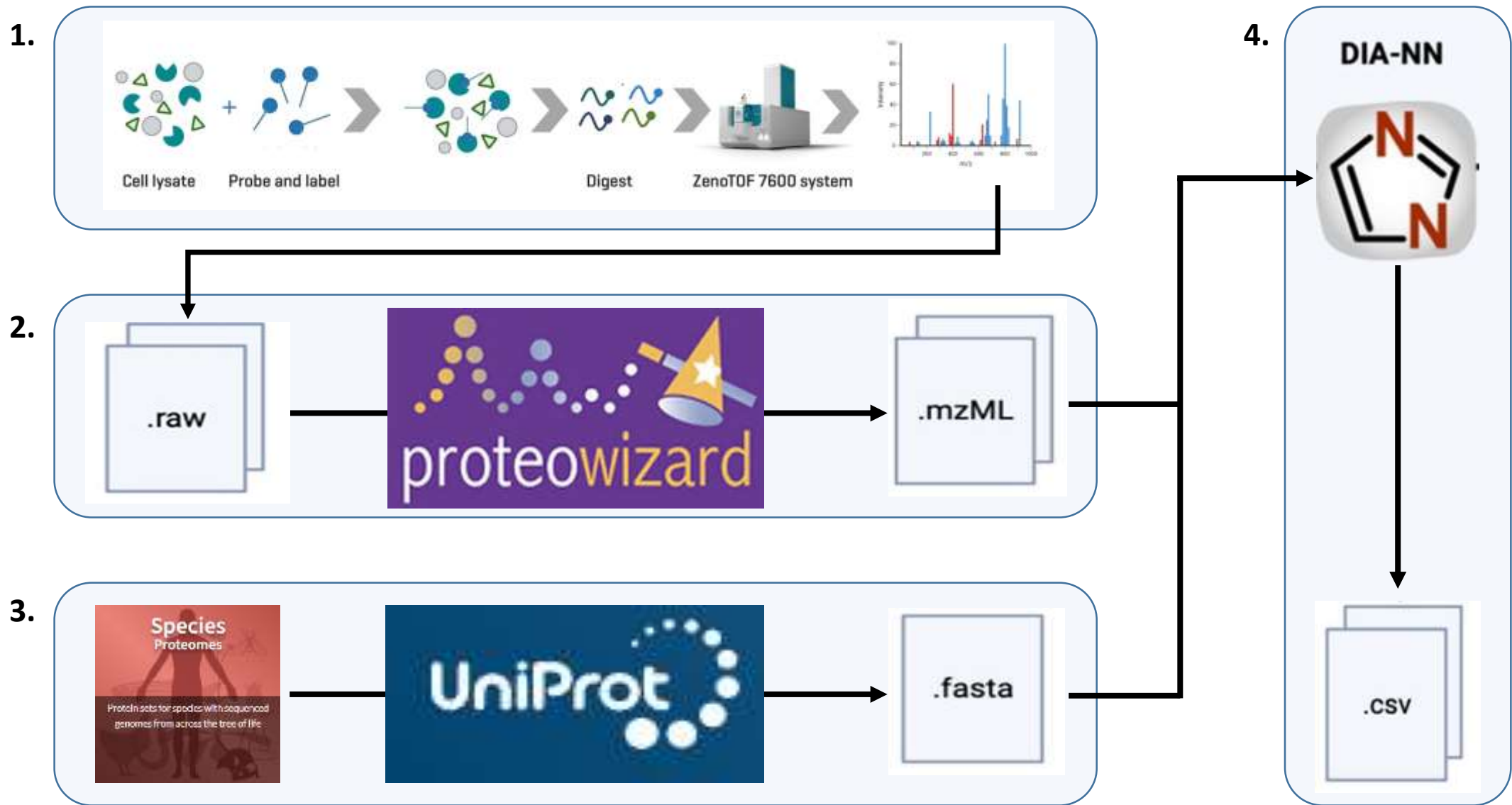


Data conversion and Database search

MS-facility



Raw data conversion

Proteowizard

Proteowizard Download

The screenshot shows the ProteoWizard download page. At the top, there is a navigation menu with links for Projects, Download, Publications, Team, Documentation, and Support. The main heading is "Download". Below this, there are three dropdown menus: "Project" (set to ProteoWizard), "Platform" (set to Windows 32-bit installer (able to convert vendor files)), and "Email (optional)". There is a checkbox for "I have read and accepted the license agreements" which is checked. Below the forms are two buttons: "DOWNLOAD" and "DONATE". To the right of the forms, there is text asking for citations and providing a reference: "Chambers, M.C., MacLean, B., ... Mallick, P. Nature Biotechnology 30, 918-920 (2012). Article".

Download

Project: ProteoWizard

Platform: Windows 32-bit installer (able to convert vendor files)

Email (optional):

I have read and accepted the [license agreements](#)

Please cite us in your publications

A cross-platform toolkit for mass spectrometry and proteomics
Chambers, M.C., MacLean, B., ... Mallick, P. Nature
Biotechnology 30, 918-920 (2012). [Article](#) →

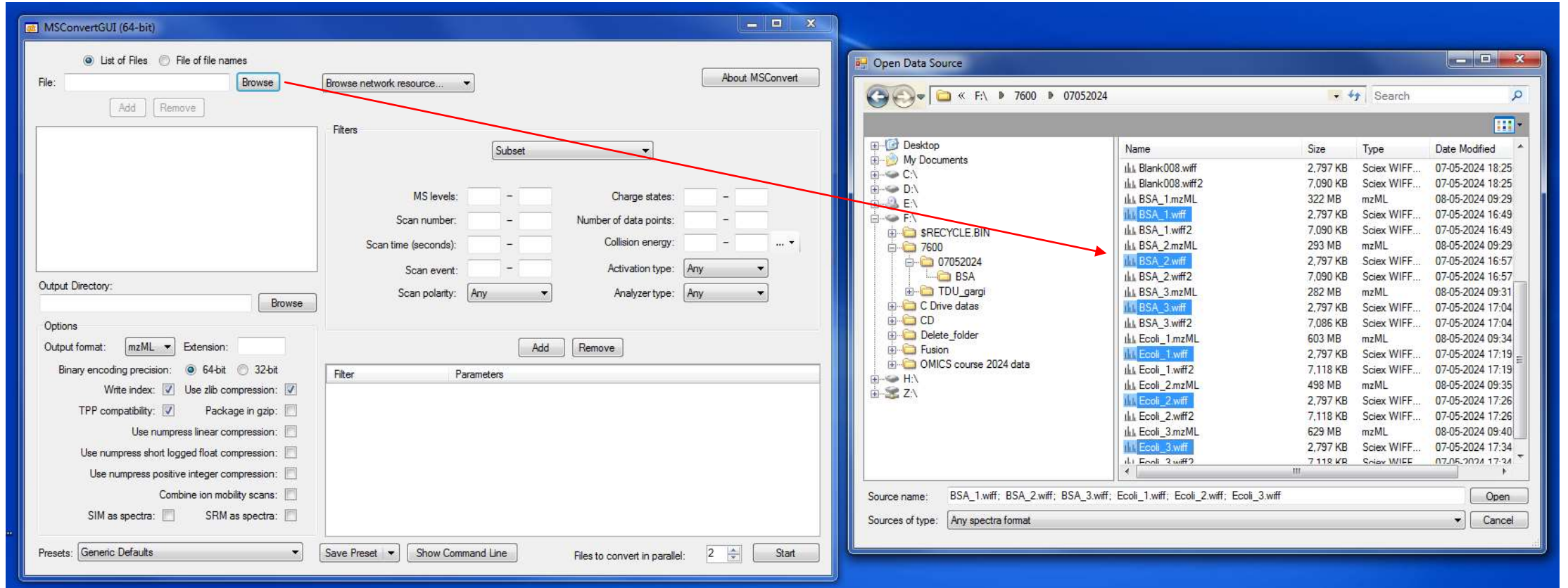
If you would like to support the efforts of ProteoWizard,
donations are happily accepted. Thank you for your support!

Installation

Windows Users: The only requirement is [.NET Framework 4.7.2](#) (or newer). If you get a message saying .NET 3.5 is required on Windows 8, 8.1, or 10, enable it with [this process](#).

Linux Users: There are two options for using ProteoWizard on Linux: native executables or running the Windows executables via Wine/Docker. Only the Wine/Docker option allows converting vendor formats. For Docker instructions, see [here](#). For a native installation, download the native binary tarball, navigate to the download directory. Once there, use 'tar xvf [archive name].tar.bz2' on the archive and move the pwiz directory into a directory of your choice. PRO-TIP: For convenience, adjust your "PATH" environment variable to include your ProteoWizard tools directory, so that the command line tools may be run from any directory.

MSConvert-GUI



The image shows the MSConvertGUI (64-bit) interface on the left and an Open Data Source dialog box on the right. A red arrow points from the 'Browse' button in the MSConvertGUI to the file list in the Open Data Source dialog.

