META-IPEP: A GALAXY-BASED METAIMMUNOPEPTIDOMICS BIOINFORMATICS PIPELINE RIGOROUSLY CHARACTERIZES MICROBIAL PEPTIDE ANTIGENS BOUND TO THE HUMAN LEUKOCYTE ANTIGEN (HLA) COMPLEX

Katherine Do[1], Subina Mehta[1], Reid Wagner[2], Fengchao Yu[3], Alexey Nesvizhskii[3,4], Timothy J. Griffin[1,5], Pratik D. Jagtap[1,5]

[1]Biochemistry, Mol. Biology and Biophysics, University of Minnesota, Minneapolis, MN; [2]Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, MN; [3]Department of Pathology, University of Michigan, Ann Arbor, MI; [4]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI; [5]Masonic Cancer Center, University of Minnesota, Minneapolis, MN

INTRODUCTION

- Emerging studies have described the presentation of intratumoral microbial peptides by the human leukocyte antigen (HLA) complex.
- Mass spectrometry (MS)-based metaproteomics can be used to characterize these "meta-immunopeptides," bettering our understanding of tumor development and progression.
- However, the complex tumor microbiome presents numerous bioinformatic challenges.
- The Meta-iPep pipeline was developed on the Galaxy platform[2] to enable bioinformatic analysis of microbial peptides from immunopeptidomic tandem MS (MS/MS) data, allowing researchers to investigate microorganisms that may contribute to cancer.

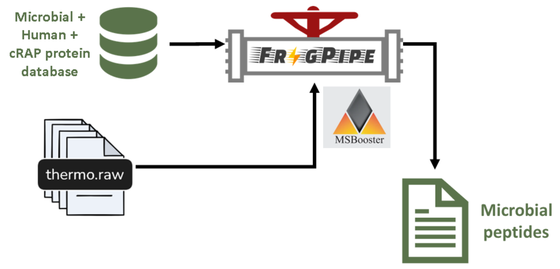
EXAMPLE DATASETS

- Datasets from melanoma samples[1] were used for demonstration (samples 86B2, 92B3, 27). Raw MS/MS data were obtained via data-dependent acquisition (DDA) mode (Q Exactive Plus).
- Protein sequence databases: Microbial database of 40 bacterial species (697K seqs); Microbial + Human + cRAP database (780K seqs); Human + cRAP database (83.2K seqs)

METHODS: META-IPEP PIPELINE

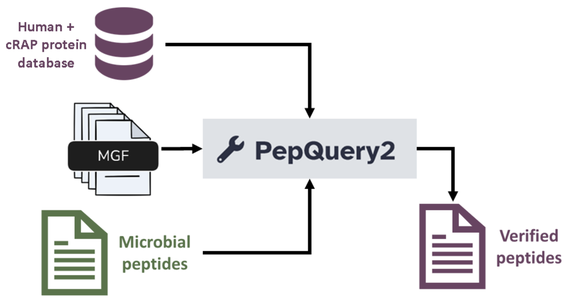
Module 1: Peptide Identification

FragPipe matched MS/MS data against a customized protein sequence database and identified microbial peptides.



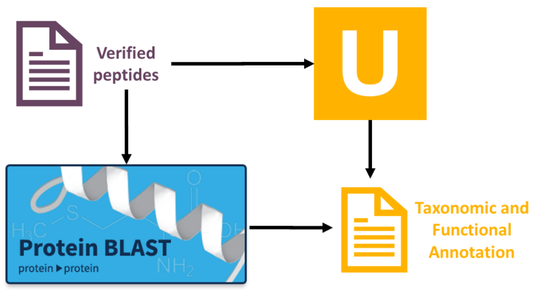
Module 2: Peptide Verification

Microbial peptides identified using FragPipe were then validated using PepQuery2, resulting in verified non-host microbial peptides.



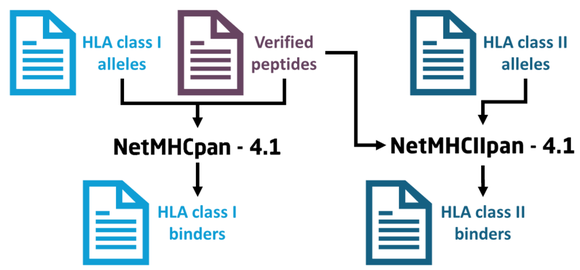
Module 3: Peptide Annotation

Microbial peptides verified by PepQuery2 were analyzed by Unipept and BLAST-P for taxonomic and functional annotation.



Module 4: HLA Binding Prediction

The verified peptides were used as input for NetMHCpan and NetMHCIIpan to predict peptide binding to HLA class I and class II molecules, respectively.



Key Takeaways

- **"Meta-immunopeptides"** are intratumoral microbial peptides that are presented by the HLA and can be characterized via MS-based metaproteomics.
- The **Meta-iPep pipeline** was developed to address the many bioinformatic challenges posed by the complexity of the tumor microbiome.
- Meta-iPep is a modular bioinformatics pipeline for **rigorous analysis of MS-based meta-immunopeptidomics data** that empowers community adoption via the Galaxy ecosystem.
- This pipeline can be expanded to **analyze samples from different cancer types** to better our understanding of **tumor development and progression**.

