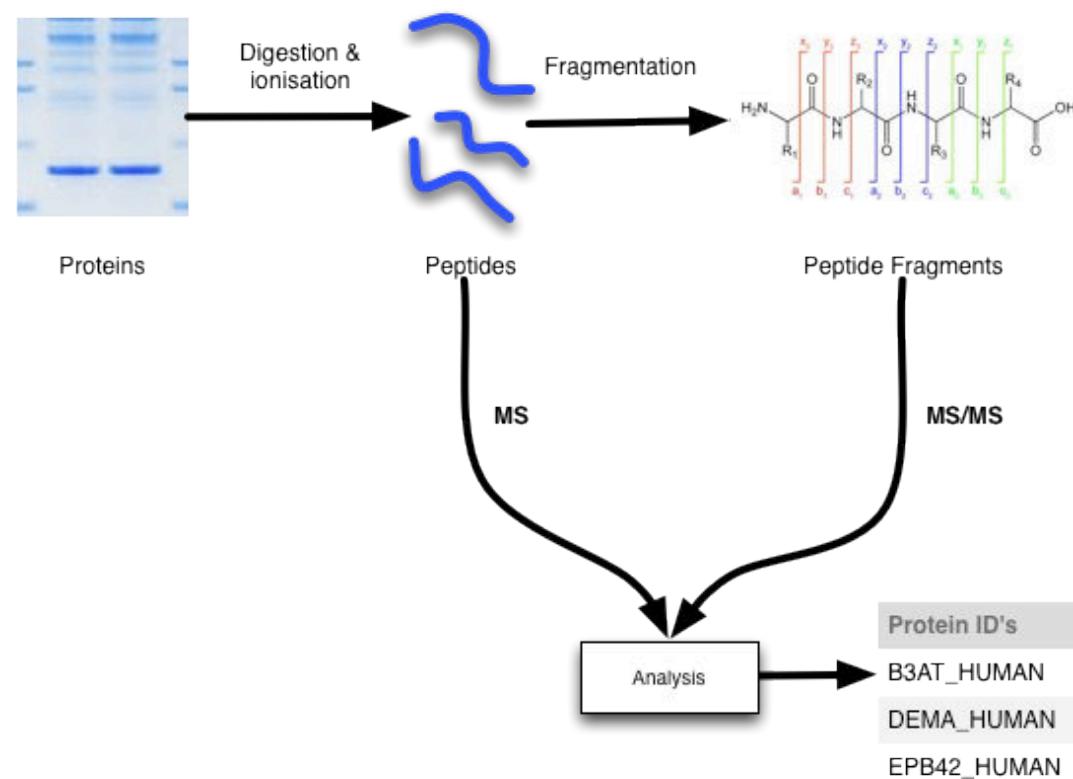


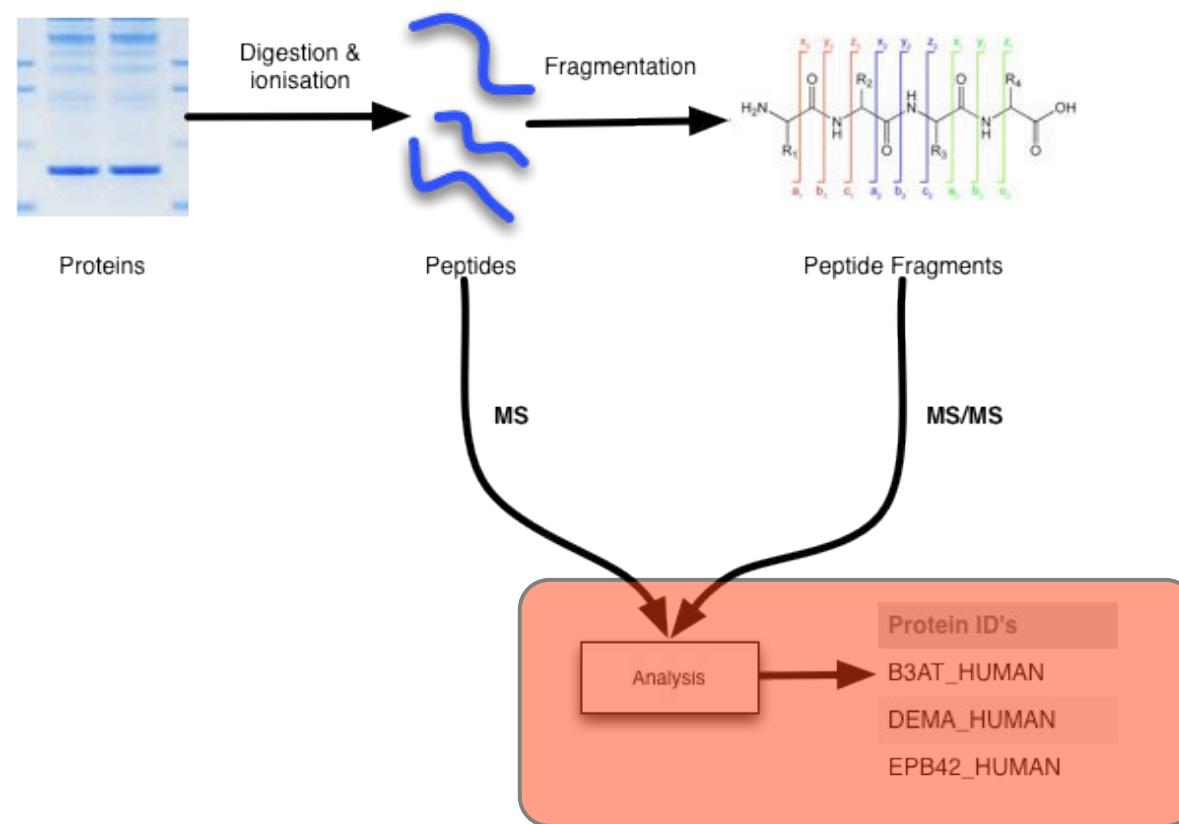
# Proteomics with Galaxy

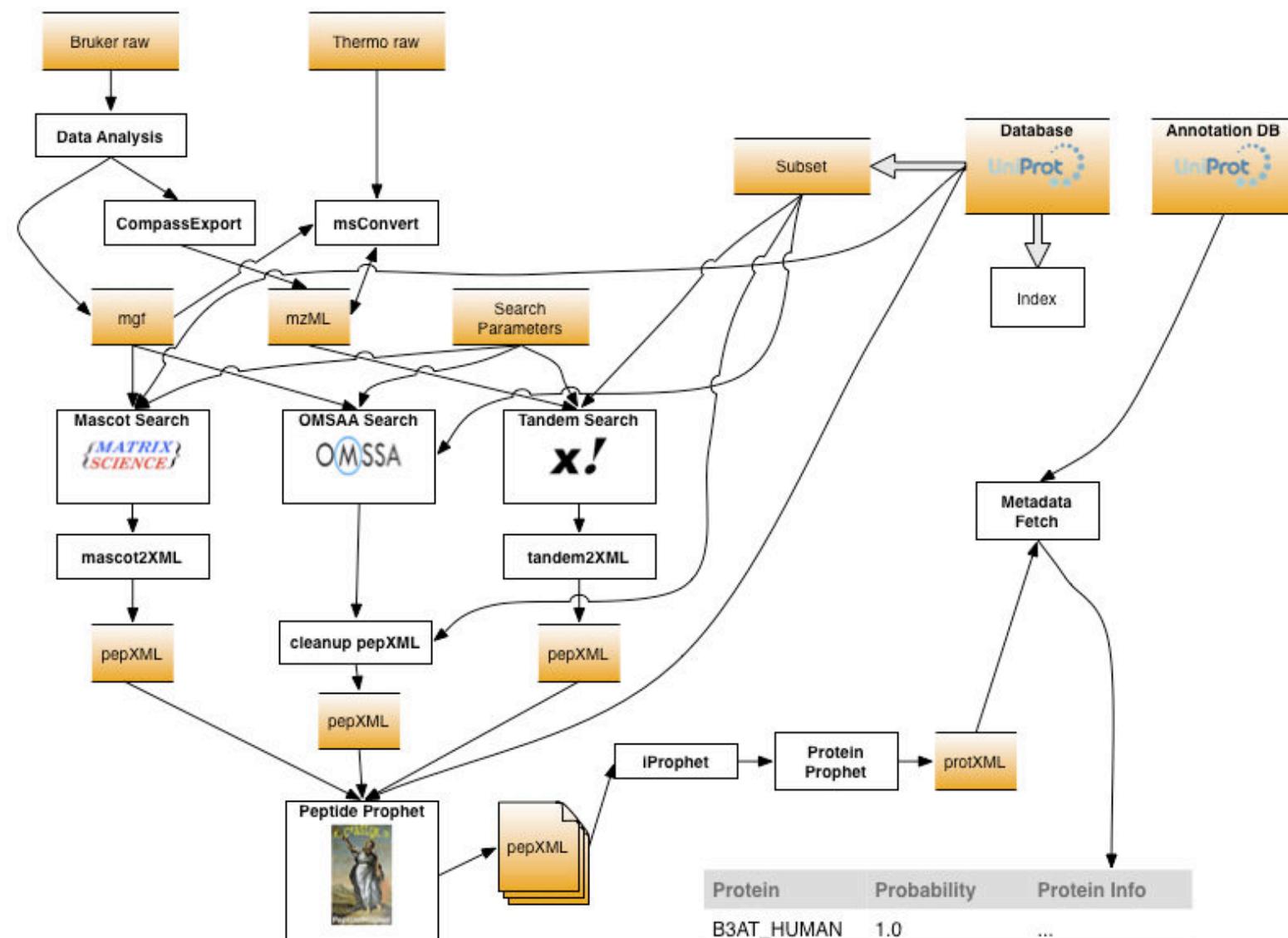
Ira Cooke