MSConvertGUI (64-bit) Interface:

- File selection: List of Files, File of file names. File: Browse. Browse network resource...
- Filters: Subset (dropdown)
- MS levels: - Charge states: -
- Scan number: - Number of data points: -
- Scan time (seconds): - Collision energy: -
- Scan event: - Activation type: Any (dropdown)
- Scan polarity: Any (dropdown) Analyzer type: Any (dropdown)
- Output Directory: Browse
- Options:
 - Output format: mzML (dropdown) Extension:
 - Binary encoding precision: 64-bit 32-bit
 - Write index: Use zlib compression:
 - TPP compatibility: Package in gzip:
 - Use numpress linear compression:
 - Use numpress short logged float compression:
 - Use numpress positive integer compression:
 - Combine ion mobility scans:
 - SIM as spectra: SRM as spectra:
- Presets: Generic Defaults (dropdown) Save Preset Show Command Line
- Files to convert in parallel: 2 Start

Open Data Source Dialog:

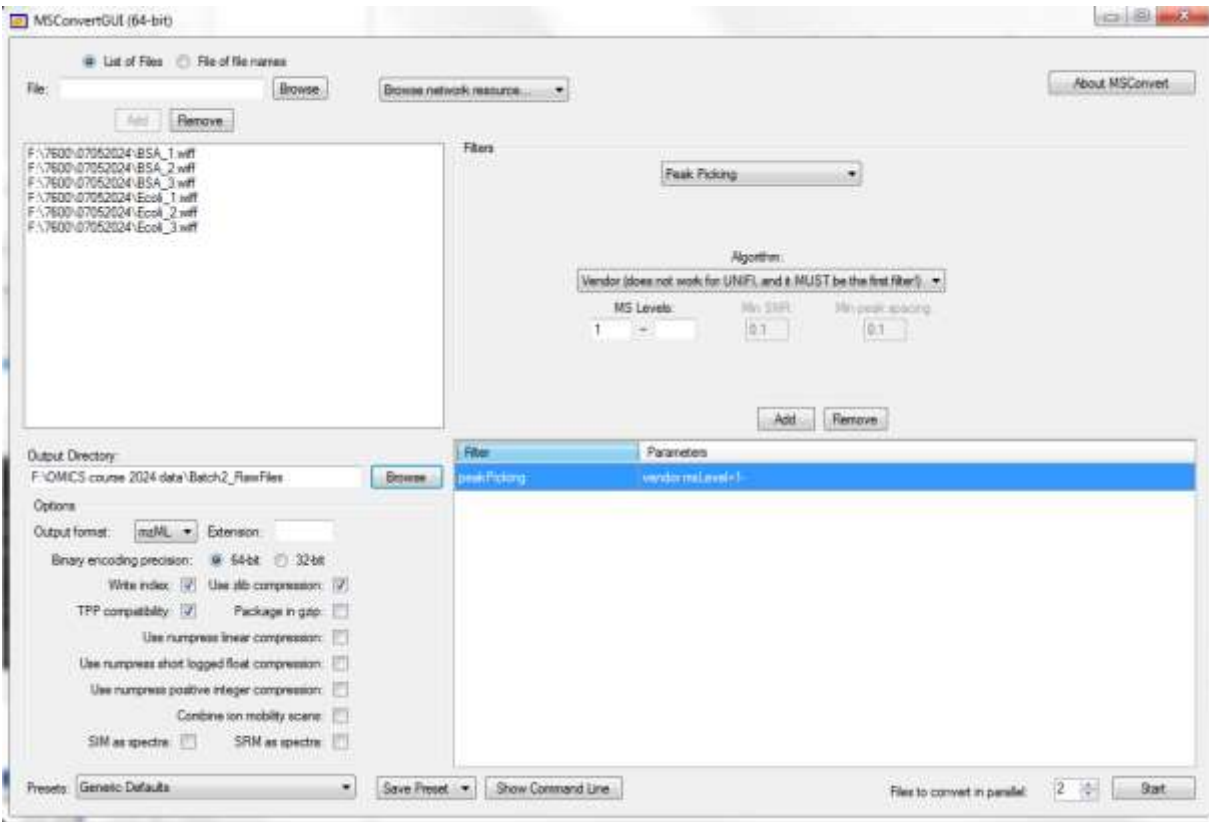
Path: F:\ 7600 \ 07052024

Name	Size	Type	Date Modified
Blank008.wiff	2,797 KB	Sciex WIFF...	07-05-2024 18:25
Blank008.wiff2	7,090 KB	Sciex WIFF...	07-05-2024 18:25
BSA_1.mzML	322 MB	mzML	08-05-2024 09:29
BSA_1.wiff	2,797 KB	Sciex WIFF...	07-05-2024 16:49
BSA_1.wiff2	7,090 KB	Sciex WIFF...	07-05-2024 16:49
BSA_2.mzML	293 MB	mzML	08-05-2024 09:29
BSA_2.wiff	2,797 KB	Sciex WIFF...	07-05-2024 16:57
BSA_2.wiff2	7,090 KB	Sciex WIFF...	07-05-2024 16:57
BSA_3.mzML	282 MB	mzML	08-05-2024 09:31
BSA_3.wiff	2,797 KB	Sciex WIFF...	07-05-2024 17:04
BSA_3.wiff2	7,086 KB	Sciex WIFF...	07-05-2024 17:04
Ecoli_1.mzML	603 MB	mzML	08-05-2024 09:34
Ecoli_1.wiff	2,797 KB	Sciex WIFF...	07-05-2024 17:19
Ecoli_1.wiff2	7,118 KB	Sciex WIFF...	07-05-2024 17:19
Ecoli_2.mzML	498 MB	mzML	08-05-2024 09:35
Ecoli_2.wiff	2,797 KB	Sciex WIFF...	07-05-2024 17:26
Ecoli_2.wiff2	7,118 KB	Sciex WIFF...	07-05-2024 17:26
Ecoli_3.mzML	629 MB	mzML	08-05-2024 09:40
Ecoli_3.wiff	2,797 KB	Sciex WIFF...	07-05-2024 17:34
Ecoli_3.wiff2	7,118 KB	Sciex WIFF...	07-05-2024 17:34

Source name: BSA_1.wiff; BSA_2.wiff; BSA_3.wiff; Ecoli_1.wiff; Ecoli_2.wiff; Ecoli_3.wiff Open

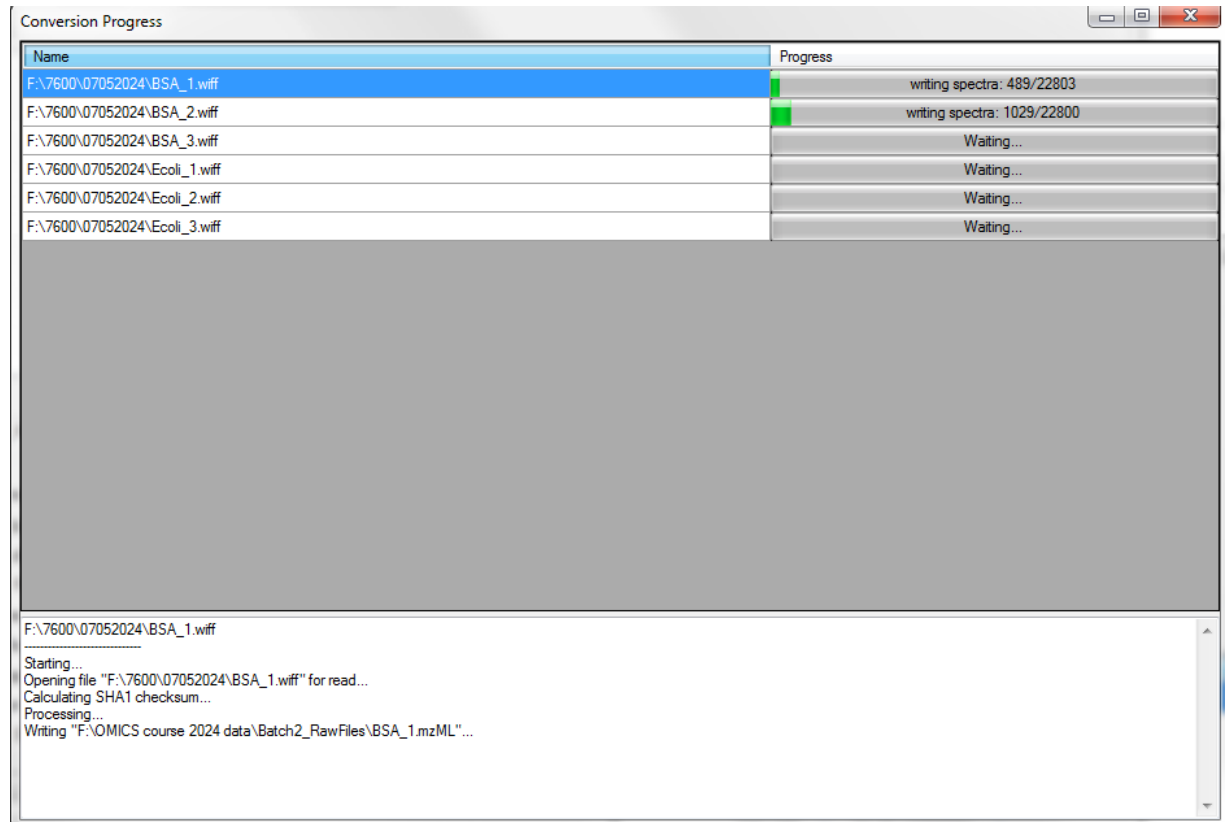
Sources of type: Any spectra format Cancel

Raw file conversion



The MSConvertGUI (64-bit) window displays the following configuration:

- File List:** F:\7600\07052024\BSA_1.wiff, F:\7600\07052024\BSA_2.wiff, F:\7600\07052024\BSA_3.wiff, F:\7600\07052024\Ecoli_1.wiff, F:\7600\07052024\Ecoli_2.wiff, F:\7600\07052024\Ecoli_3.wiff
- Output Directory:** F:\OMICS course 2024 data\Batch2_RawFiles
- Options:**
 - Output format: mzML
 - Binary encoding precision: 64-bit
 - Write index: Use zlib compression:
 - TPP compatibility: Package in gdf:
 - Use numpress linear compression:
 - Use numpress short logged float compression:
 - Use numpress positive integer compression:
 - Combine ion mobility scans:
 - SIM as spectra: SRM as spectra:
- Parameters:**
 - File: peakPicking
 - Parameter: vendor msLevel=1



Name	Progress
F:\7600\07052024\BSA_1.wiff	writing spectra: 489/22803
F:\7600\07052024\BSA_2.wiff	writing spectra: 1029/22800
F:\7600\07052024\BSA_3.wiff	Waiting...
F:\7600\07052024\Ecoli_1.wiff	Waiting...
F:\7600\07052024\Ecoli_2.wiff	Waiting...
F:\7600\07052024\Ecoli_3.wiff	Waiting...

