compare has an option to output pepXML, which is compatible with the TPP.

•Protein Prospector was modified to allow direct searching of mzML peak list files and an option to output the FASTA database file that was searched..

•The TPP (v5.0.0) was modified to add a Protein Prospector option which uses the Protein Prospector expectation values as the discriminant for calculating reliability using Peptide Prophet and Protein Prophet.

•The peak lists for the various data sets are uploaded to their respective directories using the TPP File Browser. A form was created for selecting the Batch-Tag, Search Compare and TPP parameters (Fig.1) to be used.

•The ProtXML files can be downloaded for further processing in ProHits.

## Conclusions

•Approximately 125 unique affinity experiments with two to five replicate IP-MS samples of each were processed through the Protein Prospector => TPP => ProHits workflow.

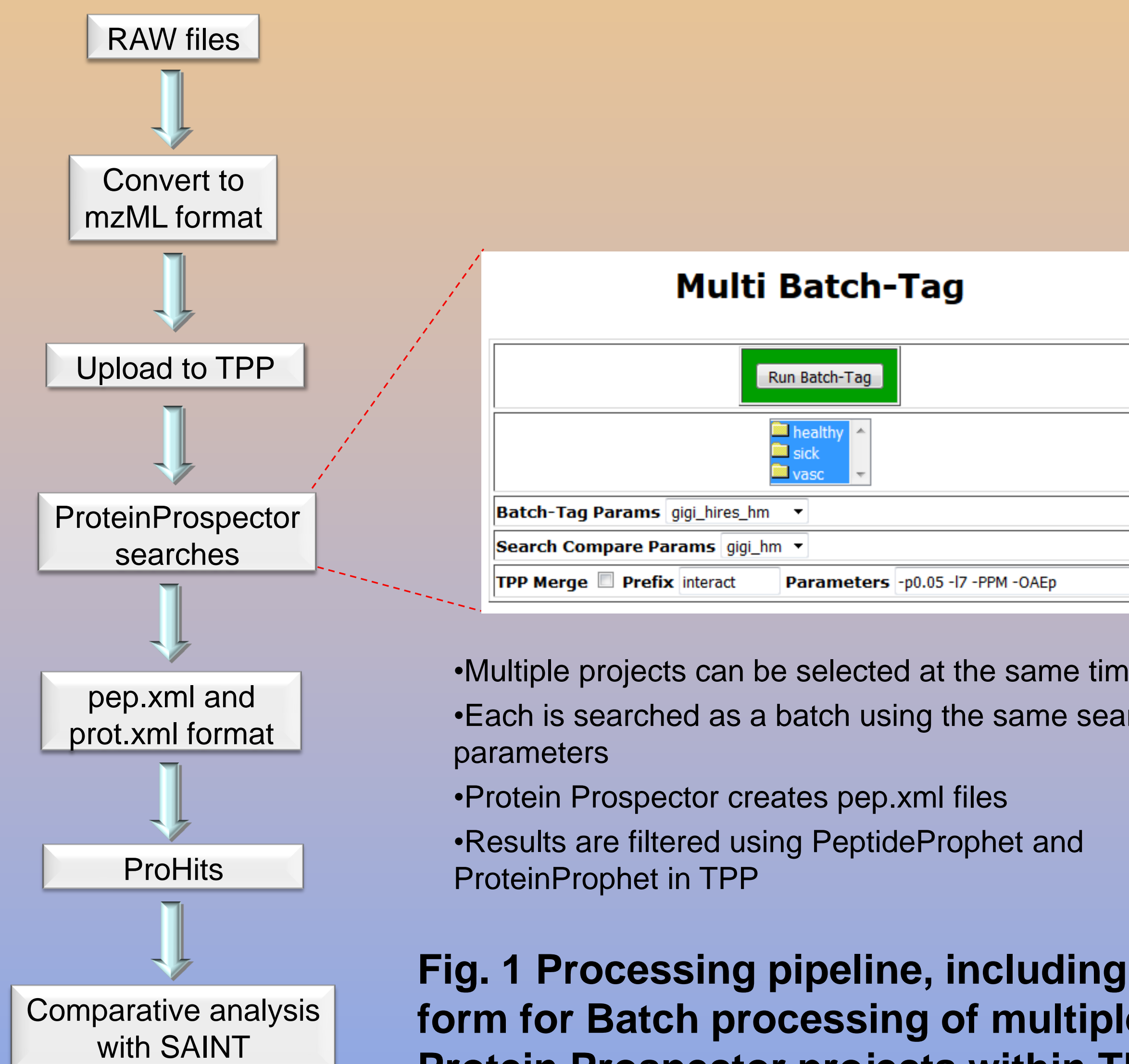
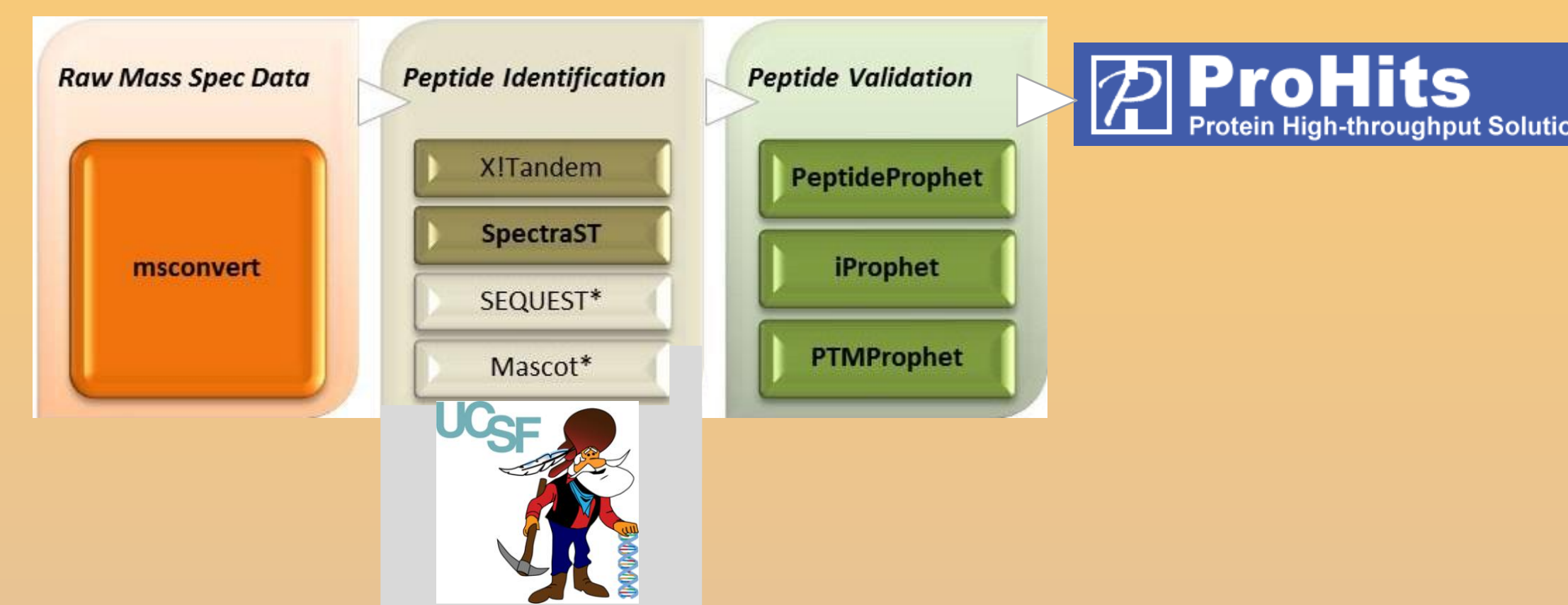
•By comparing such a large number of baits one is able to determine proteins that bind non-specifically in the IP process (e.g. to the magnetic beads).