# Protein ID by Tandem Mass Spectrometry

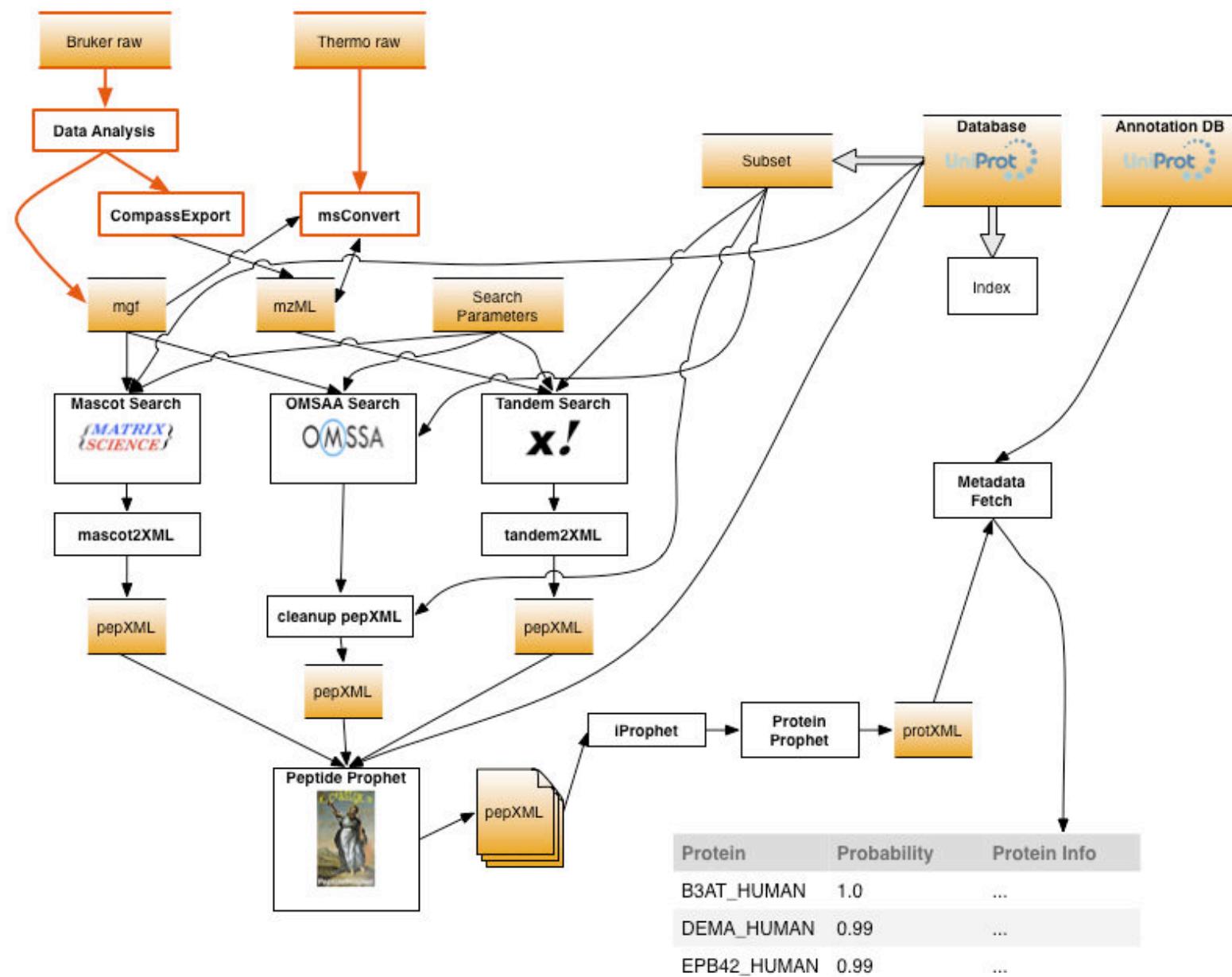


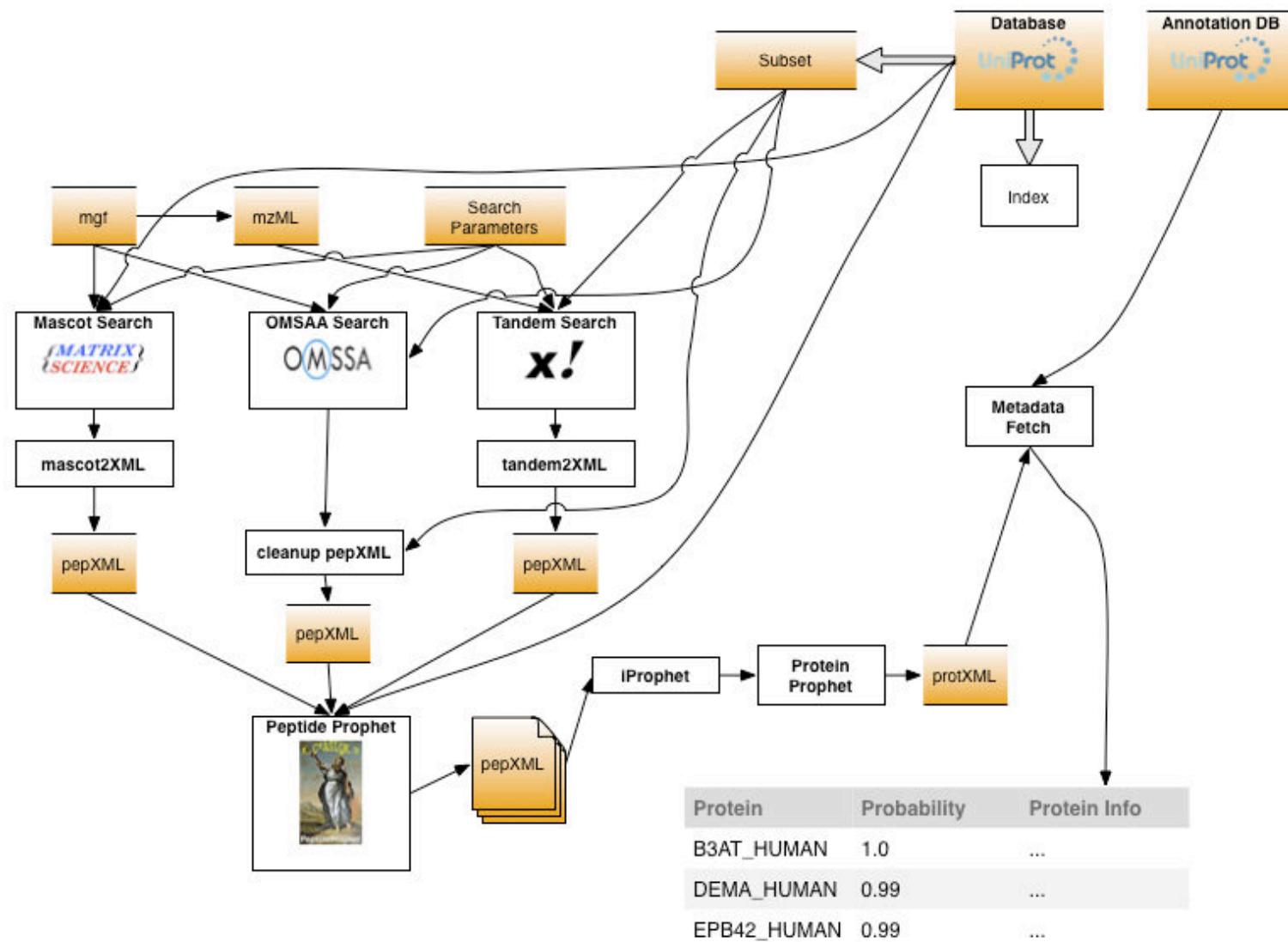
# Protein ID by Tandem Mass Spectrometry

