

A Galaxy-based immunopeptidogenomic pipeline for neoantigen discovery

Subina Mehta, Kyle Richards, James E. Johnson, Reid Wagner, Tyler Jubenville, Flavia Popescu, Suzanne Coleman, Fengchao Yu, Alexey Nesvizhskii, Pratik Jagtap, David A. Largaespada, Timothy J. Griffin



Acknowledgements



UNIVERSITY OF MINNESOTA



Subina Mehta

James E. Johnson (JJ)*

Reid Wagner

Pratik Jagtap

Tim Griffin

Kyle Richards

Tyler Jubenville

Flavia Popescu

Suzanne Coleman

David A. Largaespada



**In memoriam*



Minnesota
Supercomputing Institute
UNIVERSITY OF MINNESOTA
Driven to Discover®



Fengchao Yu

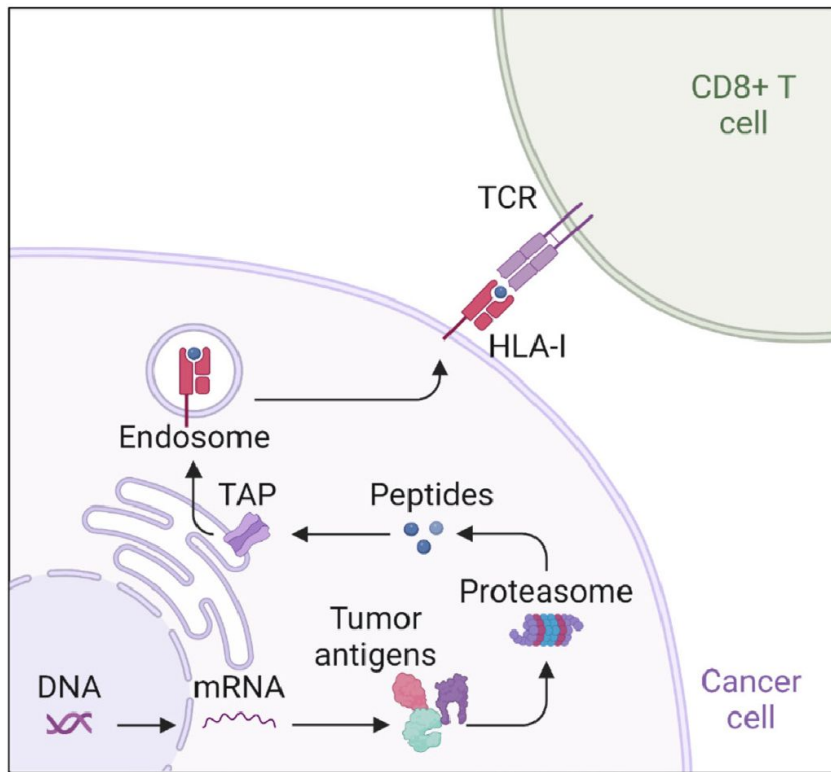
Alexey Nesvizhskii



NATIONAL CANCER INSTITUTE
Informatics Technology for
Cancer Research



Antigen presentation on the major histocompatibility complex (MHC, aka HLA) and immunopeptidomics