•Several exciting targets were identified that will be followed up on to determine their significance in neurological autoimmune disorders.

## References

1. Deutsch EW, Mendoza L, Shteynberg D, Slagel J, Sun Z, Moritz RL. (2015) Trans-Proteomic Pipeline, a standardized data processing pipeline for large-scale reproducible proteomics informatics. *Proteomics Clin Appl.* 9(7-8), 745-54
2. Liu G, Zhang J, Choi H, Lambert JP, Srikanth T, Larsen B, Nesvizhskii AI, Raught B, Tyers M, Gingras AC. (2012) Using ProHits to Store, Annotate, and Analyze Affinity Purification-Mass Spectrometry (AP-MS) Data. *Curr. Protoc. Bioinform.* 39:8.16.1-8.16.32
3. Choi H, Liu G, Mellacheruvu D, Tyers M, Gingras AC, Nesvizhskii AI. (2012) Analyzing Protein-Protein Interactions from Affinity Purification-Mass Spectrometry Data with SAINT. *Curr. Protoc. Bioinform.* 39:8.15.1-8.15.23

## Workflow for IP-MS Analysis

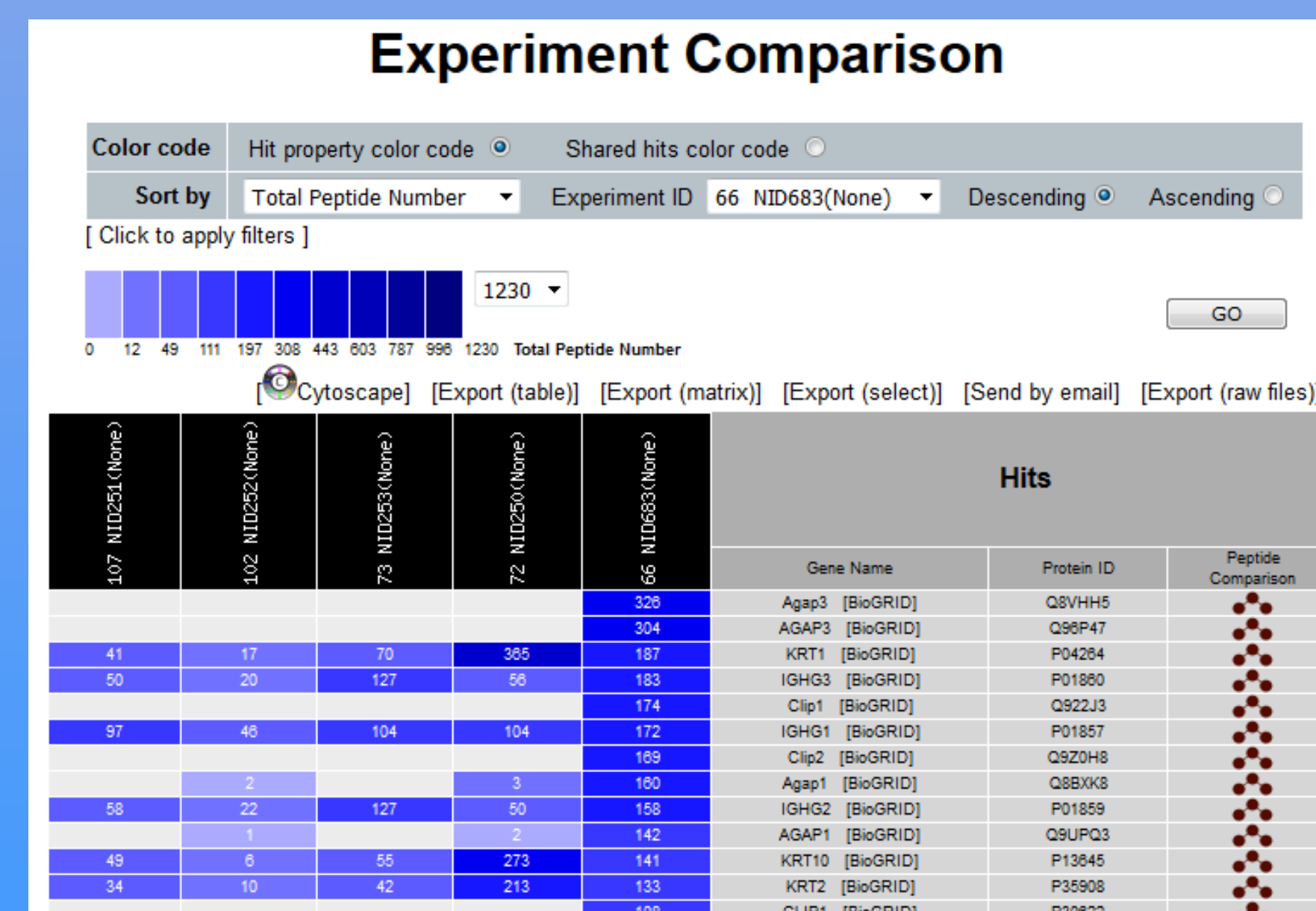


- Multiple projects can be selected at the same time
- Each is searched as a batch using the same search parameters
- Protein Prospector creates pep.xml files
- Results are filtered using PeptideProphet and ProteinProphet in TPP

**Fig. 1 Processing pipeline, including new form for Batch processing of multiple Protein Prospector projects within TPP.**

## Comparative Analysis with ProHits

ProteinProphet files in pep.xml and prot.xml format are uploaded as search results into ProHitsVM software for scoring with SAINT<sup>3</sup> and comparative analysis. The ProHits analyst module was used for results comparisons to visualize data sets.