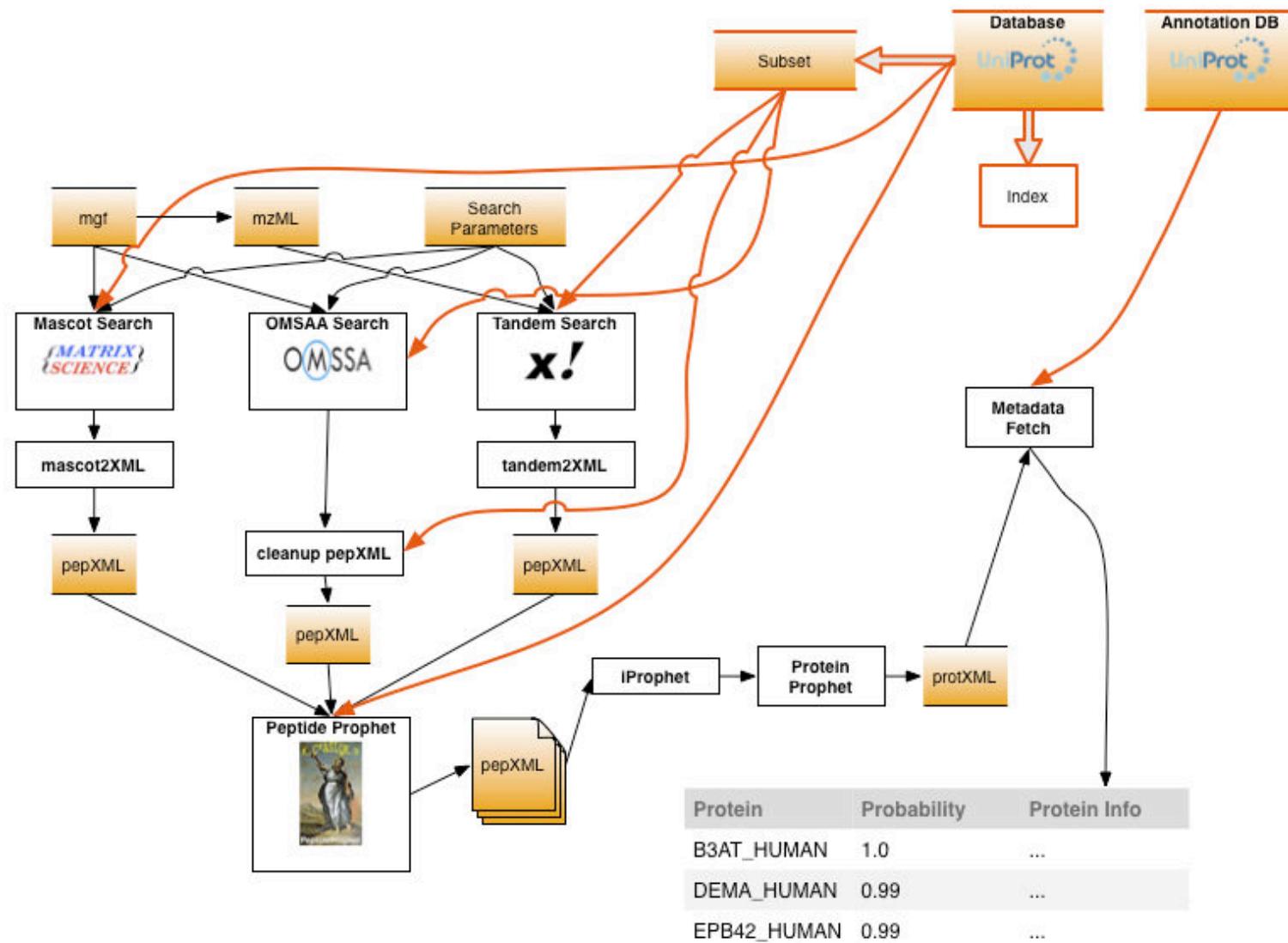


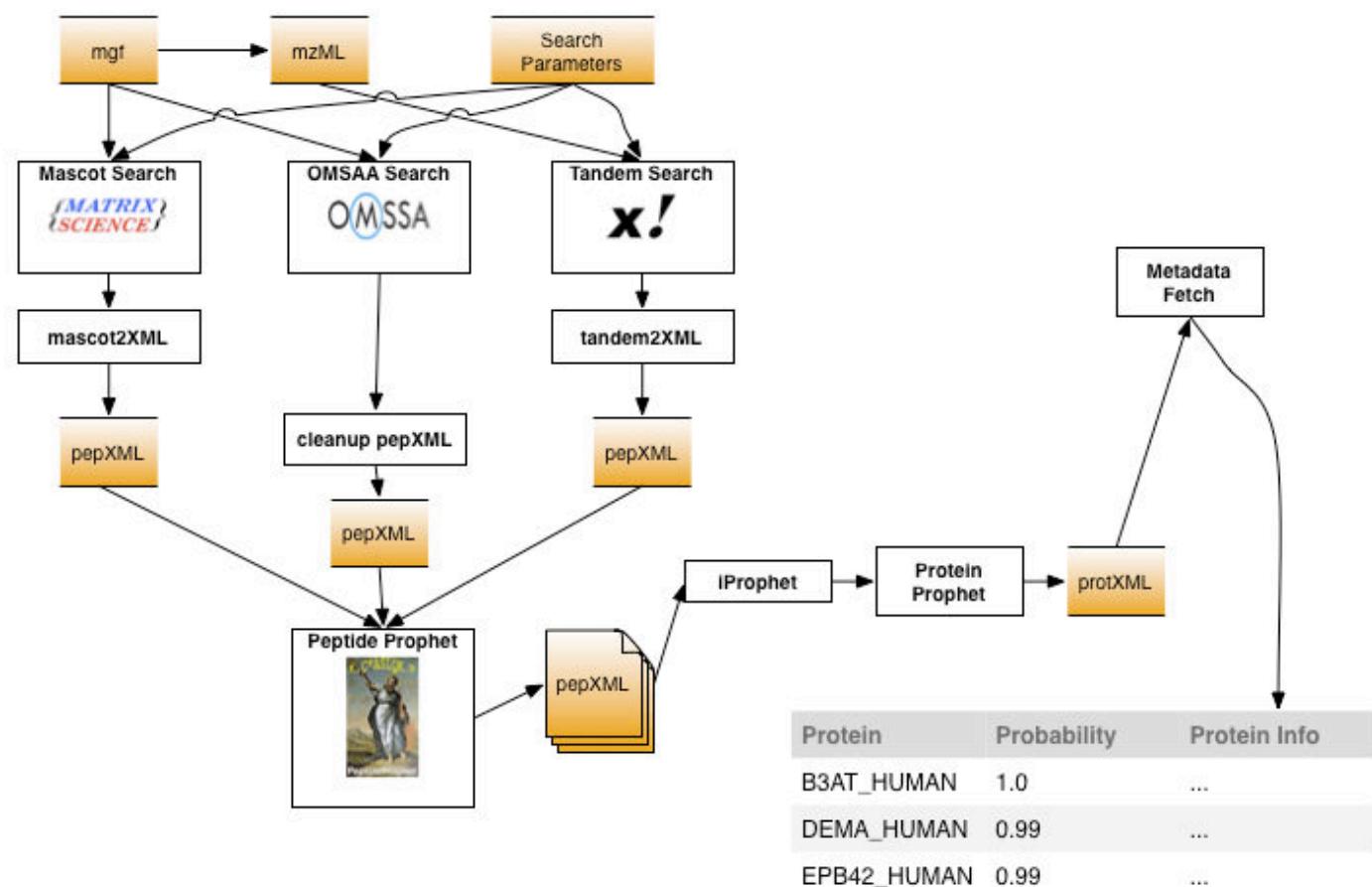


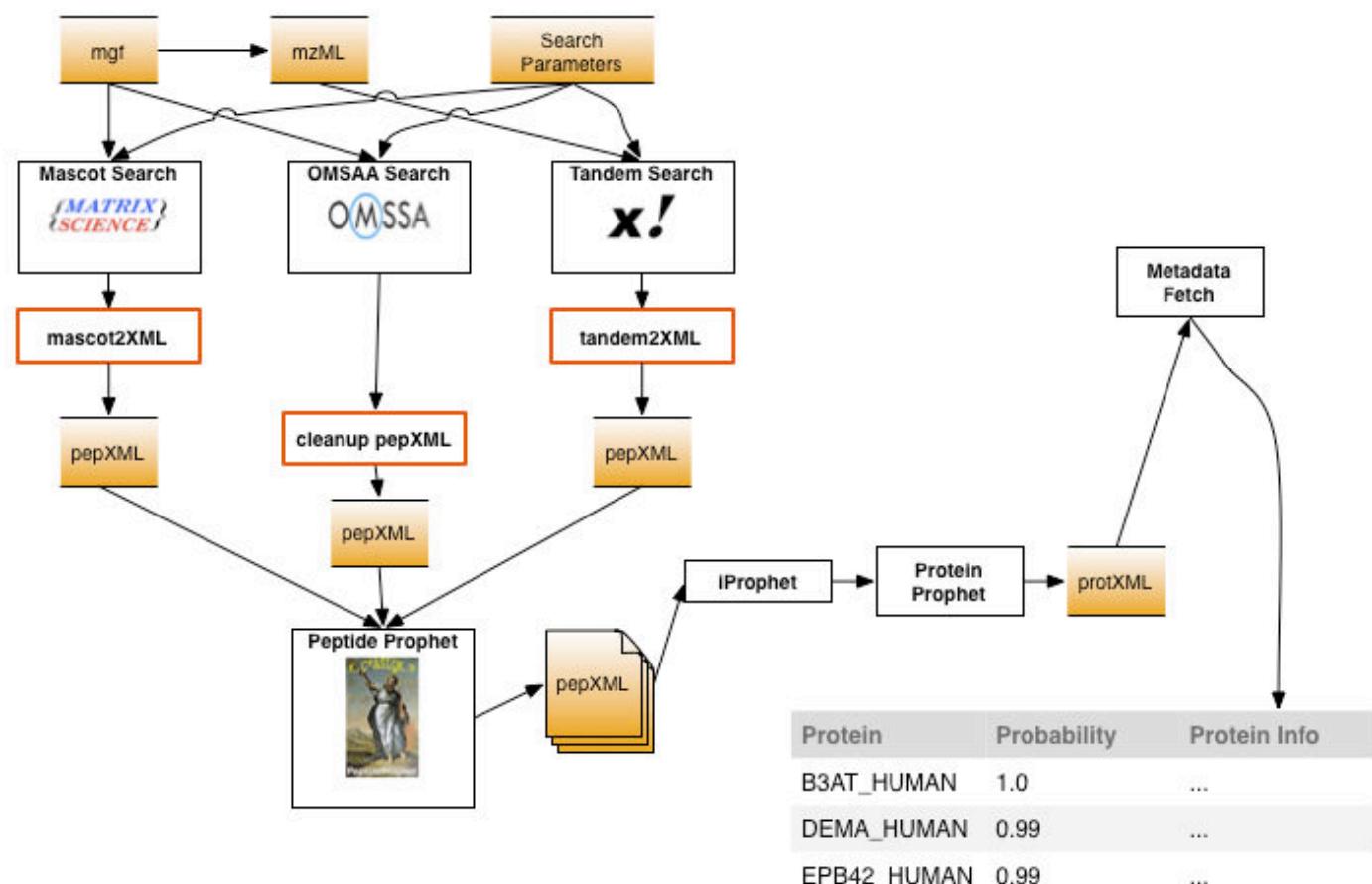
Protein	Probability	Protein Info
B3AT_HUMAN	1.0	...
DEMA_HUMAN	0.99	...
EPB42_HUMAN	0.99	...