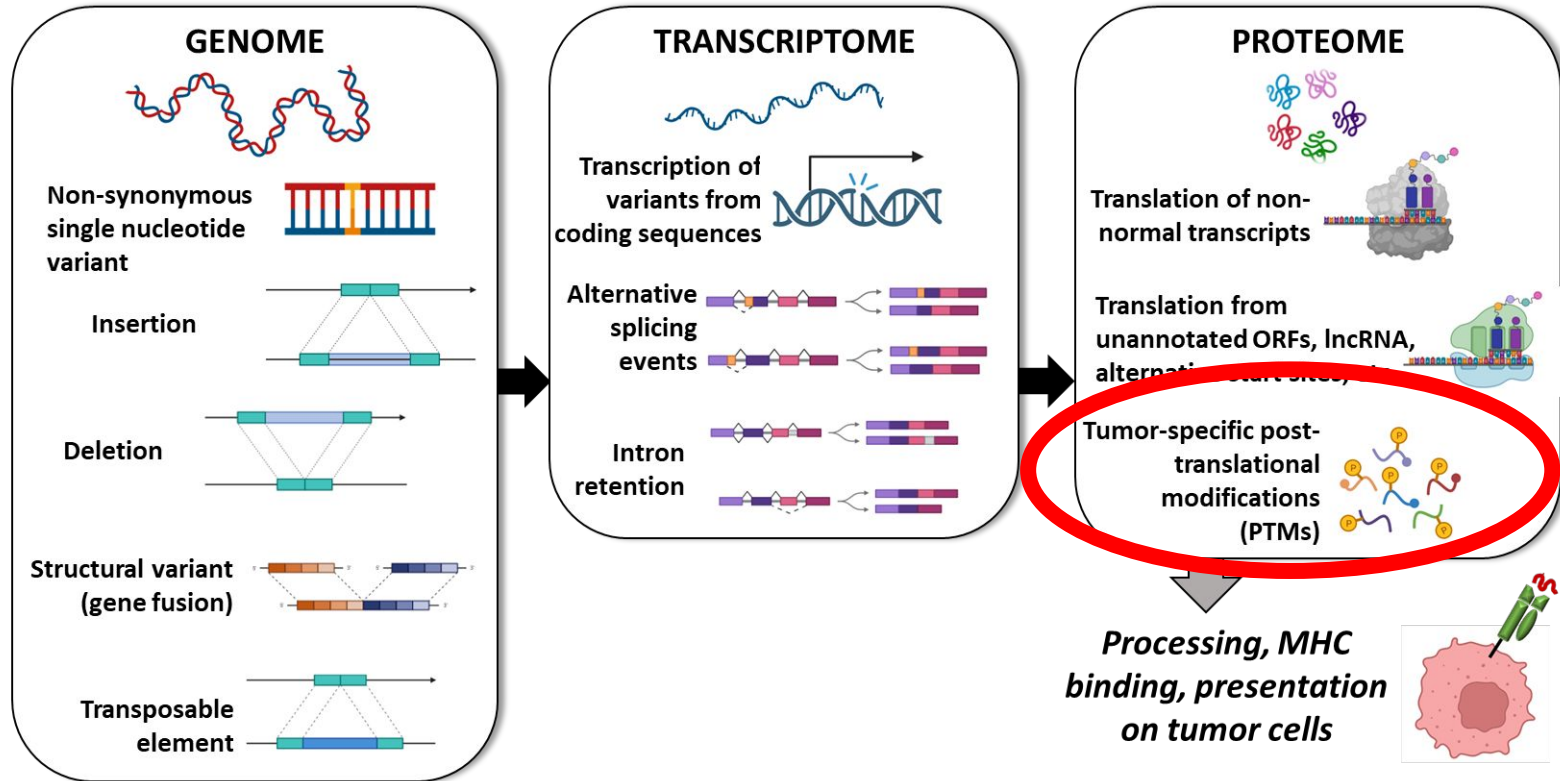


Zhang B, Bassani-Sternberg M. *J Immunother Cancer* 2023;11:e007073.