**Fig. 2 Comparative analysis report generated by ProHits**

## Example IPMS Project

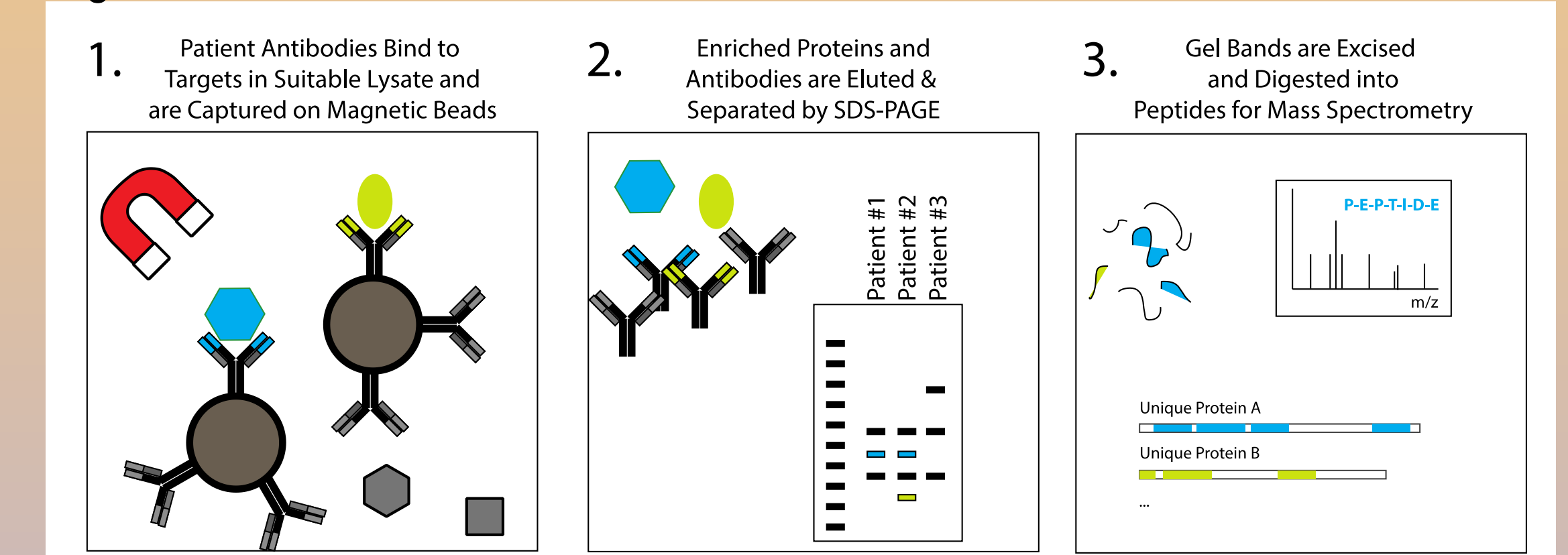
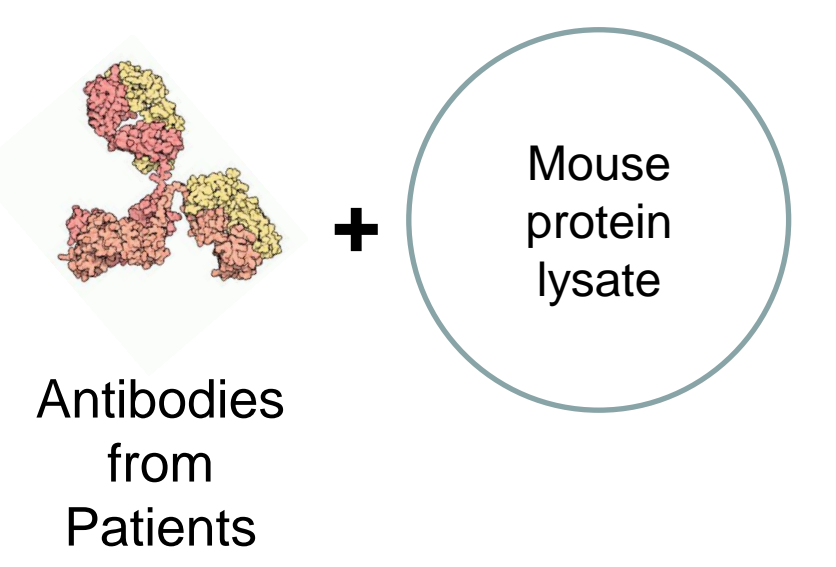
**Motivation: To identify the targets of autoantibodies in patient CSF samples.**

Patient antibodies are captured with protein A + protein G magnetic beads and incubated with protein lysates prepared from mouse tissues as a source of prey proteins.

Data are searched as appropriate with the SwissProt human + mouse database.

Assumption: If antibody binds a particular mouse protein its normal target is probably the human homolog.