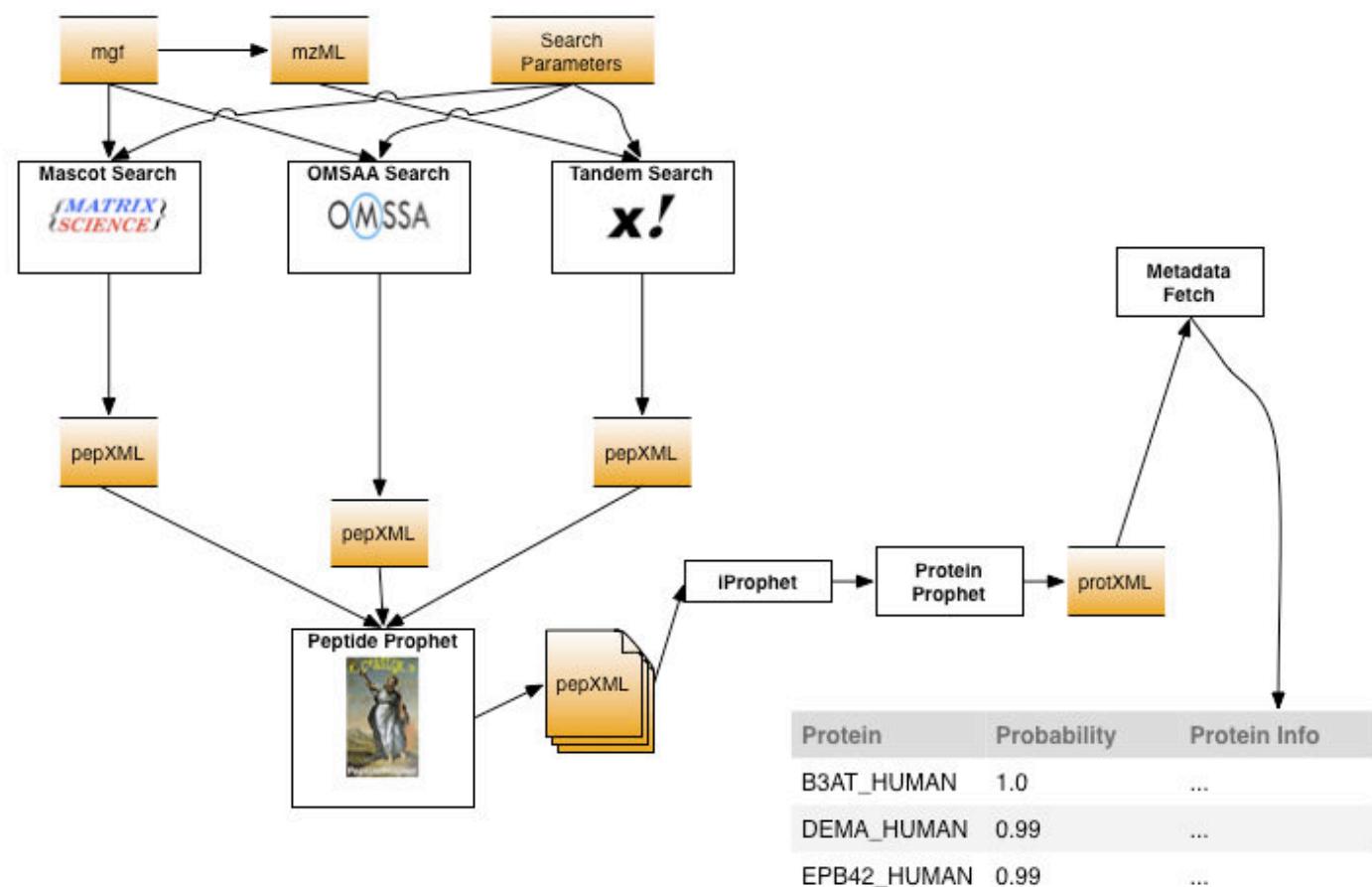


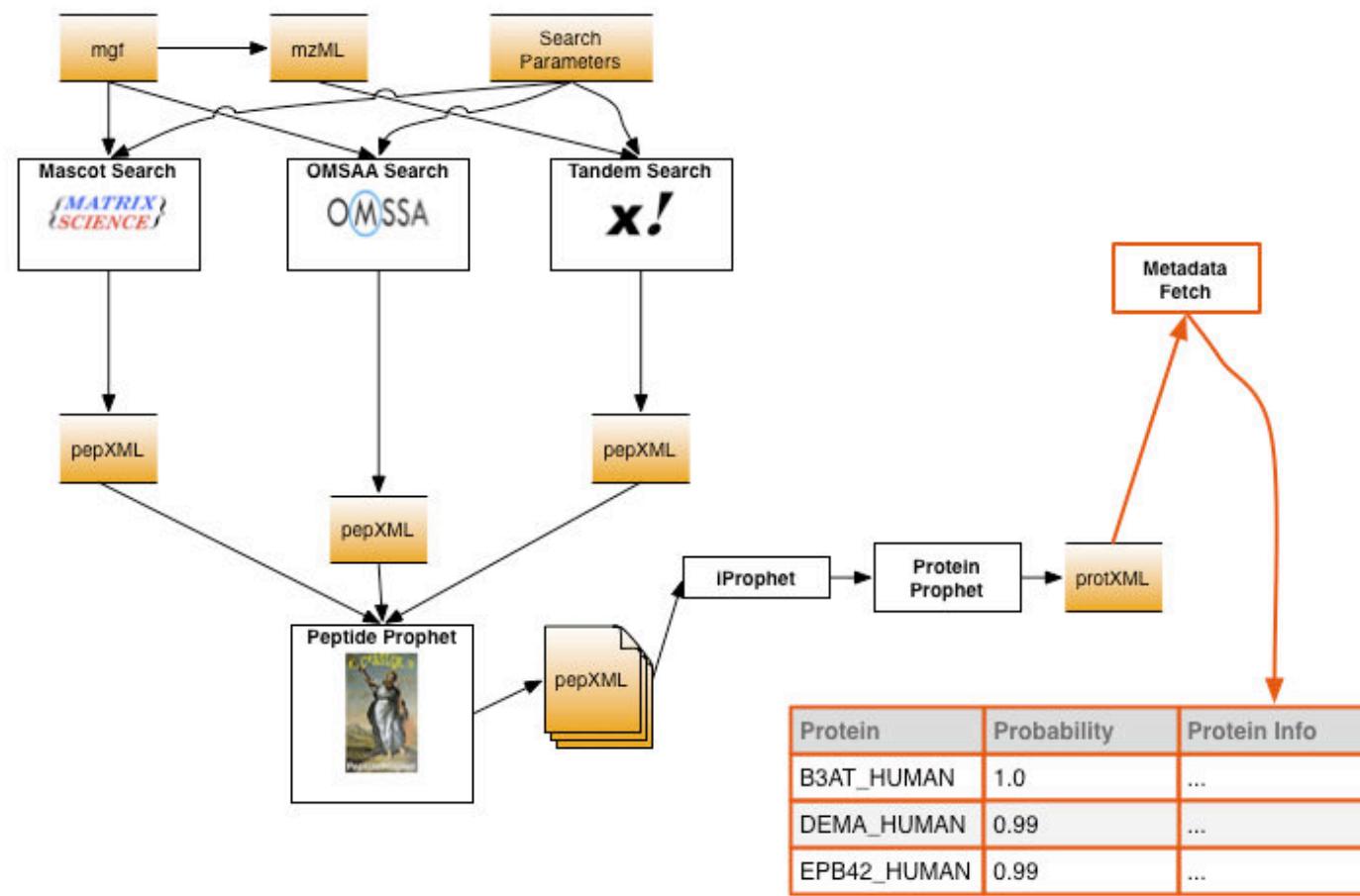


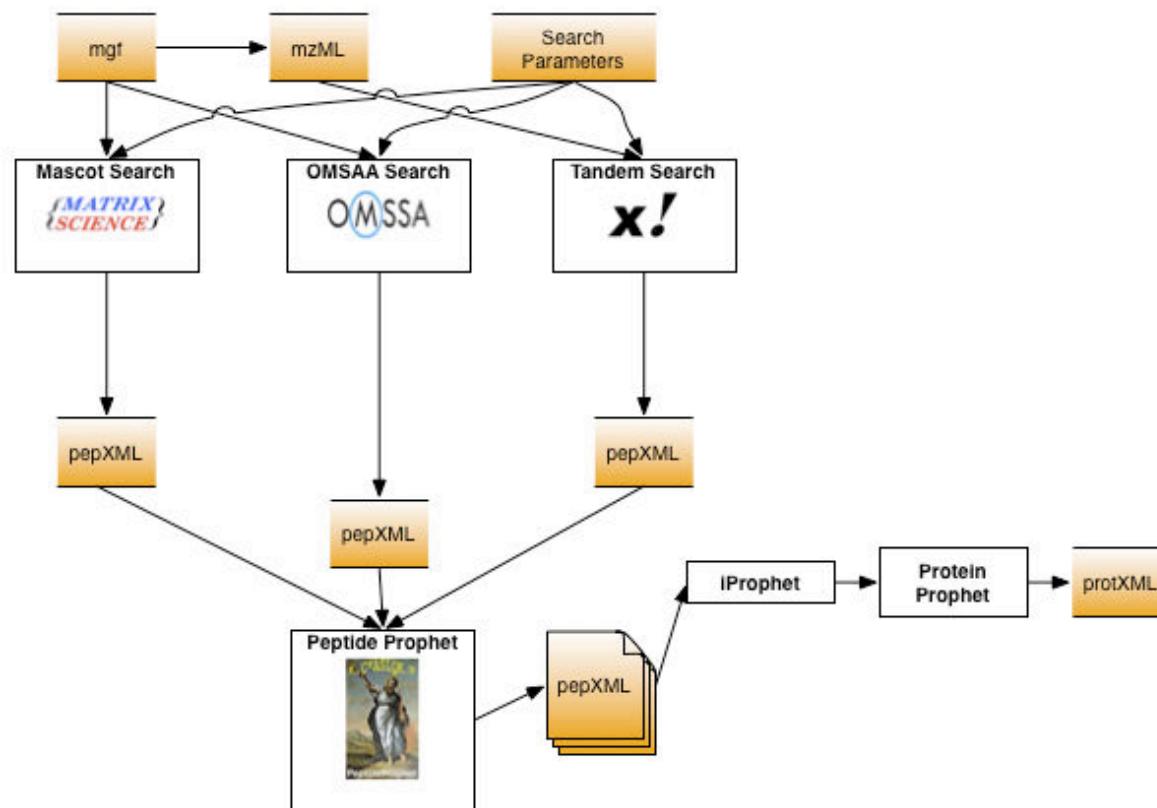












# Proteomics Toolkit

## “protk”

# Proteomics Toolkit

## “protk”

TPP

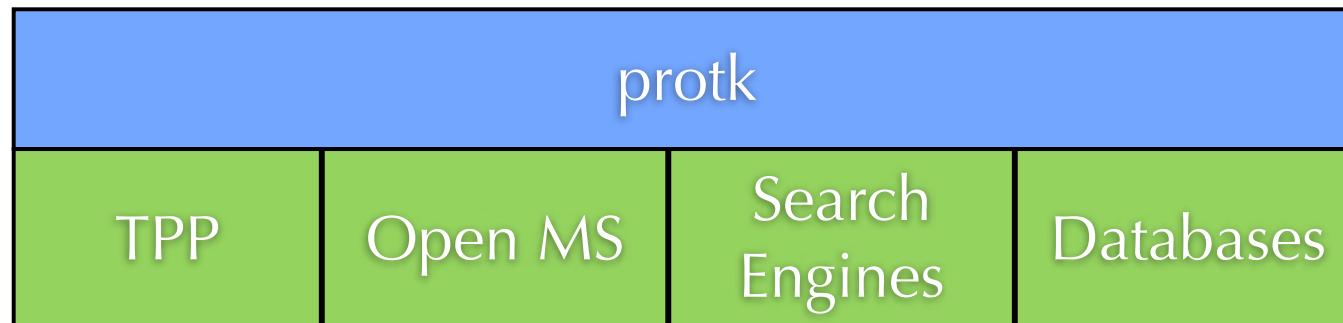
Open MS

Search  
Engines

Databases

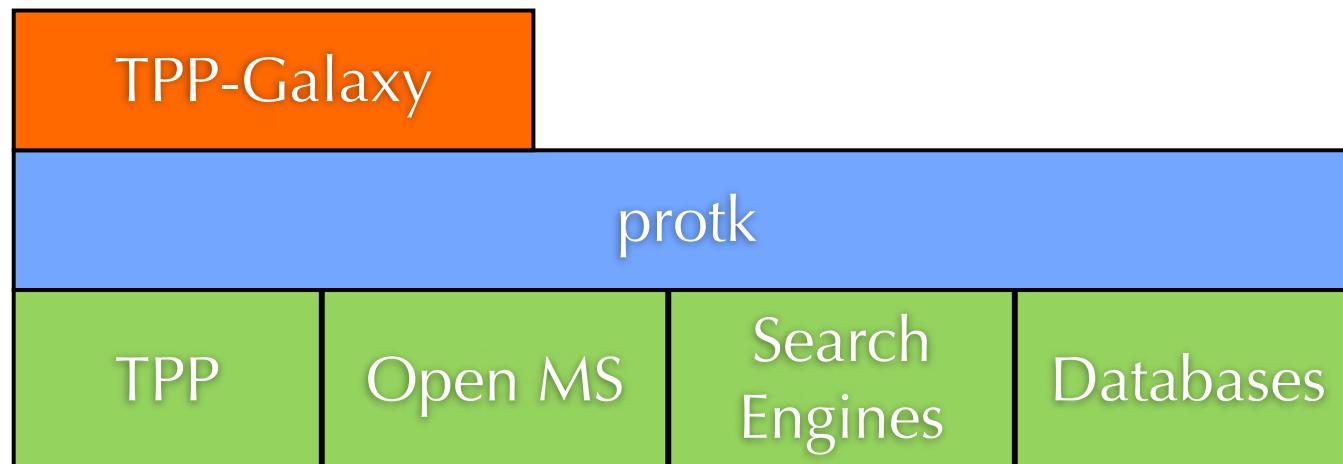
# Proteomics Toolkit

## “protk”



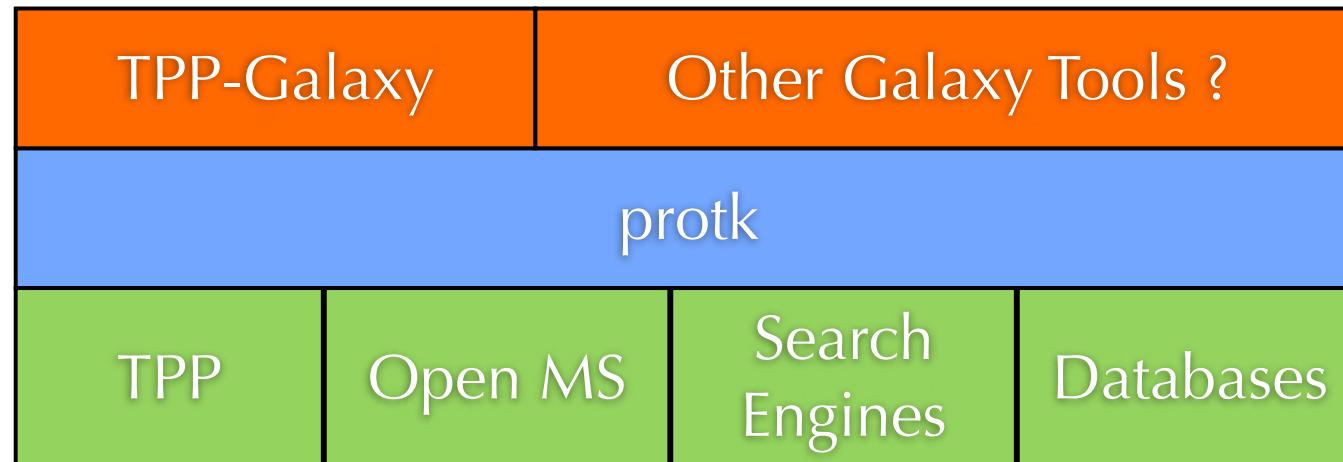
# Proteomics Toolkit

## “protk”



# Proteomics Toolkit

## “protk”



211.55.24/root

W Rede - Wikipedia, the free encyclopedia ProtVis Results

Read Later Welcome to Sage Galaxy Galaxy-dev Ion Torrent Systems ms-utils.org - Software biomaRt and Genom Other Bookmarks Using 2.7 Gb

Analyze Data Workflow Shared Data Admin Help User

Displaying all 31 results

Protein	Peptides	Coverage	Probability																																																												
41_HUMAN	HSNLMLEDLDSQEEIKK <a href="#">2 more</a>	5.7%	1.0																																																												
B3AT_HUMAN	GTVLLDLQETSLAGVANQLLDR <a href="#">1 more</a>	4.5%	1.0																																																												
<a href="#">Hide</a>																																																															
SPTA1_HUMAN	<table border="1"> <thead> <tr> <th>Peptide</th> <th>Initial Probability</th> <th>Position</th> <th>Length</th> </tr> </thead> <tbody> <tr><td>C[160]QALSAADPGSDLFSVQALQR</td><td>0.9989</td><td>1202</td><td>21</td></tr> <tr><td>DLNTLAEDLLSSGTFNVNDQIVK</td><td>0.999</td><td>1664</td><td>22</td></tr> <tr><td>DTEDEEAWIQTETPSATSTYLGK</td><td>0.9988</td><td>800</td><td>23</td></tr> <tr><td>GLNYYLPMVEEDEHEPKFEK</td><td>0.9989</td><td>2301</td><td>20</td></tr> <tr><td>LTLSHPSDAPQIQEMKEDLVSSWEHIR</td><td>0.9989</td><td>329</td><td>27</td></tr> <tr><td>TGQEMIEGGHYASDNVTTR</td><td>0.999</td><td>642</td><td>19</td></tr> <tr><td>VVEVNQYANEC[160]AEENHPDLPLIQSKE</td><td>0.999</td><td>213</td><td>25</td></tr> <tr><td>AYFLDGSSLK</td><td>0.9984</td><td>2177</td><td>10</td></tr> <tr><td>EKEPVLTSEDYGKDLVASEGLFHSHK</td><td>0.9975</td><td>281</td><td>26</td></tr> <tr><td>VADDLLFEGLLTPEGQAQIR</td><td>0.997</td><td>1139</td><td>19</td></tr> <tr><td>QEAFLENEIDLGNISLGSAEALLQK</td><td>0.9948</td><td>494</td><td>23</td></tr> <tr><td>LPEITDLKDKLISAQHNQSK</td><td>0.9856</td><td>1984</td><td>20</td></tr> <tr><td>QNDLEANVQFOQYLAIDLHEAFTWIR</td><td>0.8468</td><td>891</td><td>25</td></tr> <tr><td>FQSADETQQLVNNANHEASDEVR</td><td>0.7494</td><td>425</td><td>23</td></tr> </tbody> </table>	Peptide	Initial Probability	Position	Length	C[160]QALSAADPGSDLFSVQALQR	0.9989	1202	21	DLNTLAEDLLSSGTFNVNDQIVK	0.999	1664	22	DTEDEEAWIQTETPSATSTYLGK	0.9988	800	23	GLNYYLPMVEEDEHEPKFEK	0.9989	2301	20	LTLSHPSDAPQIQEMKEDLVSSWEHIR	0.9989	329	27	TGQEMIEGGHYASDNVTTR	0.999	642	19	VVEVNQYANEC[160]AEENHPDLPLIQSKE	0.999	213	25	AYFLDGSSLK	0.9984	2177	10	EKEPVLTSEDYGKDLVASEGLFHSHK	0.9975	281	26	VADDLLFEGLLTPEGQAQIR	0.997	1139	19	QEAFLENEIDLGNISLGSAEALLQK	0.9948	494	23	LPEITDLKDKLISAQHNQSK	0.9856	1984	20	QNDLEANVQFOQYLAIDLHEAFTWIR	0.8468	891	25	FQSADETQQLVNNANHEASDEVR	0.7494	425	23	12.5%	1.0
	Peptide	Initial Probability	Position	Length																																																											
	C[160]QALSAADPGSDLFSVQALQR	0.9989	1202	21																																																											
	DLNTLAEDLLSSGTFNVNDQIVK	0.999	1664	22																																																											
	DTEDEEAWIQTETPSATSTYLGK	0.9988	800	23																																																											
	GLNYYLPMVEEDEHEPKFEK	0.9989	2301	20																																																											