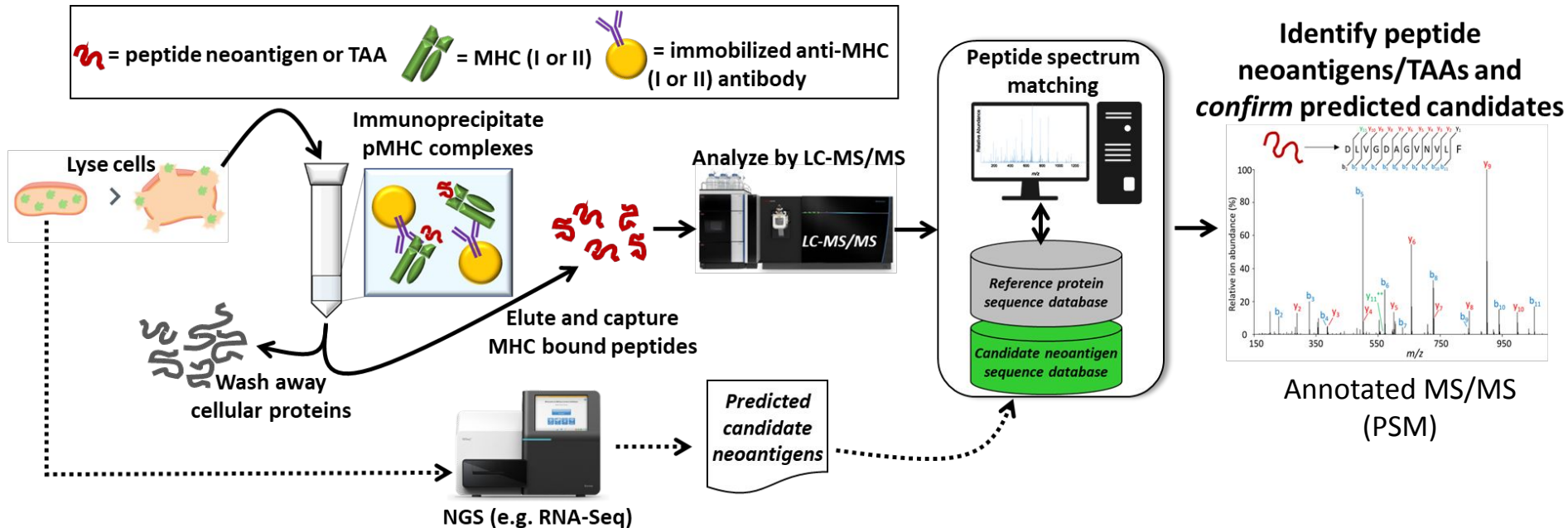
- Processing of “foreign” proteins expressed in the cancer cell
- Bound peptides act as cancer biomarkers
- Peptide-based immunotherapies, potentially personalized to individual tumors
- Potential for “tumor vaccine” development



Sources of antigens (reference sequences) and neoantigens (non-normal sequences)



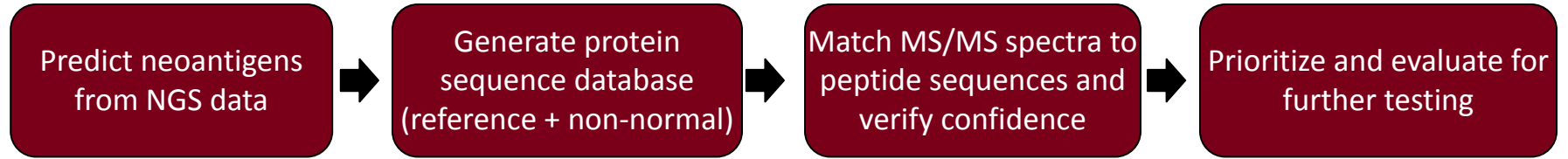
Direct confirmation of MHC-bound peptides: Mass spectrometry (MS)-based immunopeptidomics



- Confirms binding of peptides to MHC that are *predicted* from NGS
- Identifies peptides with post-translational modifications (PTMs) that regulate processing and MHC binding



Immunopeptidogenomic bioinformatics: Modular, multi-omic analysis ideally suited for Galaxy