Log output:

```

F:\7600\07052024\BSA_1.wiff
Starting...
Opening file "F:\7600\07052024\BSA_1.wiff" for read...
Calculating SHA1 checksum...
Processing...
Writing "F:\OMICS course 2024 data\Batch2_RawFiles\BSA_1.mzML" ...
  
```

Reference proteome download

Uniprot

Database download from Uniprot

UniProtKB 20,434 results

Entry	Entry Name	Protein Names
Q95905	ECD_HUMAN	Protein ectodysless homolog[...]
P04439	HLAA_HUMAN	HLA class I histocompatibility antigen, A alpha chain[...]
P08246	ELNE_HUMAN	Neutrophil elastase[...]
P43694	HELZ_HUMAN	Probable helicase with zinc finger domain[...]
Q62WK4	RHEX_HUMAN	Regulator of hemoglobinization and erythroid cell expansion protein[...]
Q8FHJ7	SFRP4_HUMAN	Secreted frizzled-related protein 4[...]
P15829	MCP_HUMAN	Membrane cofactor protein[...]
Q08648	SGT1B_HUMAN	Sperm-associated antigen 11B[...]
Q9Y2H6	FN3A_HUMAN	Fibronectin type-III domain-containing protein 3A[...]
P93244	RACK1_HUMAN	Small ribosomal subunit protein RACK1[...]
P78310	CXAR_HUMAN	Coxsackievirus and adenovirus receptor[...]
P01911	DRB1_HUMAN	HLA class II histocompatibility antigen, DRB1 beta chain[...]
P25990	PML_HUMAN	Protein PML[...]
Q89VV6	STING_HUMAN	Stimulator of interferon genes protein[...]
P01889	HLAB_HUMAN	HLA class I histocompatibility antigen, B alpha chain[...]

Download

Download selected (0)

Download all (20,434)

Format

FASTA (canonical)

Compressed¹

Yes

No

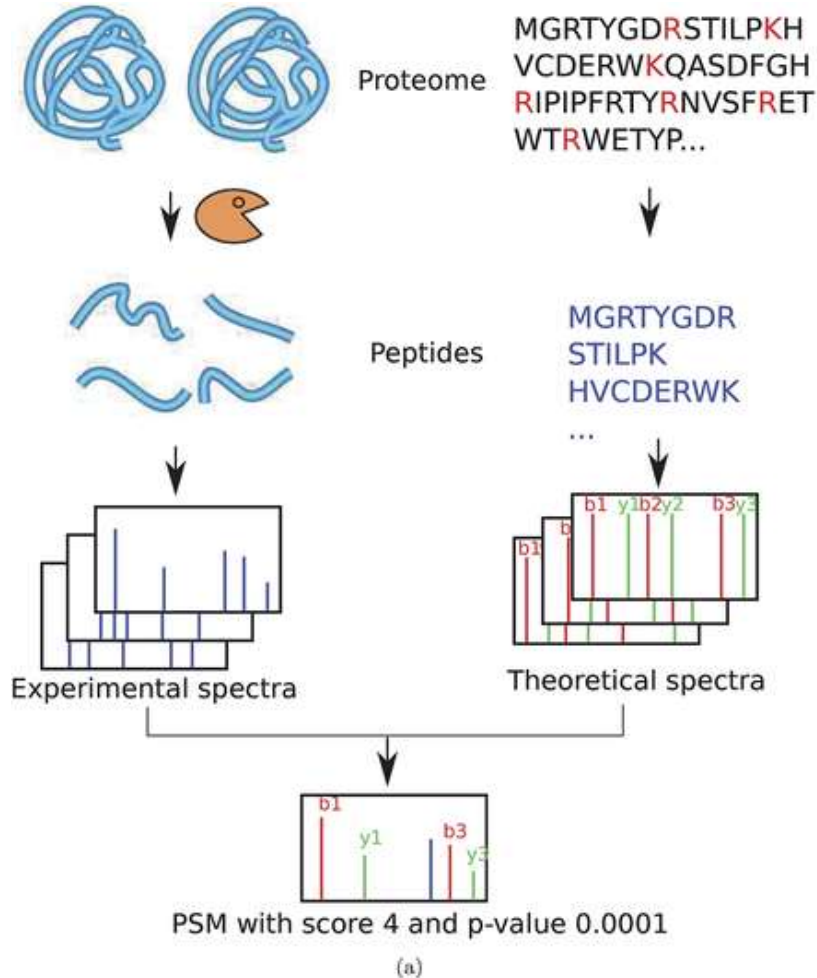
Generate URL for API Preview 10 Cancel Download

pwiz-setup-3.0.24054.3456c11-x86_64	Windows Installer ...	25-02-2024 11:37	81,842 KB
uniprotkb_e_coli_AND_model_organism_833_2024_02_24	FASTA File	24-02-2024 08:45	1,852 KB

Database Search tool

DIA-NN

Database Search Pipeline



Dynamic PTMs grow search space

Because multiple PTMs may be in each peptide, adding PTMs to a search creates an *exponential* cost.

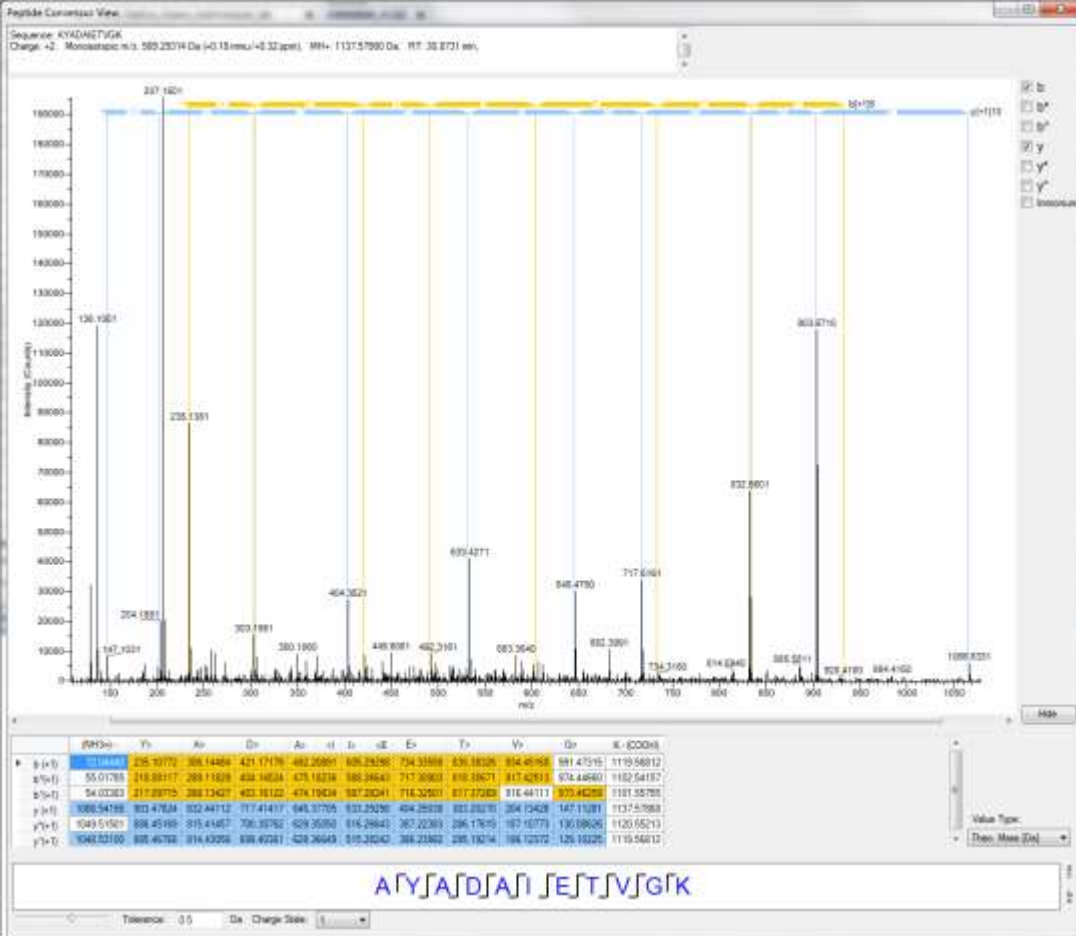
Here, three sites lead to eight PTM variants.

