NPvis: interactive visualizer for MS/MS fragmentation of natural products

Olga Kynyavskaya[1,2], Alexey Gurevich[1,*]

[1] Center for Algorithmic Biotechnology, St. Petersburg State University, St. Petersburg, Russia;

[2] National Research University Higher School of Economics, St. Petersburg, Russia; *aleksey.gurevich@spbu.ru

Overview

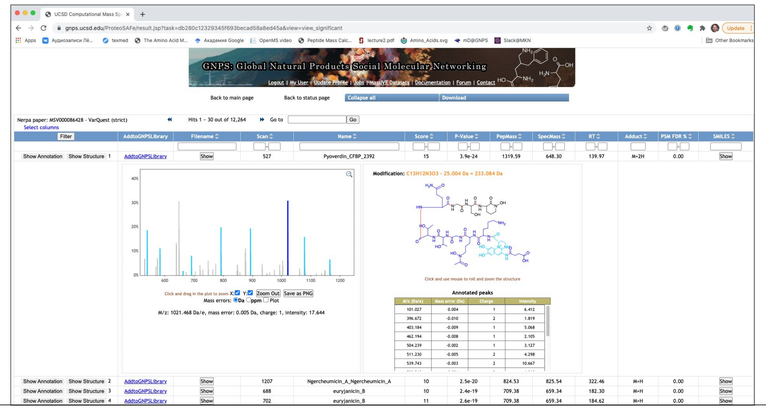
- Natural products (NPs) include many antibiotics and other bioactive compounds [1]
- High-throughput NP discovery is driven by tandem mass spectrometry (MS/MS). Analysis of experimental MS/MS data involves computational methods
- MS/MS can be searched against reference spectral libraries of previously analysed compounds, however they are currently small for NPs. Another approach is to use *in silico* fragmentation software. These tools compare theoretical MS/MS fragmentations of target structures from large chemical NP databases (e.g. the Dictionary of Natural Products [2]) to experimental MS/MS peaks. Despite the recent progress in the NP database search tools development, these instruments are still in their infancy and require validation of their outputs
- Visualization of MS/MS compound fragmentation is a powerful method for validation of *in silico* identification results [3]. Standard proteomics visualization tools are not applicable to NPs due to
  - complex chemical structure (e.g. cyclic, branched)
  - non-standard amino acid residues or absence of amino acids at all (e.g. polyketides, lipids)
  - unclear MS/MS fragmentation patterns (e.g. in the absence of peptide bonds)
- We present NPvis, a versatile interactive visualizer for MS/MS fragmentation of wide range of NPs

Implementation

- NPvis is implemented in JavaScript/JQuery (visualization part) and Python/RDKit (preprocessing)
- Build on top of Lorikeet, an MS/MS visualizer for proteomics http://uwpr.github.io/Lorikeet/
- Fragmentation and compound-MS/MS matching is done with NPDtools (VarQuest [4] / Dereplicator+ [5] /etc) https://cab.spbu.ru/software/npdtools/

Integration with GNPS

- Global Natural Products Social Molecular Networking (GNPS, https://gnps.ucsd.edu/ [6]) is a community-driven platform for storing, sharing and analysing NP MS/MS data
- NPvis is integrated in the output of the GNPS In silico Tools



Input data

Mass spectra: MGF, mzML, mzXML
Compounds: MDL MOL, SMILES, InChIKey, peptide sequence