RESULTS

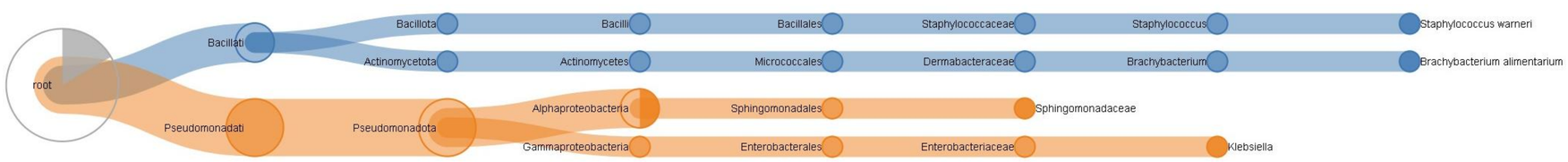


FIGURE 1. Example of Microbial Taxonomy (Sample 86B2).

Example Sample 86B2

Taxonomy (FIG. 1)

- Three genera identified: ***Klebsiella***, ***Staphylococcus***, and ***Brachybacterium***.
- Two bacterial species identified: ***Staphylococcus warneri*** and ***Brachybacterium alimentarium***.

Microbial peptides of interest

- Out of **92 microbial peptides** identified **(TABLE 1)**, **four microbial peptides were of interest**, based on the quality of their mass spectra: KAGMLSITY, DTEIPRKELY, FAEGAGRLAG, AADKAAADKAAADK **(TABLE 2, FIG. 2)**.

TABLE 1. Results summary for three melanoma samples. SBs: strong binders; WBs: weak binders.

Tool	Data	Sample 86B2	Sample 92B3	Sample 27
FragPipe	**Total peptides**	11,262	7,424	5,998
	Microbial peptides	92	88	63
PepQuery2	**Validated PSMs**	17	15	12
	Microbial peptides	12	11	6
NetMHCpan 4.1	**HLA class I SBs**	21	11	16
	HLA class I WBs	40	27	30
NetMHCiipan 4.1	**HLA class II SBs**	5	3	2
	HLA class II WBs	6	4	2

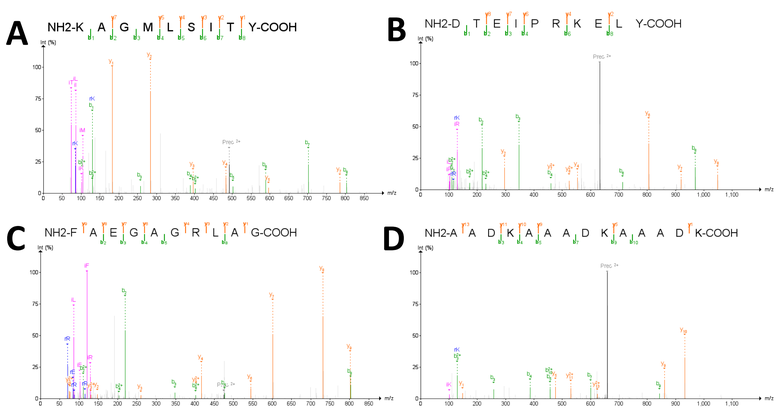


FIGURE 2. Mass spectra for four microbial peptides of interest:
A) KAGMLSITY, **B)** DTEIPRKELY, **C)** FAEGAGRLAG, and **D)** AADKAAADKAAADK.

TABLE 2. Excerpt of HLA binding predictions for sample 86B2.

Microbial Peptide (MS scan)	Length	Taxonomic and Functional Annotation	Predicted Binding Core	Select Predicted HLA Alleles
KAGMLSITY (Seq47093_QE3_raw:31656:2)	9	*Staphylococcus warneri* Transmembrane transporter activity	KAGMLSITY	HLA-B*35:01, HLA-C*02:02
DTEIPRKELY (Seq47093_QE3_raw:25390:2)	10	*Brachybacterium alimentarium* DNA Partitioning ATPase	DTEIPRKELY	HLA-A*01:01, HLA-B*44:02
FAEGAGRLAG (Seq47093_QE3_raw:20028:2)	10	Family Sphingomonadaceae Phage tail assembly chaperone protein (TAC6)	FAEGAGRL	HLA-C*03:03, HLA-C*08:02
AADKAAADKAAADK (Seq47095_QE3_raw:4568:2)	14	root dextransucrase	AAADKAAAD	HLA-DQA1*01:02/ DQB1*03:02

CONCLUSIONS

- Meta-iPep is a modular, Galaxy-based pipeline for the identification, rigorous verification, and characterization of HLA-I- and HLA-II-binding microbial peptides from immunopeptidomics data.
- In this study, we used the Meta-iPep pipeline to examine bacterial peptides in the HLA immunopeptidome of melanoma samples.
- These workflows will be made publicly accessible on the European Galaxy server, and accompanying tutorial materials will be disseminated on the Galaxy Training Network (GTN).

FUTURE DIRECTIONS

- We envision that the Meta-iPep pipeline can be used for different samples besides melanoma to investigate microorganisms that may contribute to specific cancers.
- Considering peptide inversion is one future area of exploration that may yield greater understanding of factors that influence peptide—MHC binding, thereby improving MHC prediction models.

REFERENCES

1. Kalaora S, Nagler A, Nejman D, Alon M, Barbolin C, Barnea E, et al. Identification of bacteria-derived HLA-bound peptides in melanoma. Nature. 2021;592(7852):138–43.
2. Galaxy Community. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. Nucleic Acids Res. 2024;52(W1):W83-W94. doi: 10.1093/nar/gkae410.

ACKNOWLEDGMENTS

This work is supported by NIH grants UG3CA244687 and 1U01CA288888, as well as grant P30CA077598 to the Masonic Cancer Center at the University of Minnesota.
COI DISCLOSURE: A.I.N. and F.Y. receive royalties from the University of Michigan for the sale of MSFragger, IonQuant, and diaTracer software licenses to commercial entities. All license transactions are managed by the University of Michigan Innovation Partnerships office, and all proceeds are subject to university technology transfer policy. Other authors declare no other competing interests.