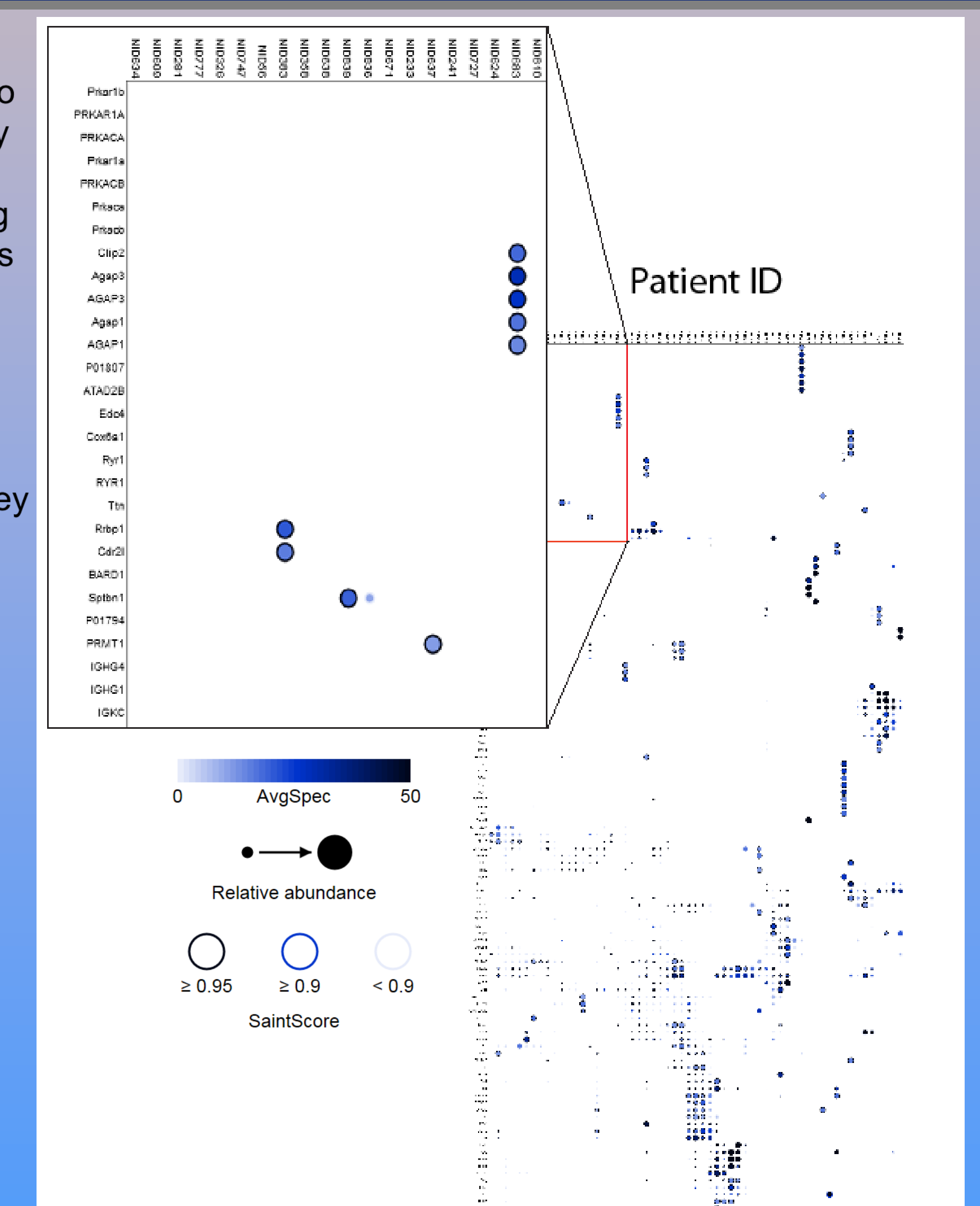
Probing antibody – antigen pairs in auto-immune disease.



**Fig. 3 IPMS workflow**

## Scored Results

ProHits Analyst is used to define control and affinity experiments for comparisons and scoring with SAINT. ProHits-viz is then used to visualize SAINT score data. Here the SAINT score filter is  $\geq 0.9$ . Alternative views include heat maps for patient vs patient and prey vs prey visualization.



**Fig. 4 Dot plot generated by ProHits-viz**

## Acknowledgements

•This work was supported by grant number 8P41GM103481, The Sandler Family Foundation, The William K. Bowes Foundation, The Weill Foundation, and The Dr Miriam and Sheldon G. Adelson Medical Research Foundation.