D	I	G	S	E	S	T	E	K
D	I	G	S*	E	S	T	E	K
D	I	G	S	E	S*	T	E	K
D	I	G	S	E	S	T*	E	K
D	I	G	S*	E	S*	T	E	K
D	I	G	S*	E	S	T*	E	K
D	I	G	S	E	S*	T*	E	K
D	I	G	S*	E	S*	T*	E	K

CASA1_BOVIN

Yates et al *Anal. Chem.* (1995) 67: 1426-1436.

Spectral Assignment



Peptide Spectrum Match Identification Details

Peptide Summary
Sequence: AYADAIETVGGK, Charge: +2, Monoisotopic m/z: 569.29314 Da (+0.15 mDa/+0.32 ppm), MH+: 1137.57900 Da, RT: 30.8731 min.
Identified with: Sequest HT (v1.17); XCorr: 2.36, Percolator q-Value: 0.060, Percolator PEP: 3.2e-3, Ions matched by search engine: 0/0
Fragment match tolerance used for search: 0.6 Da
Fragments used for search: b; b-H₂O; y; y-H₂O; y-NH₃

Fragment Matrix

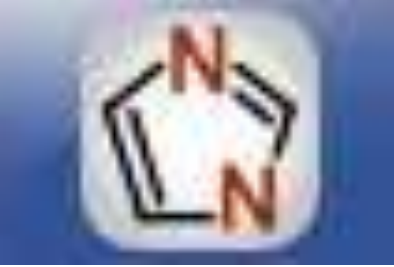
Ion Score	Neutral Losses	Precursor Ions
#1	a ¹	b ¹
1	44.54948	22.52838
2	207.1120	72.04435
3	278.14852	154.06004
4	393.17686	139.57860
5	454.21337	187.29207
6	577.29804	232.61063
7	706.34063	289.32688
8	807.38831	353.67395
9	908.45612	434.33779
10	983.47819	482.73200
11		569.29314

Match Tolerance: 0.6 Da
Match Tolerance: 0.0 Da
Mass analyzer: ITMS
Match Tolerance: 0.0 Da

Fragment Spectrum

12542024_4.cw #16813 RT: 30.8731 min
ITMS, 569.2933@h0226.00, z=+2, Mono m/z=569.29314 Da, MH+=1137.57900 Da, Match Tol=0.6 Da

Peptide: AYADAIETVGGK



Download DIA-NN

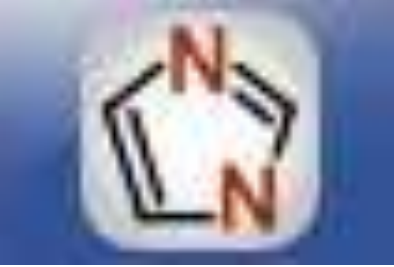
The screenshot shows the GitHub release page for DIA-NN 1.8.1. The page title is "DIA-NN 1.8.1" with a "Latest" badge. It indicates that vdemichev released this version on April 15, 2022, with 20 commits to master since the release. The release includes a list of features and improvements:

- Multiplexing support
- Improved dia-PASEF performance
- 'Peak height' quantification mode
- Fixed handling of cases when a spectral library annotates the precursor ion as fragment or includes yL/b1 fragments
- Stability issues under Linux solved

Below the list, there is an "Assets" section with a table of download links:

Asset	Size	Released
DIA-NN_1.8.1.Setup.exe	132 MB	Apr 15, 2022
diann_1.8.1.deb	136 MB	Apr 15, 2022
diann_1.8.1.tar.gz	136 MB	Apr 15, 2022
Source code (zip)		Apr 15, 2022
Source code (tar.gz)		Apr 15, 2022

At the bottom of the assets section, it shows 35 people reacted to the release.



Interface

05/07/2024 12:41:32 x +

Experiment name: 05/07/2024 12:41:32

Input

Raw diaPASEF .d Clear list Convert to .dia

Spectral library

Add FASTA

Clear list

Reannotate

DIA-NN exe diann.exe

Precursor ion generation

FASTA digest for library-free search / library generation

Deep learning-based spectra, RTs and IMs prediction

Protease Trypsin/P Missed cleavages 1

Maximum number of variable modifications 1

N-term M excision C carbamidomethylation

Ox(M) Ac(N-term) Phospho K-GG

Peptide length range 7 - 30

Precursor charge range 2 - 5

Precursor m/z range 400 - 900

Fragment ion m/z range 200 - 1800

Output

Use existing .quant files when available

Main output

Temp/.dia dir

Generate spectral library Quantities matrices

Output library F:\OMICS course 2024 data\Data for

Generate Prosit input from FASTA or spectral library

Precursor FDR (%) 1.0 Threads 12

Generate PDF report Log level 1

Additional options

Algorithm

Mass accuracy 0.0 Unrelated runs

MS1 accuracy 0.0 Use isotopologues

Scan window 0 MBR

Heuristic protein inference No shared spectra

Protein inference Genes

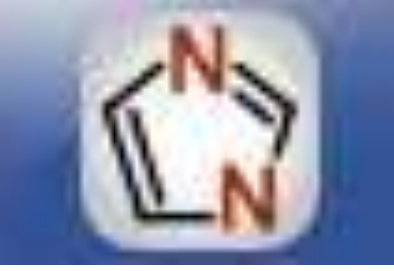
Neural network classifier Double-pass mode

Quantification strategy Robust LC (high precision)

Cross-run normalisation RT-dependent

Library generation Smart profiling

Speed and RAM usage Optimal results



Final Run window

Experiment name: 05/07/2024 12:41:32

Input

Raw diaPASEF .d Clear list Convert to .dia

\UT-2K.mzML
F:\OMICS course 2024 data\Data for DIANN search
\UT-2L.mzML
F:\OMICS course 2024 data\Data for DIANN search
\UT-OA.mzML
F:\OMICS course 2024 data\Data for DIANN search
\UT-OB.mzML
F:\OMICS course 2024 data\Data for DIANN search
\UT-OC.mzML
F:\OMICS course 2024 data\Data for DIANN search
\UT-Test-A.mzML
F:\OMICS course 2024 data\Data for DIANN search
\UT-Test-B.mzML

Spectral library

Add FASTA F:\OMICS course 2024 data
\Data for DIANN search
\uniprotkb_proteome_UP000000
625_2023_Ecoli.fasta

Clear list

Reannotate

DIA-NN exe diann.exe

Output

Use existing .quant files when available

Main output F:\OMICS course 2024 data\Data for

Temp/.dia dir

Generate spectral library Quantities matrices

Output library for DIANN search\DA2\report-lib.tsv

Generate Prosit input from FASTA or spectral library

Precursor FDR (%) 1.0 Threads 12

Generate PDF report Log level 1

Additional options

Run Running... Stop

Precursor ion generation

FASTA digest for library-free search / library generation

Deep learning-based spectra, RTs and IMs prediction

Protease Trypsin/P Missed cleavages 1

Maximum number of variable modifications 1

N-term M excision C carbamidomethylation

Ox(M) Ac(N-term) Phospho K-GG

Peptide length range 7 - 30

Precursor charge range 2 - 5

Precursor m/z range 400 - 900

Fragment ion m/z range 200 - 1800

Algorithm

Mass accuracy 0.0 Unrelated runs

MS1 accuracy 0.0 Use isotopologues

Scan window 0 MBR

Heuristic protein inference No shared spectra

Protein inference Genes

Neural network classifier Double-pass mode

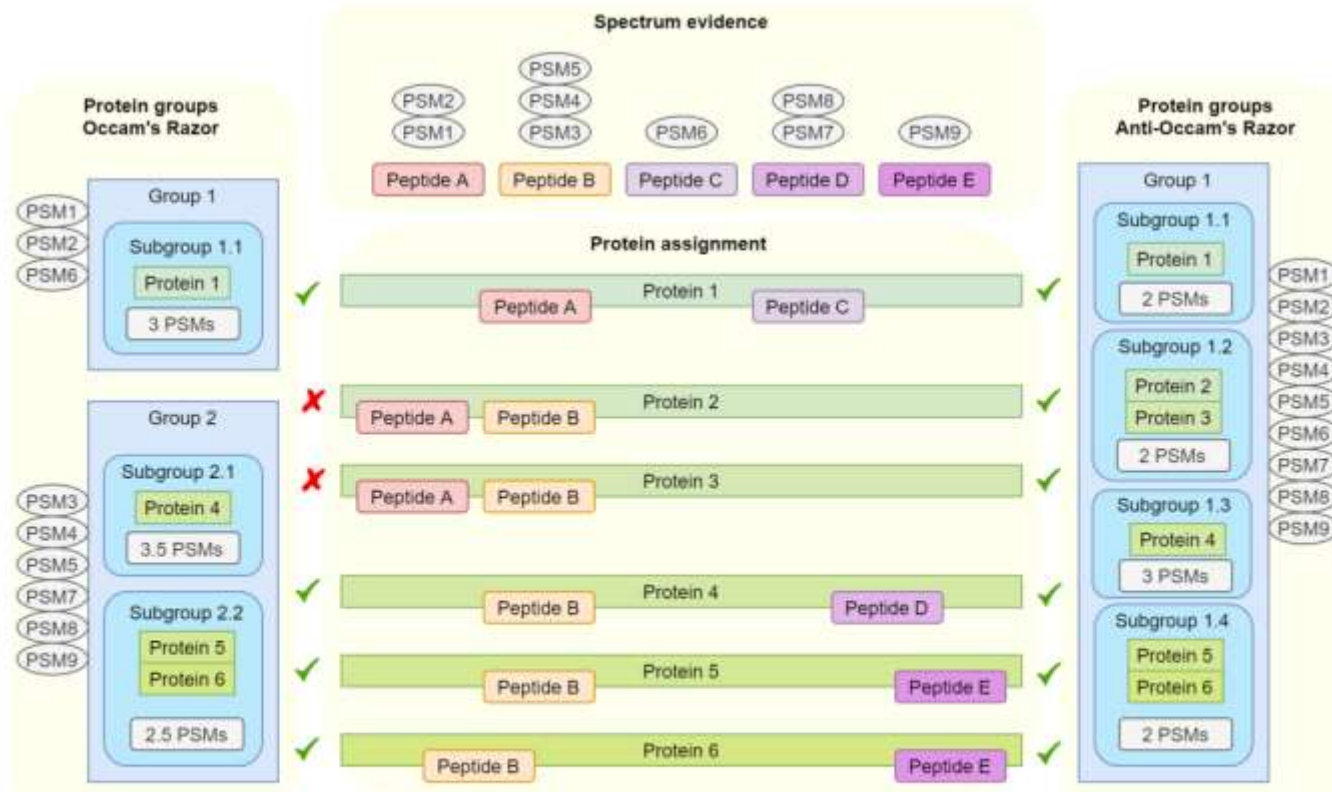
Quantification strategy Robust LC (high precision)