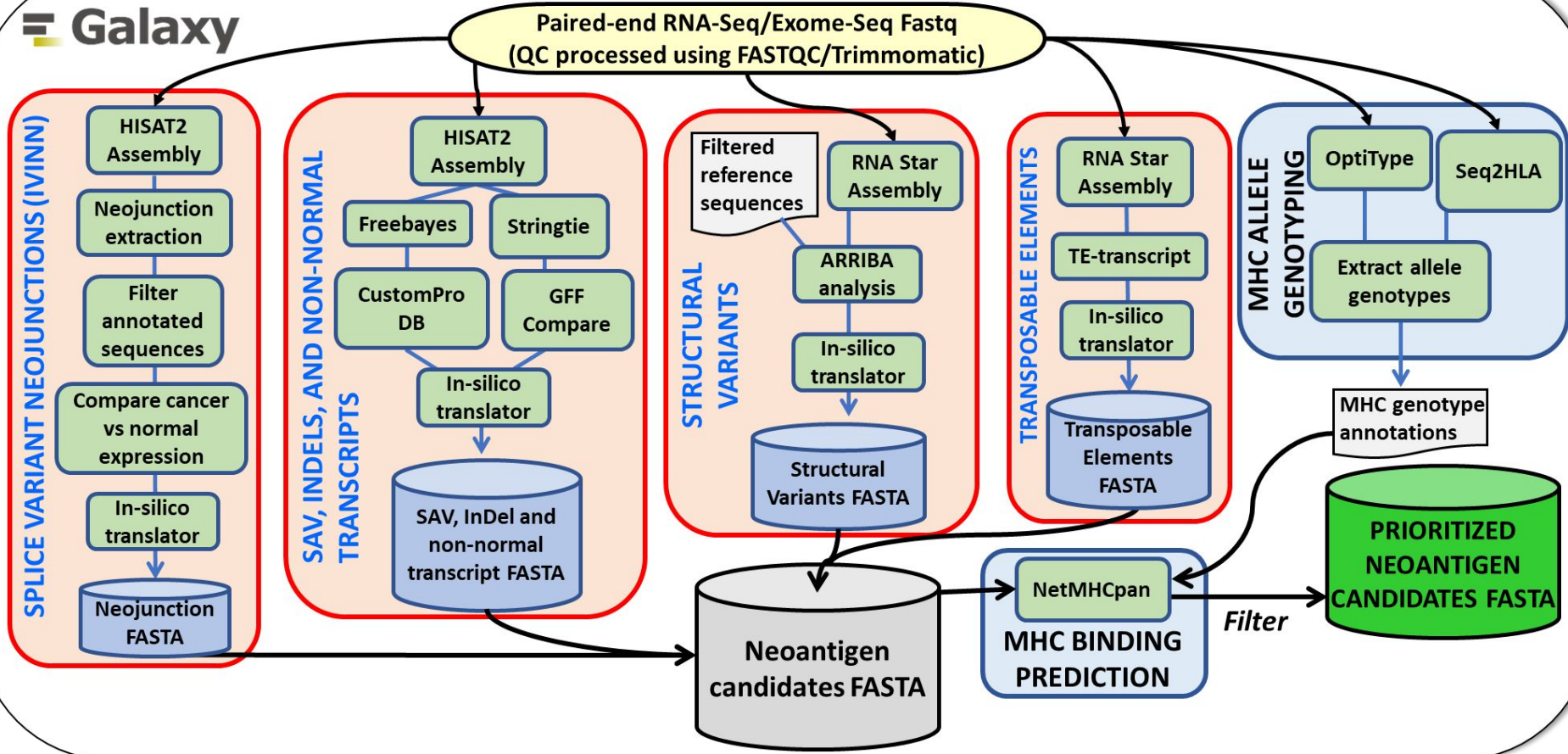


 **Galaxy**

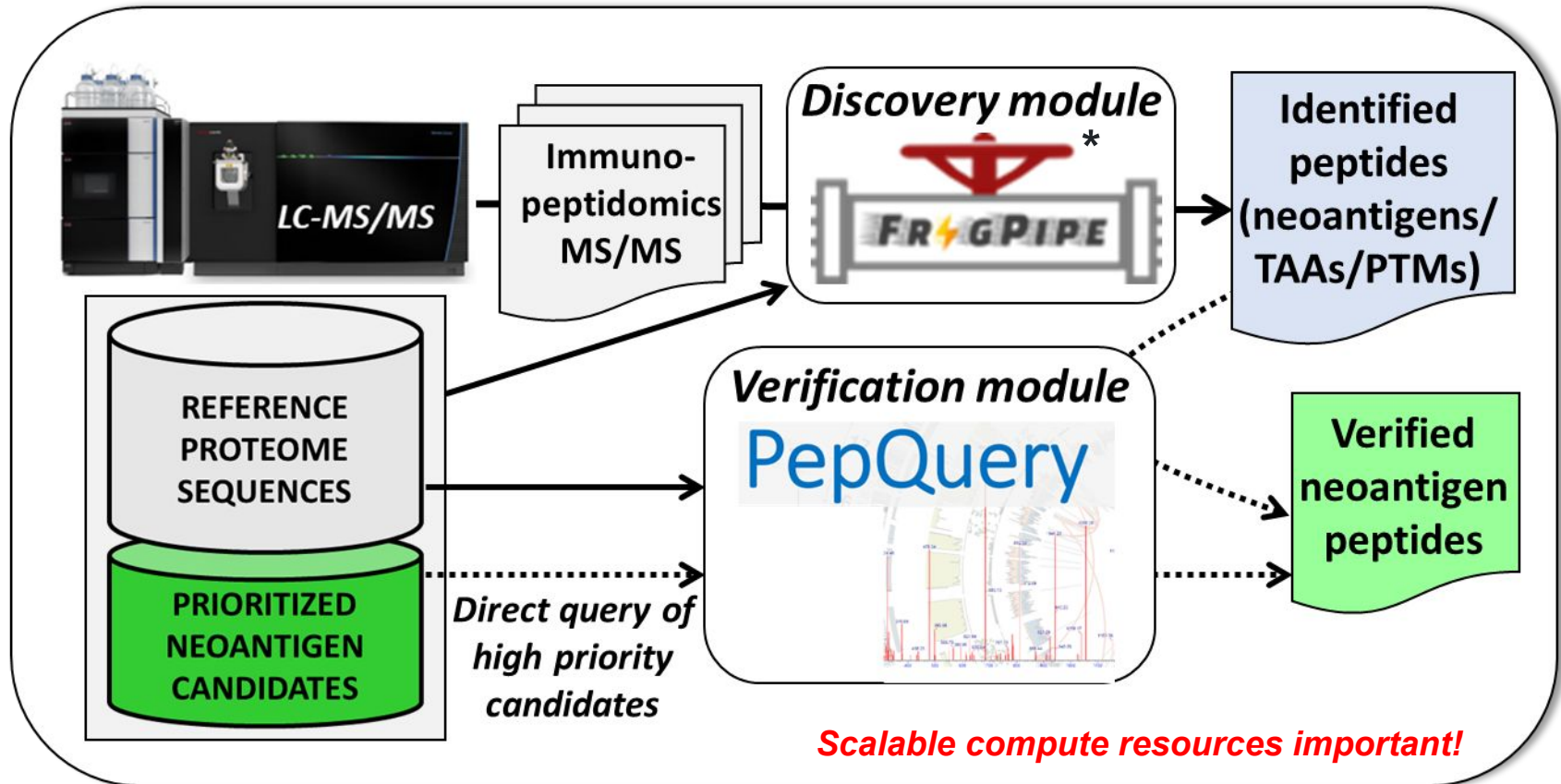


Predicting neoantigens and generating customized protein sequence databases

Galaxy



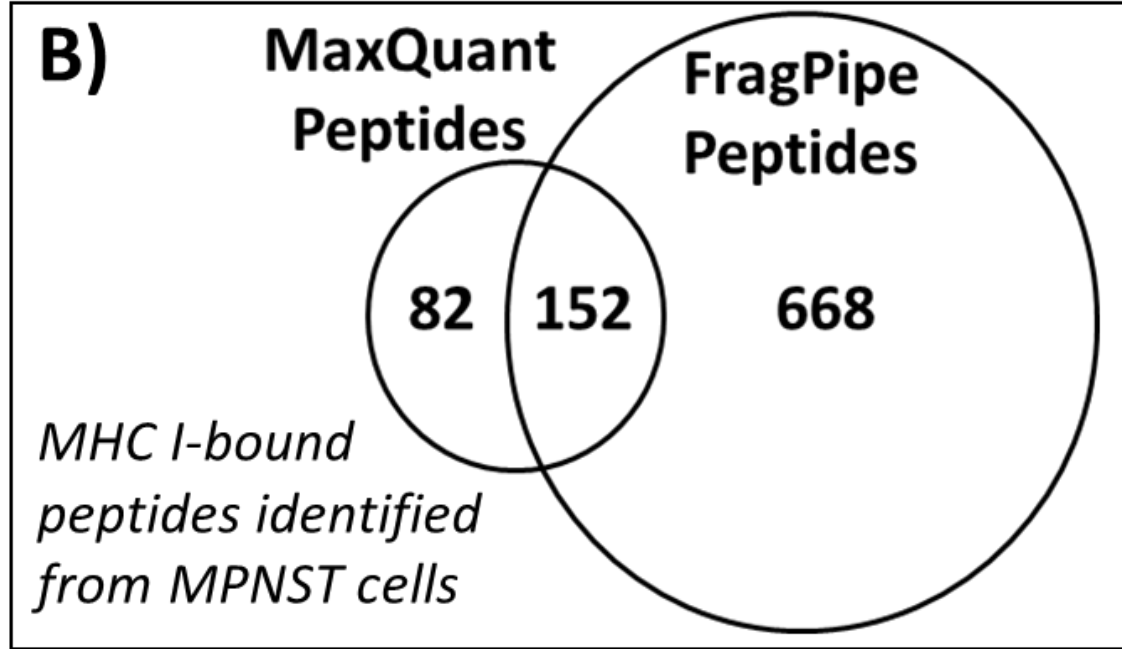
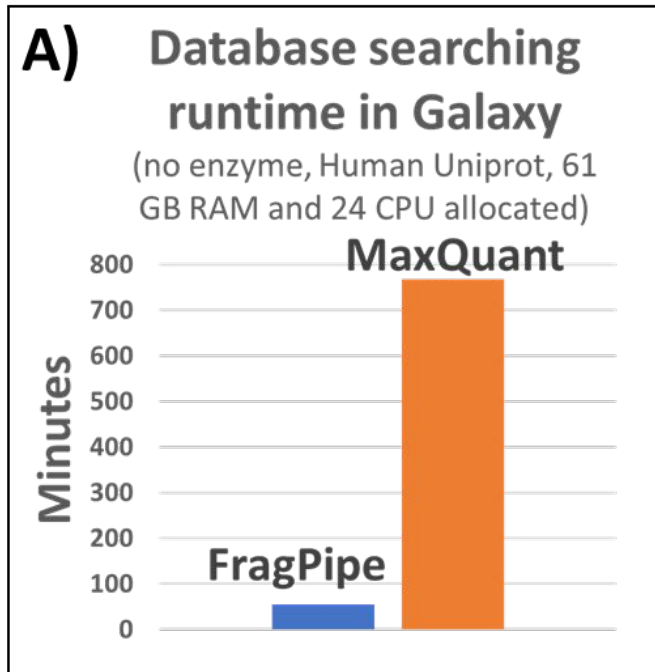
Matching peptide sequences to MS/MS and verifying confidence



*An Update on the FragPipe Galaxy Tool for Proteomics Workflows; Reid Wagner, poster 55



A look at performance on scalable compute resources