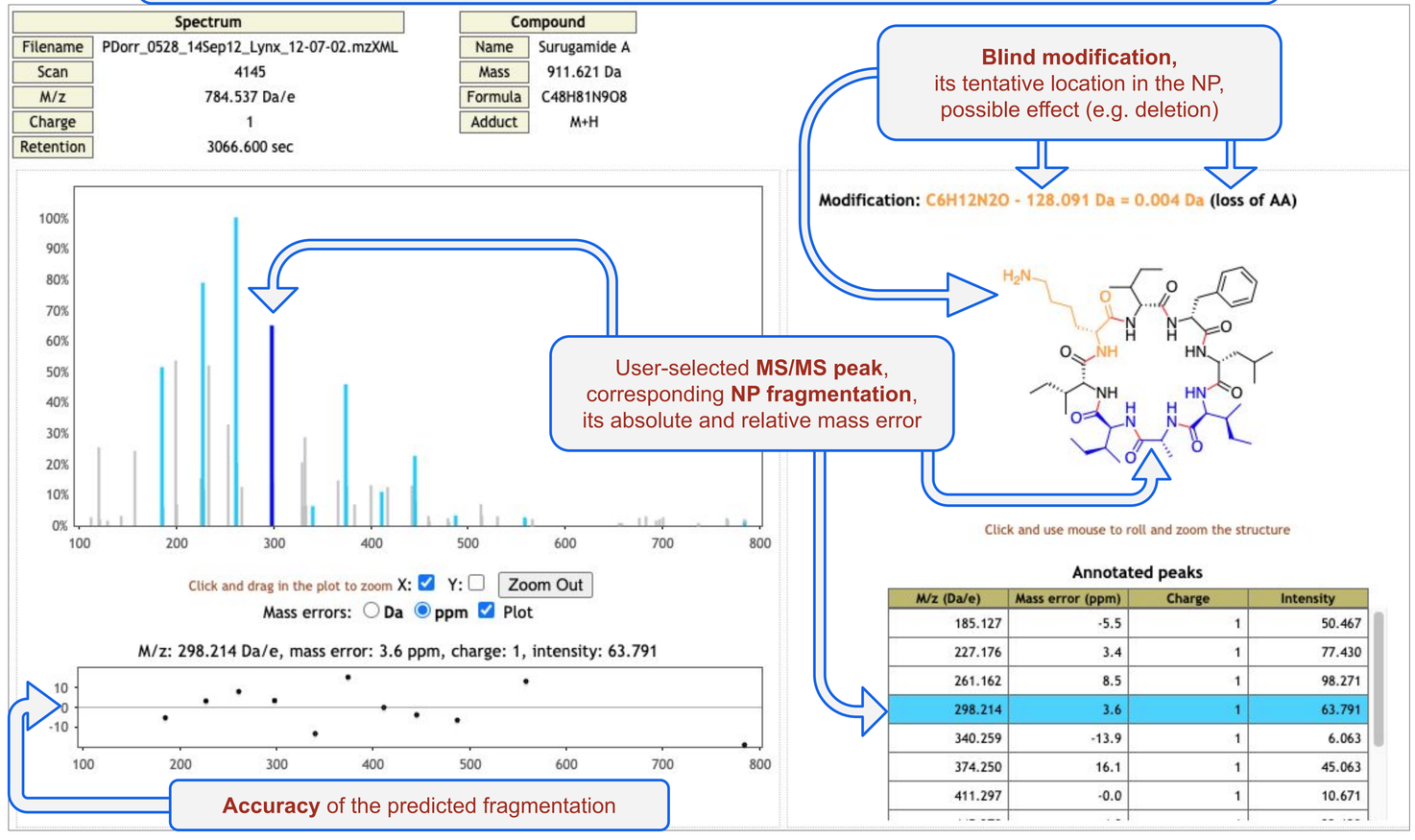
Key features

- Compounds of virtually arbitrary chemical structure
- Non-peptidic MS/MS fragmentation
- Single blind modifications

Coming soon

Improved support for non-peptidic NPs
Release of the command-line version
Web-service at http://cab.cc.spbu.ru/

More examples and info at https://cab.spbu.ru/software/npvis. Stay tuned!



Acknowledgement

We would like to thank Hosein Mohimani (CMU) for the initiation of the project, Alla Mikheenko (CAB, SPbU) for help with the implementation of the NPvis prototype, and Ming Wang (UCSD) for the assistance in the integration of NPvis with GNPS.

Funding: Russian Science Foundation (grant 20-74-00032)

References

[1] Cragg G & Newman D (2013) Natural products: a continuing source of novel drug leads. BBA-General Subjects, 1830(6), 3670-3695

[2] Buckingham J (1994) Dictionary of natural products. London: Chapman & Hall

[3] Tyanova S et al (2015) Visualization of LC-MS/MS proteomics data in MaxQuant. Proteomics, 15(8):1453-1456

[4] Gurevich A et al. (2018) Increased diversity of peptidic natural products revealed by modification-tolerant database search of mass spectra. Nat Microbiol, 3, 319

[5] Mohimani H et al (2018) Dereplication of microbial metabolites through database search of mass spectra. Nat comm, 9:4035

[6] Wang M et al. (2016) Sharing and community curation of mass spectrometry data with Global Natural Products Social molecular networking. Nat Biotechnol, 34, 828