Cross-run normalisation RT-dependent

Library generation Smart profiling

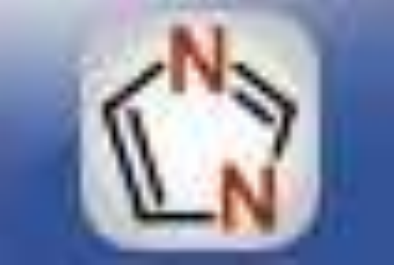
Speed and RAM usage Optimal results

Peptide Assignment and Protein grouping



Moreover, as proteins are grouped based on their identified peptides, carefully defined rules are required on when and how to group these proteins. There are two possible approaches here: the first approach consists of grouping all proteins that share one or more identified peptides (i.e., the *shared peptide rule*), while the second approach consists of only grouping proteins that share the same set (or subset) of identified peptides (i.e., the *shared peptide set rule*). These two approaches can also be interpreted as grouping at two different levels: the protein group level (based on the *shared peptide rule*) and the protein subgroup level (based on the *shared peptide set rule*). These two approaches are also visualized in **Figure 1**.

Figure 1. Protein grouping algorithms *Occam's razor* (left) and *anti-Occam's razor* (right). Groups can be based on *shared peptide rule* (protein groups) or on *shared peptide set rule* (protein subgroups). This figure also illustrates how PSMs are assigned to protein (sub)groups and shows the weighted PSM count for subgroups. When a PSM is assigned to multiple subgroups, it will be calculated as one divided by the number of subgroups, which can result in fractional PSM counts.

















Search log and output files

```
diann.exe -f "F:\OMICS course 2024 data\Data for DIANN search\T-1G.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1H.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1I.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1J.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1K.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1L.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1M.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1N.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1O.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1P.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1Q.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1R.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1S.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1T.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1U.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1V.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1W.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1X.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1Y.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-1Z.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-20.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-21.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-22.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-23.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-24.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-25.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-26.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-27.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-28.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-29.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-30.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-31.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-32.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-33.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-34.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-35.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-36.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-37.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-38.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-39.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-40.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-41.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-42.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-43.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-44.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-45.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-46.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-47.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-48.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-49.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-50.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-51.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-52.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-53.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-54.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-55.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-56.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-57.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-58.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-59.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-60.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-61.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-62.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-63.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-64.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-65.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-66.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-67.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-68.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-69.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-70.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-71.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-72.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-73.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-74.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-75.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-76.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-77.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-78.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-79.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-80.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-81.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-82.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-83.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-84.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-85.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-86.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-87.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-88.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-89.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-90.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-91.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-92.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-93.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-94.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-95.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-96.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-97.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-98.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-99.mzML" -f "F:\OMICS course 2024 data\Data for DIANN search\T-100.mzML" -lib "" -threads 12 -verbose 1 -out "F:\OMICS course 2024 data\Data for DIANN search\DataAnalysis3\report.tsv" -qvalue 0.01 -matrices -out-lib "F:\OMICS course 2024 data\Data for DIANN search\DA2\report-lib.tsv" -gen-spec-lib -predictor -reannotate -fasta "F:\OMICS course 2024 data\Data for DIANN search\uniprotkb_proteome_UP000000625_2023_Ecoli.fasta" -fasta-search -min-fr-mz 200 -max-fr-mz 1800 -met-excision -cut K,R -missed-cleavages 1 -min-pep-len 7 -max-pep-len 30 -min-pr-mz 400 -max-pr-mz 900 -min-pr-charge 2 -max-pr-charge 5 -unimod4 -var-mods 1 -var-mod UniMod:35,15.994915,M -double-search -reanalyse -relaxed-prot-inf -smart-profiling -peak-center -no-ifs-removal
DIA-NN 1.8.1 (Data-Independent Acquisition by Neural Networks)
Compiled on Apr 14 2022 15:31:19
Current date and time: Wed May 8 15:55:58 2024
CPU: GenuineIntel Intel(R) Xeon(R) CPU E5-1650 v3 @ 3.50GHz
SIMD instructions: AVX AVX2 FMA SSE4.1 SSE4.2
Logical CPU cores: 12
Thread number set to 12
Output will be filtered at 0.01 FDR
Precursor/protein x samples expression level matrices will be saved along with the main report
A spectral library will be generated
Deep learning will be used to generate a new in silico spectral library from peptides provided
Library precursors will be reannotated using the FASTA database
Library-free search enabled
Min fragment m/z set to 200
Max fragment m/z set to 1800
N-terminal methionine excision enabled
In silico digest will involve cuts at K,R*
Maximum number of missed cleavages set to 1
Min peptide length set to 7
Max peptide length set to 30
Min precursor m/z set to 400
Max precursor m/z set to 900
Min precursor charge set to 2
Max precursor charge set to 5
Cysteine carbamidomethylation enabled as a fixed modification
Maximum number of variable modifications set to 1
Modification UniMod:35 with mass delta 15.9949 at M will be considered as variable
Neural networks will be used for peak selection
A spectral library will be created from the DIA runs and used to reanalyse them; .quant files will only be saved to disk during the first step
Highly heuristic protein grouping will be used, to reduce the number of protein groups obtained; this mode is recommended for benchmarking protein ID numbers; use with caution for anything else
When generating a spectral library, in silico predicted spectra will be retained if deemed more reliable than experimental ones
Fixed-width center of each elution peak will be used for quantification
Interference removal from fragment elution curves disabled
DIA-NN will optimise the mass accuracy automatically using the first run in the experiment. This is useful primarily for quick initial analyses, when it is not yet known which mass accuracy setting works best for a particular acquisition scheme.
Exclusion of fragments shared between heavy and light peptides from quantification is not supported in FASTA digest mode - disabled; to enable, generate an in silico predicted spectral library and analyse with this library

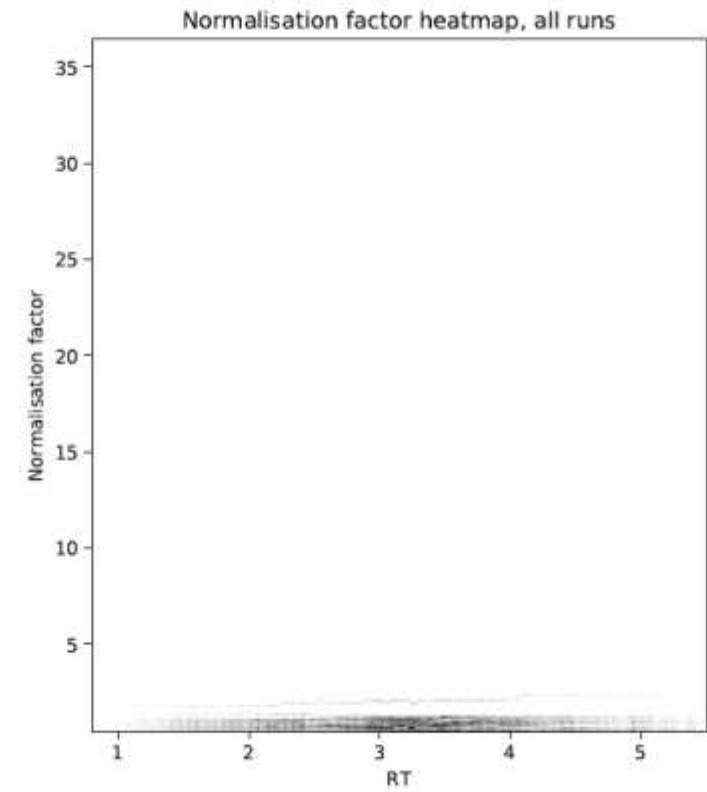
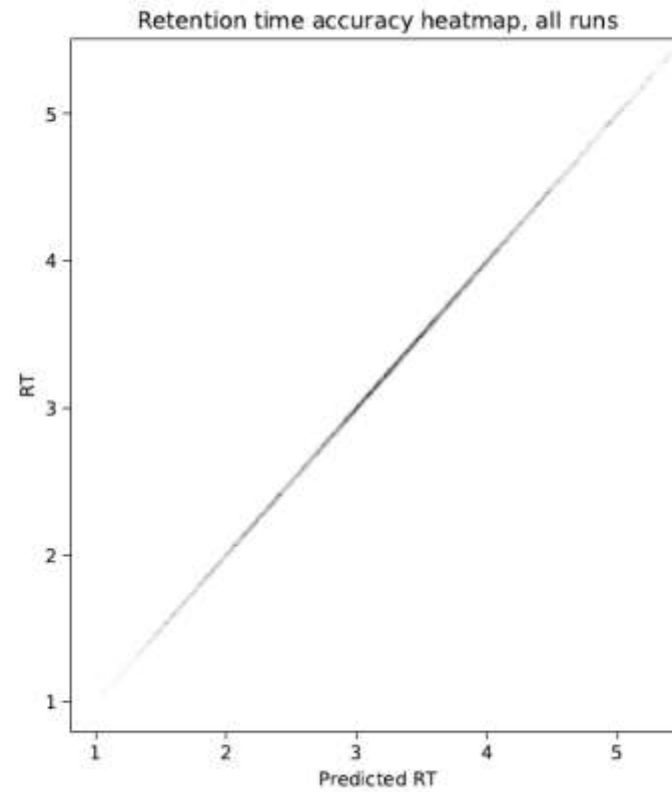
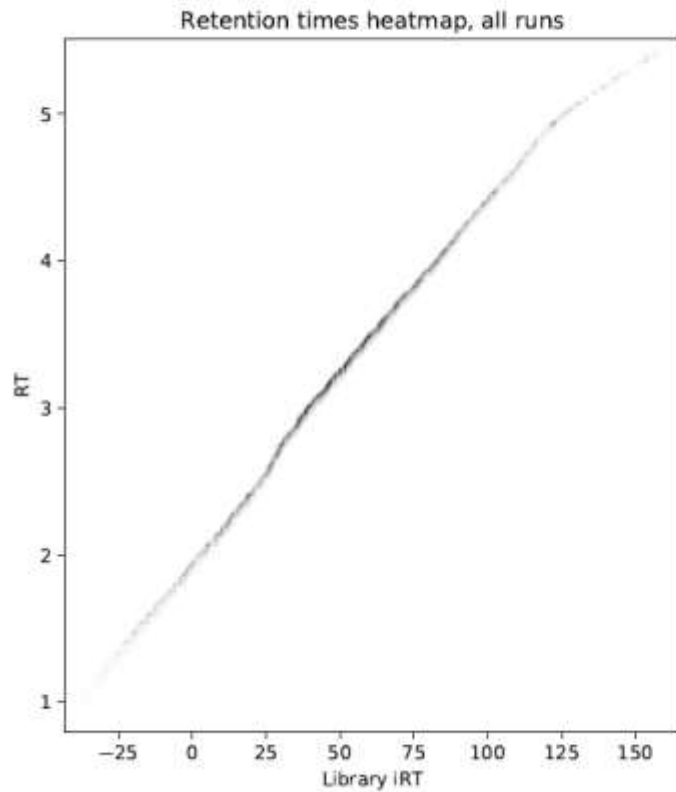
17 files will be processed
[0:00] Loading FASTA F:\OMICS course 2024 data\Data for DIANN search\uniprotkb_proteome_UP000000625_2023_Ecoli.fasta
[0:00] Processing FASTA
[0:01] Assembling elution groups
[0:02] 359300 precursors generated
[0:02] Gene names missing for some isoforms
[0:02] Library contains 4392 proteins, and 4392 genes
[0:02] Encoding peptides for spectra and RTs prediction
[0:02] Predicting spectra and IMs
```