	LTLSHPSDAPQIQEMKEDLVSSWEHIR	0.9989	329	27																																																											
	TGQEMIEGGHYASDNVTTR	0.999	642	19																																																											
	VVEVNQYANEC[160]AEENHPDLPLIQSKE	0.999	213	25																																																											
	AYFLDGSSLK	0.9984	2177	10																																																											
	EKEPVLTSEDYGKDLVASEGLFHSHK	0.9975	281	26																																																											
	VADDLLFEGLLTPEGQAQIR	0.997	1139	19																																																											
	QEAFLENEIDLGNISLGSAEALLQK	0.9948	494	23																																																											
	LPEITDLKDKLISAQHNQSK	0.9856	1984	20																																																											
	QNDLEANVQFOQYLAIDLHEAFTWIR	0.8468	891	25																																																											
FQSADETQQLVNNANHEASDEVR	0.7494	425	23																																																												
ACTB_HUMAN +1	TTGIVMDSGDGVTHTVPIYEGYALPHAILR <a href="#">1 more</a>	10.7%	1.0																																																												
ANK1_HUMAN	GFTPLYMAAQENHLEVVK <a href="#">5 more</a>	6.8%	1.0																																																												
SPTB1_HUMAN	ALEDLAQELEKENYHDQKR <a href="#">14 more</a>	13.9%	1.0																																																												
DEMA_HUMAN	HLSAEDFSR <a href="#">Show coverage</a>	2.2%	0.993																																																												
HBB_HUMAN +1	LHVDPENFR <a href="#">Show coverage</a>	6.1%	0.9841																																																												
CNPY4_HUMAN	F[111]HTDCEEKTEGEFFFQQEEREEEEECC <a href="#">Show coverage</a>	10.0%	0.9239																																																												