- Matching peptides generated by unknown protease-processing of proteins opens up large search spaces with high demands on active memory



Future: Prioritizing verified neoantigen or TAA peptides for further study

Verified neoantigen peptides



 pVACtools
 CEDAR
Cancer Epitope Database and Analysis Resource

HLA ligand atlas




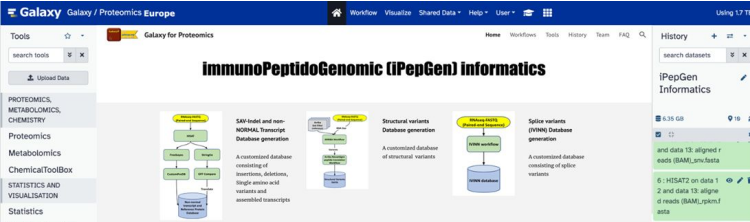
Prioritized candidates



Empirical evaluation of immunogenicity



 useGalaxy.*



Galaxy / Proteomics Europe

Tools

search tools

Upload Data

PROTEOMICS, METABOLOMICS, CHEMISTRY

Proteomics

Metabolomics

ChemicalToolBox

STATISTICS AND VISUALISATION

Statistics

immunoPeptidogenomic (IPepGen) informatics

Workflow Visualize Shared Data Help User

History

search datasets

IPepGen Informatics

6:16:08

and data 13: aligned 1 edits (SAM_snp-fasta)

6:16:52 on data 1

2 and data 13: aligned 0 reads (SAM_rpm1-fasta)