Data Report

DIA-NN

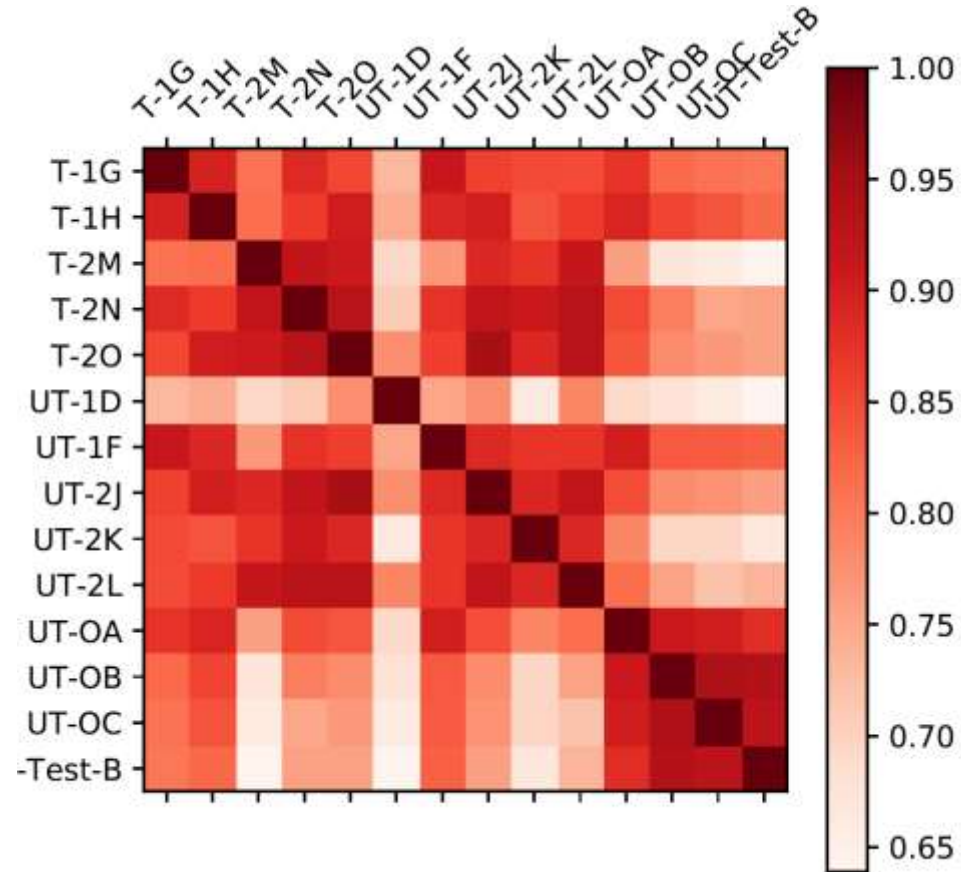
NAME	DATE MODIFIED	TYPE
 report.gg_matrix	07-05-2024 13:10	TSV File
 report.log	07-05-2024 13:10	Text Docume
 report	07-05-2024 13:10	Adobe Acrob
 report.pg_matrix	07-05-2024 13:10	TSV File
 report.pr_matrix	07-05-2024 13:10	TSV File
 report.stats	07-05-2024 13:10	TSV File
 report	07-05-2024 13:10	TSV File
 report.unique_genes_matrix	07-05-2024 13:10	TSV File
 report-first-pass.gg_matrix	07-05-2024 13:08	TSV File
 report-first-pass.pg_matrix	07-05-2024 13:08	TSV File
 report-first-pass.pr_matrix	07-05-2024 13:08	TSV File
 report-first-pass.stats	07-05-2024 13:08	TSV File
 report-first-pass	07-05-2024 13:08	TSV File
 report-first-pass.unique_genes_matrix	07-05-2024 13:08	TSV File