History

Test 45.8 MB

37: protein\_prophet.interprophet.peptide\_prophet.omssa vs sphuman.MT669- TiloS3 BAF.d\_pk.mzML.mgf.pepXML.pep.xml.protXML prot XML Search Results format: protxml, database: ? view protXML in Proteomics Visualize <?xml version="1.0" encoding="UTF-8"?> <protein\_summary xmlns="http://registerjob/working\_directory/000/37/proteins</protein\_summary\_header reference\_data /database/files/000/dataset\_36.dat.pep initial\_min\_peptide\_prob="0.05" total\_</protein\_summary>

36: interprophet.peptide\_prophet.omssa vs sphuman.MT669- TiloS3 BAF.d\_pk.mzML.mgf.pepXML.pep.xml

35: peptide\_prophet.omssa vs sphuman.MT669- TiloS3 BAF.d\_pk.mzML.mgf.pepXML.pep.xml

211.55.24/root

W Rede - Wikipedia, the free encyclopedia ProtVis Results

Read Later Welcome to Sage Galaxy Galaxy-dev Ion Torrent Systems ms-utils.org - Software biomaRt and Genom Other Bookmarks Using 2.7 Gb

Analyze Data Workflow Shared Data Admin Help User

Displaying all 31 results

Protein	Peptides	Coverage	Probability																																																												
41_HUMAN	HSNLMLEDLDSQEEIKK <a href="#">2 more</a>	5.7%	1.0																																																												
B3AT_HUMAN	GTVLLDLQETSLAGVANQLLDR <a href="#">1 more</a>	4.5%	1.0																																																												
<a href="#">Hide</a>																																																															
SPTA1_HUMAN	<table border="1"> <thead> <tr> <th>Peptide</th> <th>Initial Probability</th> <th>Position</th> <th>Length</th> </tr> </thead> <tbody> <tr><td>C[160]QALSAADPGSDLFSVQALQR</td><td>0.9989</td><td>1202</td><td>21</td></tr> <tr><td>DLNTLAEDLLSSGTFNVNDQIVK</td><td>0.999</td><td>1664</td><td>22</td></tr> <tr><td>DTEDEEAWIQTETPSATSTYLGK</td><td>0.9988</td><td>800</td><td>23</td></tr> <tr><td>GLNYYLPMVEEDEHEPKFEK</td><td>0.9989</td><td>2301</td><td>20</td></tr> <tr><td>LTLSHPSDAPQIQEMKEDLVSSWEHIR</td><td>0.9989</td><td>329</td><td>27</td></tr> <tr><td>TGQEMIEGGHYASDNVTTR</td><td>0.999</td><td>642</td><td>19</td></tr> <tr><td>VVEVNQYANEC[160]AEENHPDLPLIQSKE</td><td>0.999</td><td>213</td><td>25</td></tr> <tr><td>AYFLDGSSLK</td><td>0.9984</td><td>2177</td><td>10</td></tr> <tr><td>EKEPVLTSEDYGKDLVASEGLFHSHK</td><td>0.9975</td><td>281</td><td>26</td></tr> <tr><td>VADDLLFEGLLTPEGAQIR</td><td>0.997</td><td>1139</td><td>19</td></tr> <tr><td>QEAFLENEIDLGNISLGSAEALLQK</td><td>0.9948</td><td>494</td><td>23</td></tr> <tr><td>LPEITDLKDKLISAQHNQSK</td><td>0.9856</td><td>1984</td><td>20</td></tr> <tr><td>QNDLEANVQFOQYLAIDLHEAFTWIR</td><td>0.8468</td><td>891</td><td>25</td></tr> <tr><td>FQSADETQQLVNNANHEASDEVR</td><td>0.7494</td><td>425</td><td>23</td></tr> </tbody> </table>	Peptide	Initial Probability	Position	Length	C[160]QALSAADPGSDLFSVQALQR	0.9989	1202	21	DLNTLAEDLLSSGTFNVNDQIVK	0.999	1664	22	DTEDEEAWIQTETPSATSTYLGK	0.9988	800	23	GLNYYLPMVEEDEHEPKFEK	0.9989	2301	20	LTLSHPSDAPQIQEMKEDLVSSWEHIR	0.9989	329	27	TGQEMIEGGHYASDNVTTR	0.999	642	19	VVEVNQYANEC[160]AEENHPDLPLIQSKE	0.999	213	25	AYFLDGSSLK	0.9984	2177	10	EKEPVLTSEDYGKDLVASEGLFHSHK	0.9975	281	26	VADDLLFEGLLTPEGAQIR	0.997	1139	19	QEAFLENEIDLGNISLGSAEALLQK	0.9948	494	23	LPEITDLKDKLISAQHNQSK	0.9856	1984	20	QNDLEANVQFOQYLAIDLHEAFTWIR	0.8468	891	25	FQSADETQQLVNNANHEASDEVR	0.7494	425	23	12.5%	1.0
	Peptide	Initial Probability	Position	Length																																																											
	C[160]QALSAADPGSDLFSVQALQR	0.9989	1202	21																																																											
	DLNTLAEDLLSSGTFNVNDQIVK	0.999	1664	22																																																											
	DTEDEEAWIQTETPSATSTYLGK	0.9988	800	23																																																											
	GLNYYLPMVEEDEHEPKFEK	0.9989	2301	20																																																											
	LTLSHPSDAPQIQEMKEDLVSSWEHIR	0.9989	329	27																																																											
	TGQEMIEGGHYASDNVTTR	0.999	642	19																																																											
	VVEVNQYANEC[160]AEENHPDLPLIQSKE	0.999	213	25																																																											
	AYFLDGSSLK	0.9984	2177	10																																																											
	EKEPVLTSEDYGKDLVASEGLFHSHK	0.9975	281	26																																																											
	VADDLLFEGLLTPEGAQIR	0.997	1139	19																																																											
	QEAFLENEIDLGNISLGSAEALLQK	0.9948	494	23																																																											
	LPEITDLKDKLISAQHNQSK	0.9856	1984	20																																																											
	QNDLEANVQFOQYLAIDLHEAFTWIR	0.8468	891	25																																																											
FQSADETQQLVNNANHEASDEVR	0.7494	425	23																																																												
ACTB_HUMAN +1	TTGIVMDSGDGVTHTVPIYEGYALPHAILR <a href="#">1 more</a>	10.7%	1.0																																																												
ANK1_HUMAN	GFTPLYMAAQENHLEVVK <a href="#">5 more</a>	6.8%	1.0																																																												
SPTB1_HUMAN	ALEDLAQELEKENYHDQKR <a href="#">14 more</a>	13.9%	1.0																																																												
DEMA_HUMAN	HLSAEDFSR <a href="#">Show coverage</a>	2.2%	0.993																																																												
HBB_HUMAN +1	LHVDPENFR <a href="#">Show coverage</a>	6.1%	0.9841																																																												
CNPY4_HUMAN	F[111]HTDCEEKTEGEFFFQQEEREEEEECC <a href="#">Show coverage</a>	10.0%	0.9239																																																												