RT correction and normalisation

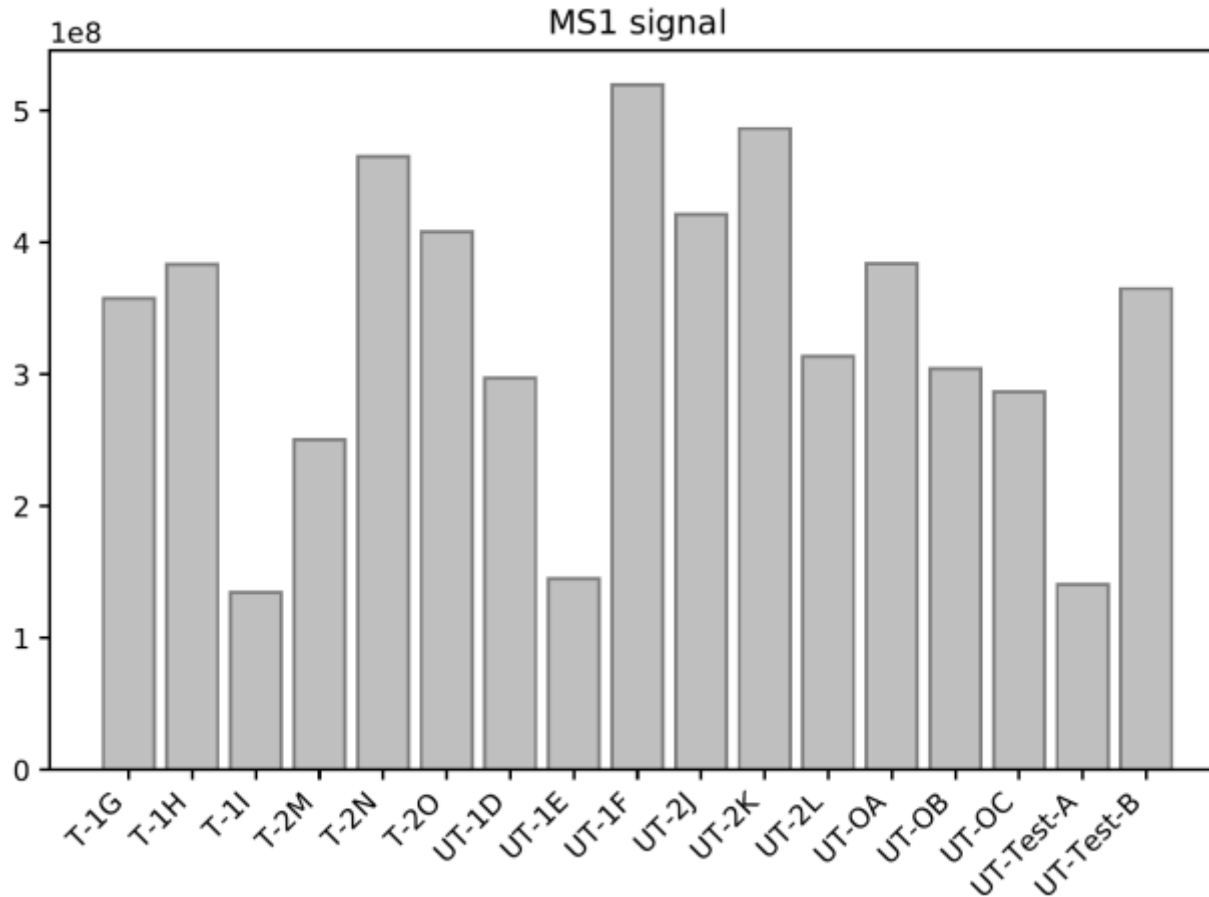




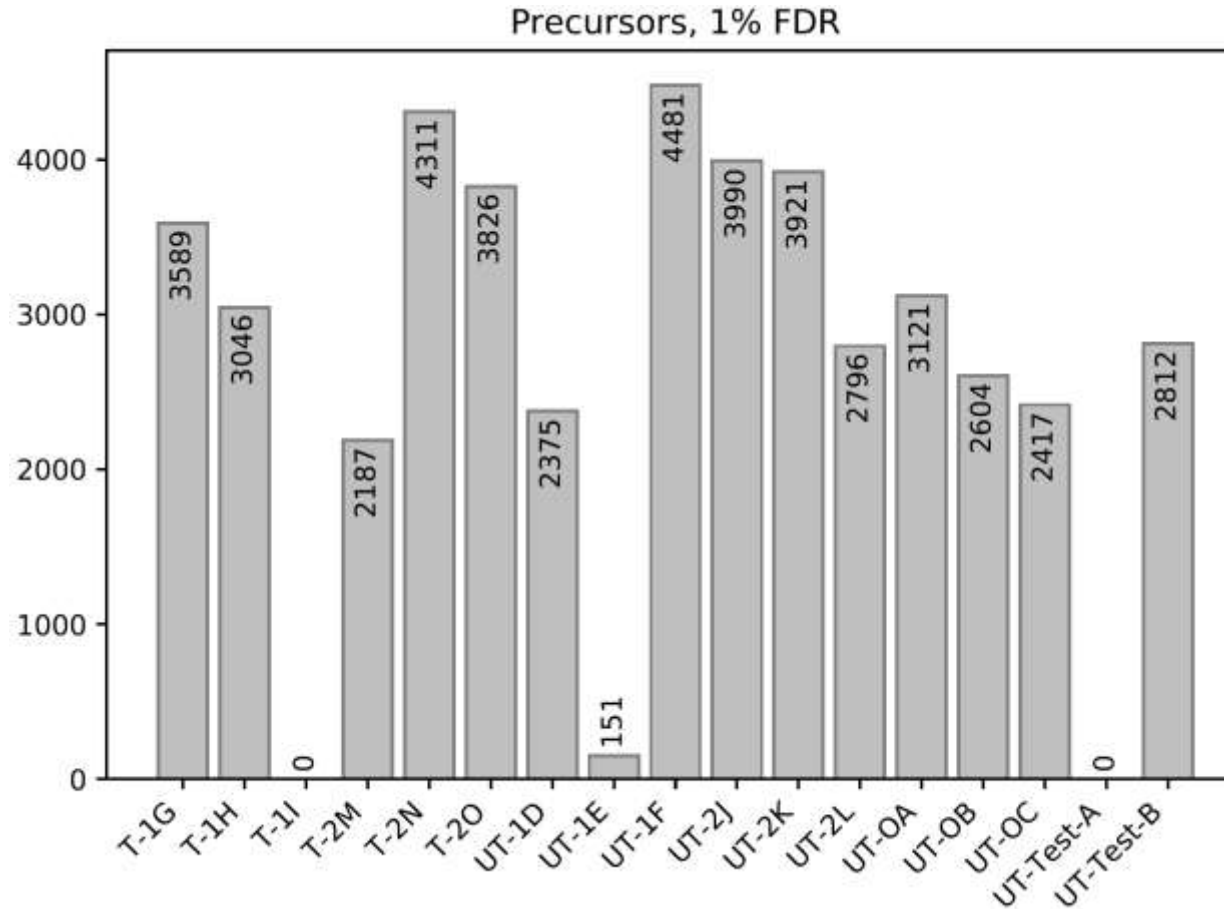
Sample to sample correlation



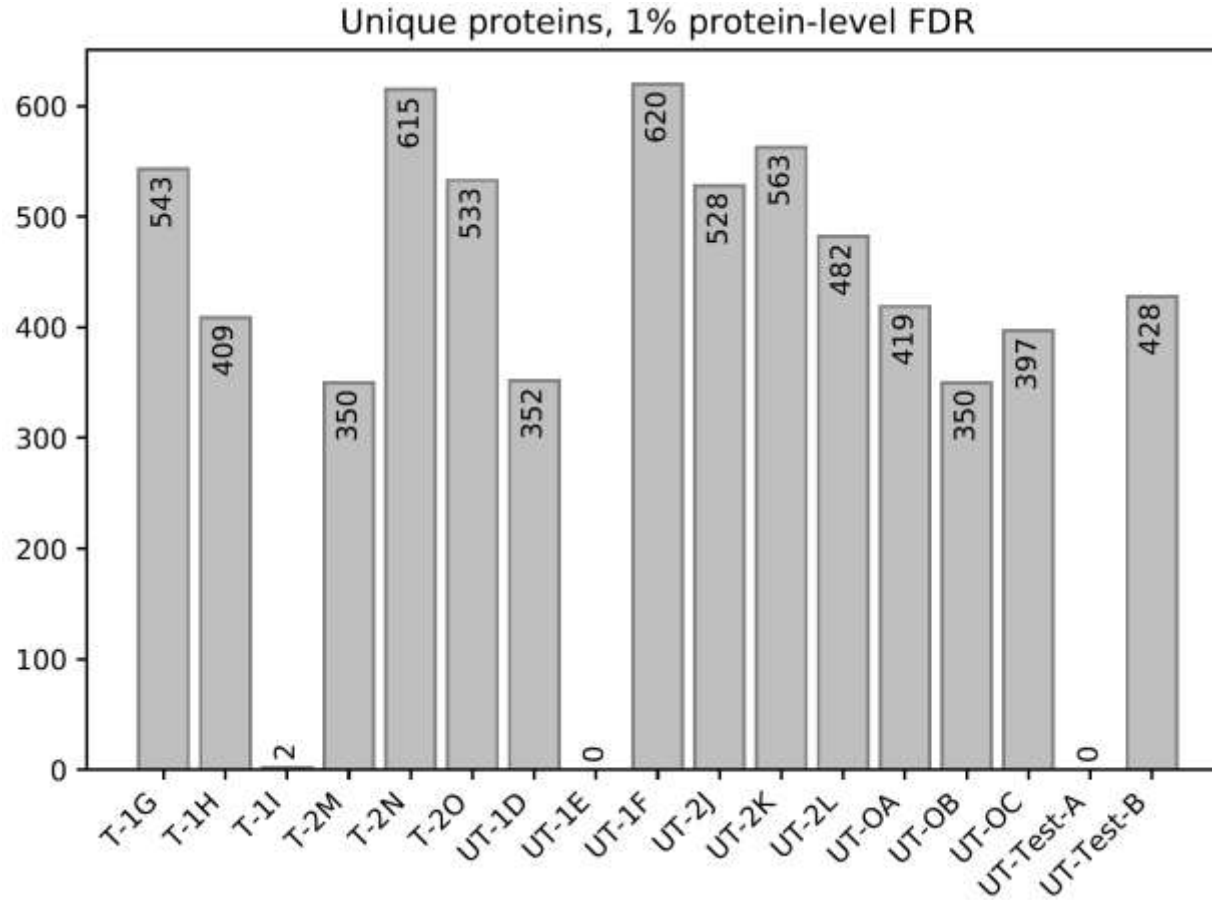
Total MS1 signal-Meaning data is not normalised



Peptides identified in each sample

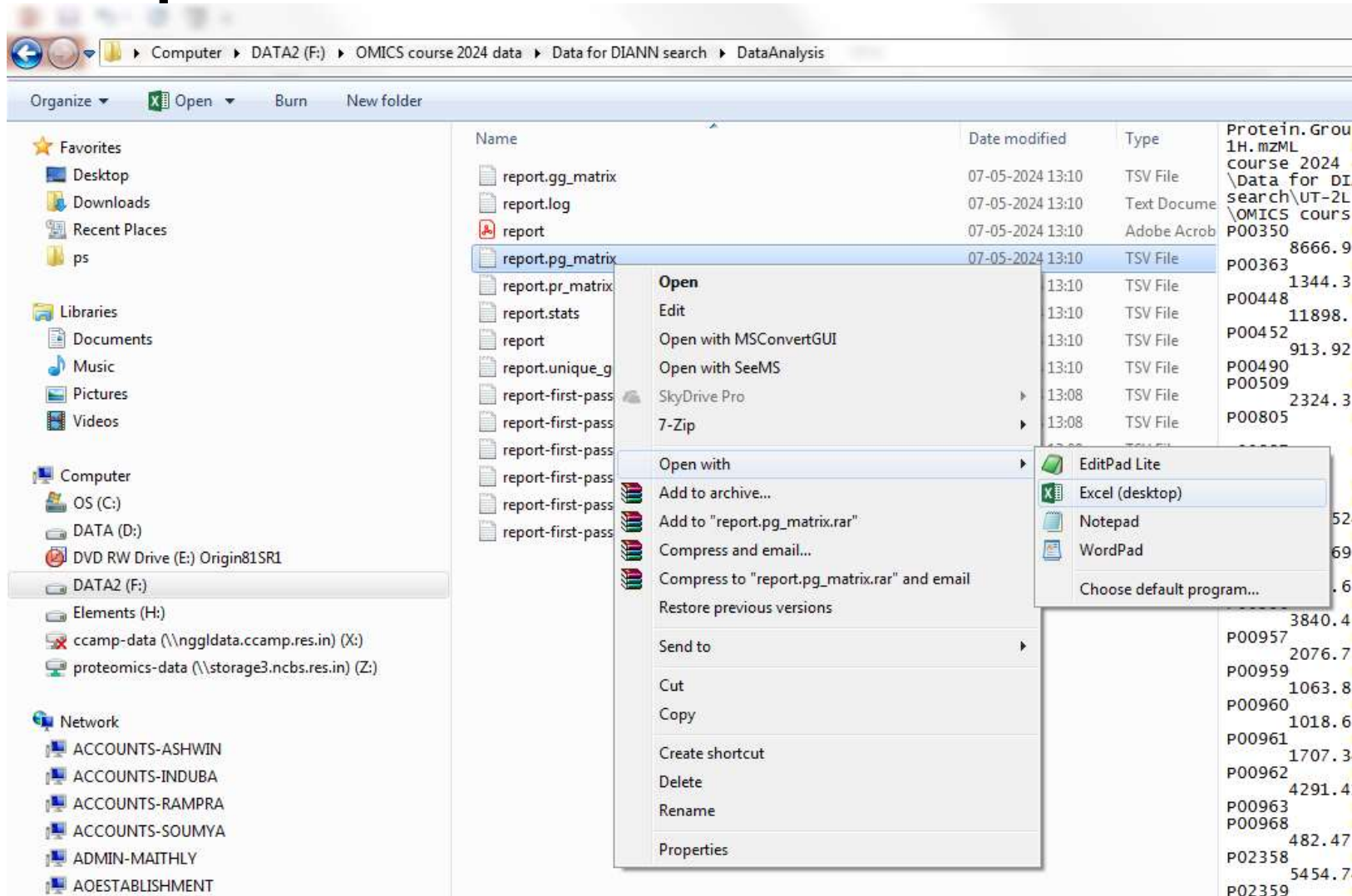


Proteins identified in each sample





Open "report PG matrix" data with MS-Excel





Protein Quantitation in the report

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
	Protein. Group	Protein.Ids	Protein. Names	Genes	First.Protein.Description	F:\OMICS course 2024 data\Data for DIANN search\T-1G.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-1H.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-1I.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-2M.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-2N.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-2O.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-1D.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-1E.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-1F.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-2J.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-2K.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-2L.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-OA.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-OB.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-OC.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-A.mzML	F:\OMICS course 2024 data\Data for DIANN search\T-B.mzML
1																						
2	P00350	P00350	6PGD_ECCgnd		6-phospho	9218.67	11752.3		10598.2	9021.98	10615.6	12927.4		10578.9	10099.1	8666.95	10584.9	10286.5	11165.4	10817.9		9485.67
3	P00363	P00363	FRDA_ECCfrdA		Fumarate	1687.88	2534.55			1000.88	2223.44	898.345		2126.16	2182.46	1597.51	1344.35	863.145				110.175
4	P00448	P00448	SODM_EC soda		Superoxid	6844.32	7549.89		7951.95	8832	9130.32	3897.57		2955.97	4920.13	6596.02	4827.69	11898.5	9109.1	10548.1		9214.31
5	P00452	P00452	RIR1_ECOI nrdA		Ribonuclé	144.154	692.111			1079.83				834.012	169.96	205.208	913.928	1023.25	1002.69	596.871		557.567
6	P00490	P00490	PHSM_ECI malP		Maltodextrin phosphorylase						488.045						544.51					
7	P00509	P00509	AAT_ECOI aspC		Aspartate	2003.61	2448.35		2163.91	2251.65	2536.75	2251.63		2413.44	2619.68	2447.59	2219.65	2324.38	1225.31	1045.87		1409.28
8	P00805	P00805	ASPG2_EC ansB		L-asparagi	8015.24	10657.6		7435.22	7605.79	9366.3	18667	31406.6	11194.4	9905.96	9541.06	11375.5	2187.62		378.585		
9	P00887	P00887	AROH_EC(aroH		Phospho-2-dehydro-3-deoxyheptonate aldolase, Tr					394.278				772.804					1356.2			680.607
10	P00926	P00926	SDHD_ECC dsdA		D-serine dehydratase										435.548	379.546						
11	P00934	P00934	THRC_ECC thrC		Threonine synthase				1425.82	1806.46	1178.86			579.938	1194.55		1219.27	487.077				364.524
12	P00946	P00946	MANA_EC manA		Mannose-	2186.81	1603.2		1350.4	1596.58	1473.49			2168.67	1493.05	1683.88	1342.04	942.698	819.033	507.834		991.91
13	P00954	P00954	SYW_ECOI trpS		Tryptophá	2357.92	3539.96		1264.73	2457.5	2437.22	804.694		4047.66	1592.96	1725.69	1765.34	3570.65	3116.98	2484.88		3314.35
14	P00956	P00956	SYI_ECOLI ileS		Isoleuciné	2579.64	3539.28		2876.24	2517.19	3198.79	4609.36		3832.75	3089.53	2315.53	3755.77	3840.46	3693.36	3455.5		3275.67
15	P00957	P00957	SYA_ECOLI alaS		Alanine--t	1856.53	2409.44		2204.26	2178.77	2055.58			2210.48	2122.99	1434.51	1728.32	2095.72	2076.78	2541.37		2515.68
16	P00959	P00959	SYM_ECOI metG		Methionir	811.187	1006.07			1047.47	977.074			990.973	682.307	878.825	1115.74	1028.95	1063.82	925.182		1410.21
17	P00960	P00960	SYGA_ECC glyQ		Glycine--t	866.576				611.169	121.282			774.734		302.202						1018.63
18	P00961	P00961	SYGB_ECC glyS		Glycine--t	901.883	1054.24			830.117	541.614			1455.03	759.956	386.284	1061.71	1707.34	722.171	1017.46		1974.91
19	P00962	P00962	SYQ_ECOLI glnS		Glutaminé	4096.21	4876.54		3316.92	3783.51	4009.21	975.018		4640.96	3681.41	3079.61	3517.44	4291.41	4051.58	3632.75		3370.75
20	P00963	P00963	ASNA_EC(asnA		Aspartate--ammonia ligase									122.99		238.081		653.773		108.622		252.085
21	P00968	P00968	CARB_ECC carB		Carbamoy	1044.67	860.999		843.496	1066.14	204.791	320.315		1107.91	444.283	360.996	482.475					
22	P02358	P02358	RS6_ECOLI rpsF		Small ribo	4870.04	5384.08		4802.67	5090.44	5296.58	8989.34		4682.21	4603.32	2906.76	5454.74	5204.59	5967.39	4441.24		5581.47
23	P02359	P02359	RS7_ECOLI rpsG		Small ribo	8828.7	6194.86		6523.38	10880.6	6900.51	1167.01		7508.02	5389.97	5182.39	5622.05	7612.2	5943.96	6622.58		10381.8
24	P02413	P02413	RL15_ECOI rplO		Large ribo	4037.69	2405.75		5163.05	4661.2	2534.05	1158.04		4119.05	2962.48	2186.19	3531.87	3990.76	4824.73	3886.1		7451.6
25	P02924	P02924	ARAF_ECC araF		L-arabinose-binding periplasmic protein					235.667	239.787					248.937						
26	P02925	P02925	RBSB_ECC rbsB		Ribose im	41864.7	25735.6		95062.2	70371.8	50705.3	31483.9	54517.2	42313.9	38996.9	79204.7	91029.5	14185.3	4046.74	4898.02		3828.54
27	P03817	P03817	MIOC_ECC mioC		Protein M	1457.92	811.593		3012.54	1937.67	1605.16	912.057		1493.21	1463.89	2550.68	1385.83	1860.76	1140.75	1234.11		1247.36
28	P03841	P03841	MALM_EC malM		Maltose o	566.876	652.948		1729.38	1673.94	1649.01			749.324	689.579	1569.47	662.62	327.227				
29	P04079	P04079	GUAA_EC(guaA		GMP syntf	3953.56	5059.24		2841.2	3798.6	3948.42	372.446		3958.71	3497.44	3254.34	2666.8	5379.92	4894.22	4401.1		4701.49
30	P04335	P04335	FRSA_ECC frsA		Esterase F	1292.97	693.338			955.622	696.028			834.576	613.299	780.095	471.105	417.542				
31	P04425	P04425	GSHB_ECC gshB		Glutathione synthet	727.702				200.214	1073.65	741.54		612.001	1095.55	575.194	1239.92	1033.61		1245.6		861.374