History

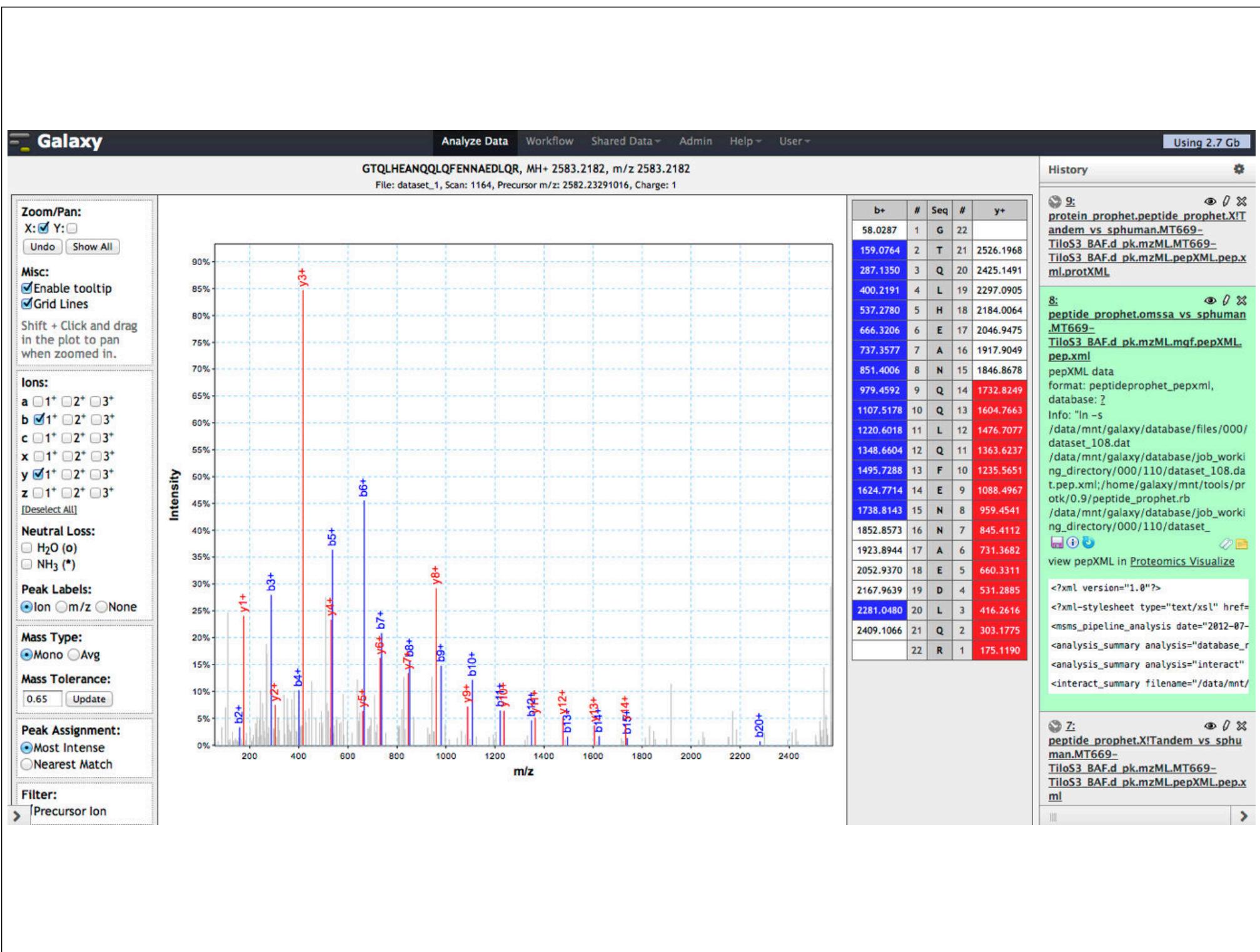
Test 45.8 MB

37: protein\_prophet.interprophet.peptide\_prophet.omssa vs sphuman.MT669- TiloS3 BAF.d\_pk.mzML.mgf.pepXML.pep.xml.protXML prot XML Search Results format: protxml, database: ? view protXML in Proteomics Visualize

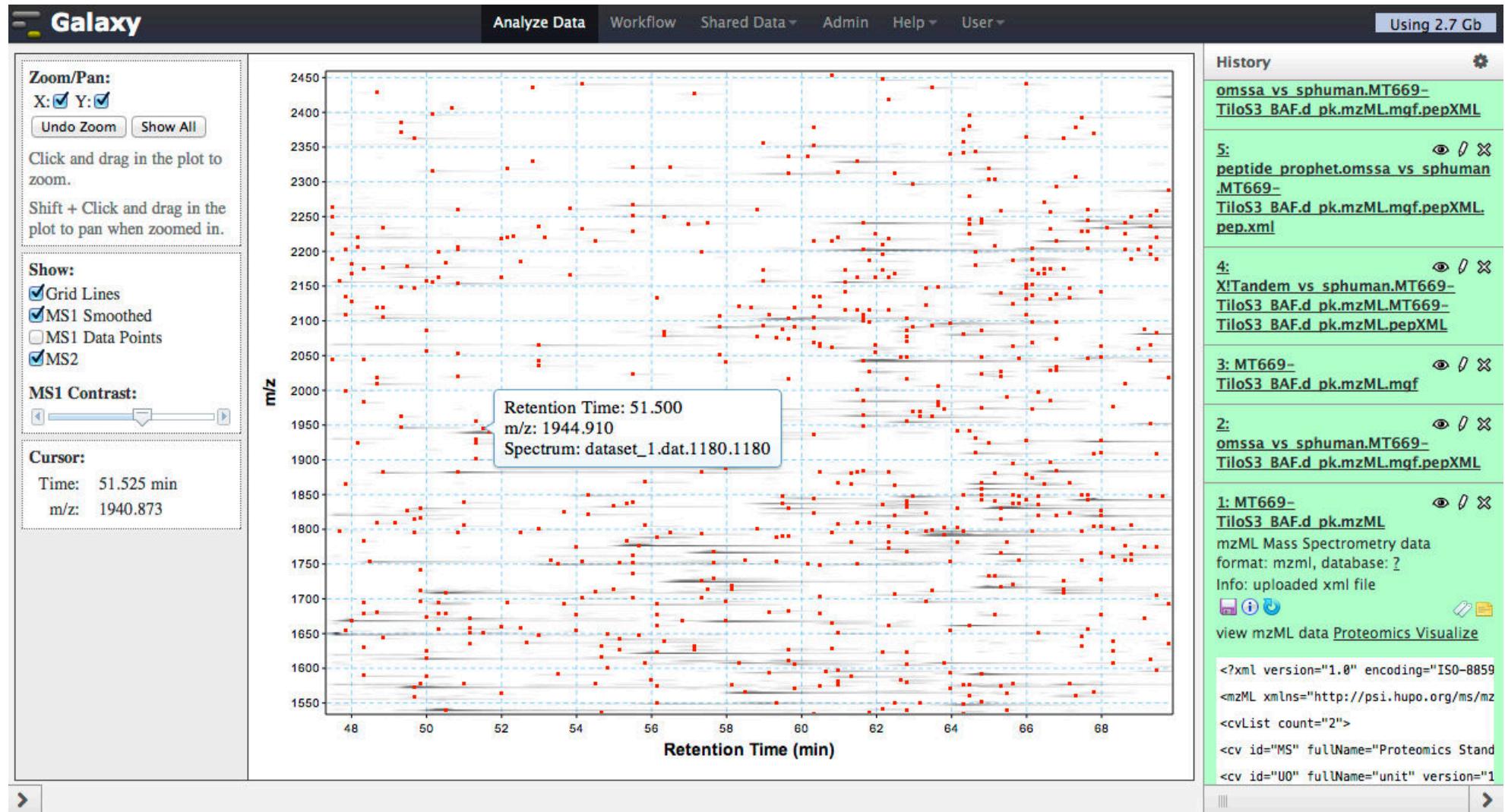
```
<?xml version="1.0" encoding="UTF-8"?>
<protein_summary>
<!-- regis-->
<!-- job_working_dir --> 000/37/prote
<protein_summary_header>ference_data
</protein_summary_header> /database/files/000/t_36.dat.pep
initial_min_peptide_prob="0.05" total_
```

36: interprophet.peptide\_prophet.omssa vs sphuman.MT669- TiloS3 BAF.d\_pk.mzML.mgf.pepXML.pep.xml

35: peptide\_prophet.omssa vs sphuman.MT669- TiloS3 BAF.d\_pk.mzML.mgf.pepXML







# Genomics and Proteomics

# Genomics and Proteomics

- Proteomics can be used to improve genome annotation. “Proteo-genomics”

# Genomics and Proteomics

- Proteomics can be used to improve genome annotation. “Proteo-genomics”  
Verify Expression

# Genomics and Proteomics

- Proteomics can be used to improve genome annotation. “Proteo-genomics”
  - Verify Expression
  - Correct gene predictions

# Genomics and Proteomics

- Proteomics can be used to improve genome annotation. “Proteo-genomics”

Verify Expression

Correct gene predictions

Find new genes

# Genomics and Proteomics

- Proteomics can be used to improve genome annotation. “Proteo-genomics”
  - Verify Expression
  - Correct gene predictions
  - Find new genes
- Genomics workflows are used to make protein databases

# Genomics and Proteomics

- Proteomics can be used to improve genome annotation. “Proteo-genomics”
  - Verify Expression
  - Correct gene predictions
  - Find new genes
- Genomics workflows are used to make protein databases
  - Emerging model organisms

# Genomics and Proteomics

- Proteomics can be used to improve genome annotation. “Proteo-genomics”
  - Verify Expression
  - Correct gene predictions
  - Find new genes
- Genomics workflows are used to make protein databases
  - Emerging model organisms
  - Variant analysis

# Where to get the tools

- Under “Proteomics” on the main toolshed
- <https://bitbucket.org/iracooke/protk>
- [https://bitbucket.org/Andrew\\_Brock/  
proteomics-visualise](https://bitbucket.org/Andrew_Brock/proteomics-visualise)
- [https://bitbucket.org/jmchilton/galaxy-p-vm-  
launcher](https://bitbucket.org/jmchilton/galaxy-p-vm-launcher)

# Acknowledgements

## ★ Mass Spec & Biochemistry

Fiona Durand, Pierre Faou, Vita Levina, Gert Talbo,  
Nick Hoogenraad, Hoogenraad Lab Members

## ★ Programming

Andrew Brock (Protvis), John Chilton

## ★ Funding

VLSCI, e-Research, CRC for Biomarker translation

