

Introduction to Proteomic Data Analysis III

Analysis of Protein Phosphorylations

University College Dublin, Conway Institute

Matthias Wilm, MD PhD,
Conway Institute, UCD,
Belfield, Dublin 4

Protein Analysis

MaxQuant



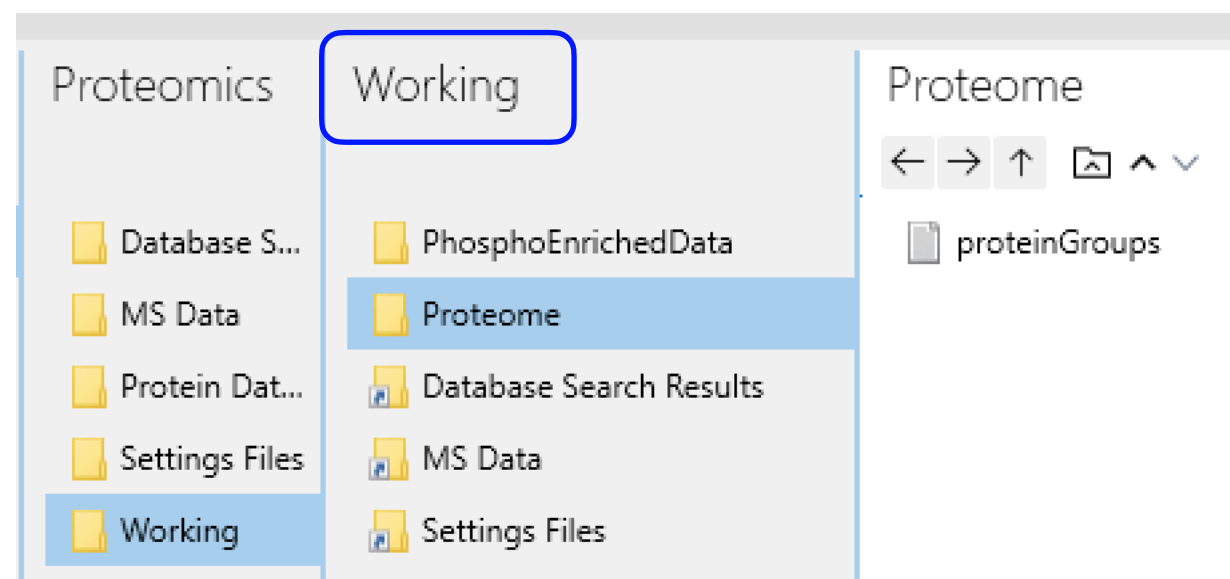
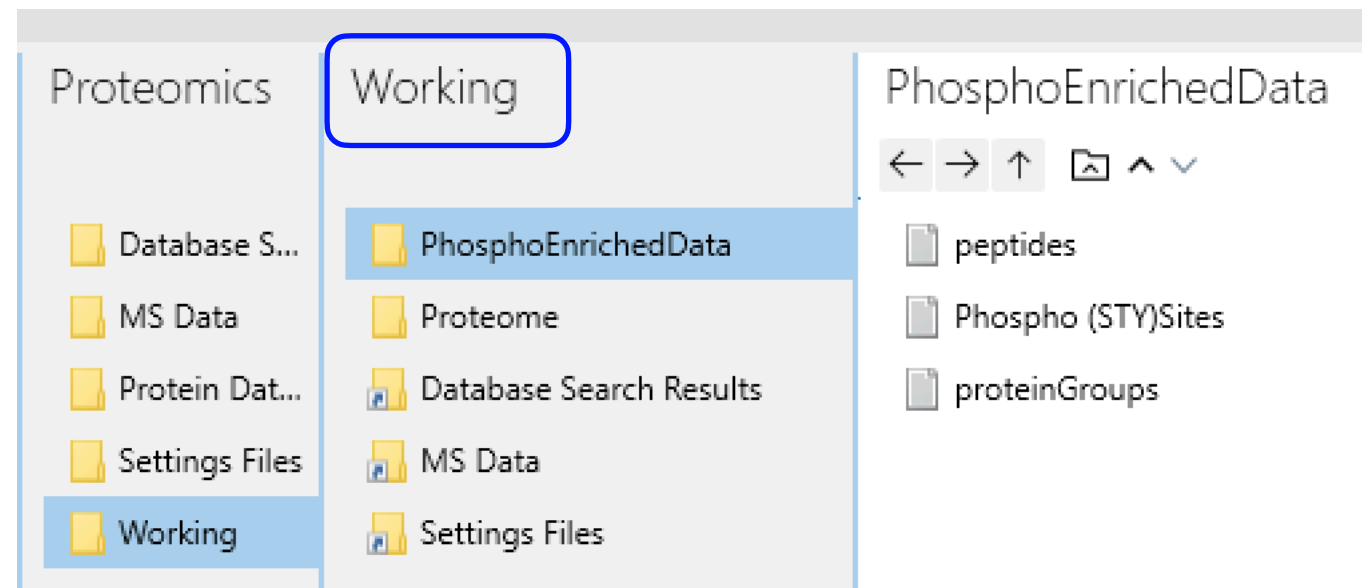
- Protein Identification
- Quantification

Perseus



- Statistical Analysis of Quantitative Results

Data Organization: Move the relevant data to the Working Folder



What is the Aim of the Analysis ?

Which proteins change their phosphorylation state in comparison to a control in a statistically significant way ?

Comparison between two samples: **t-Test**

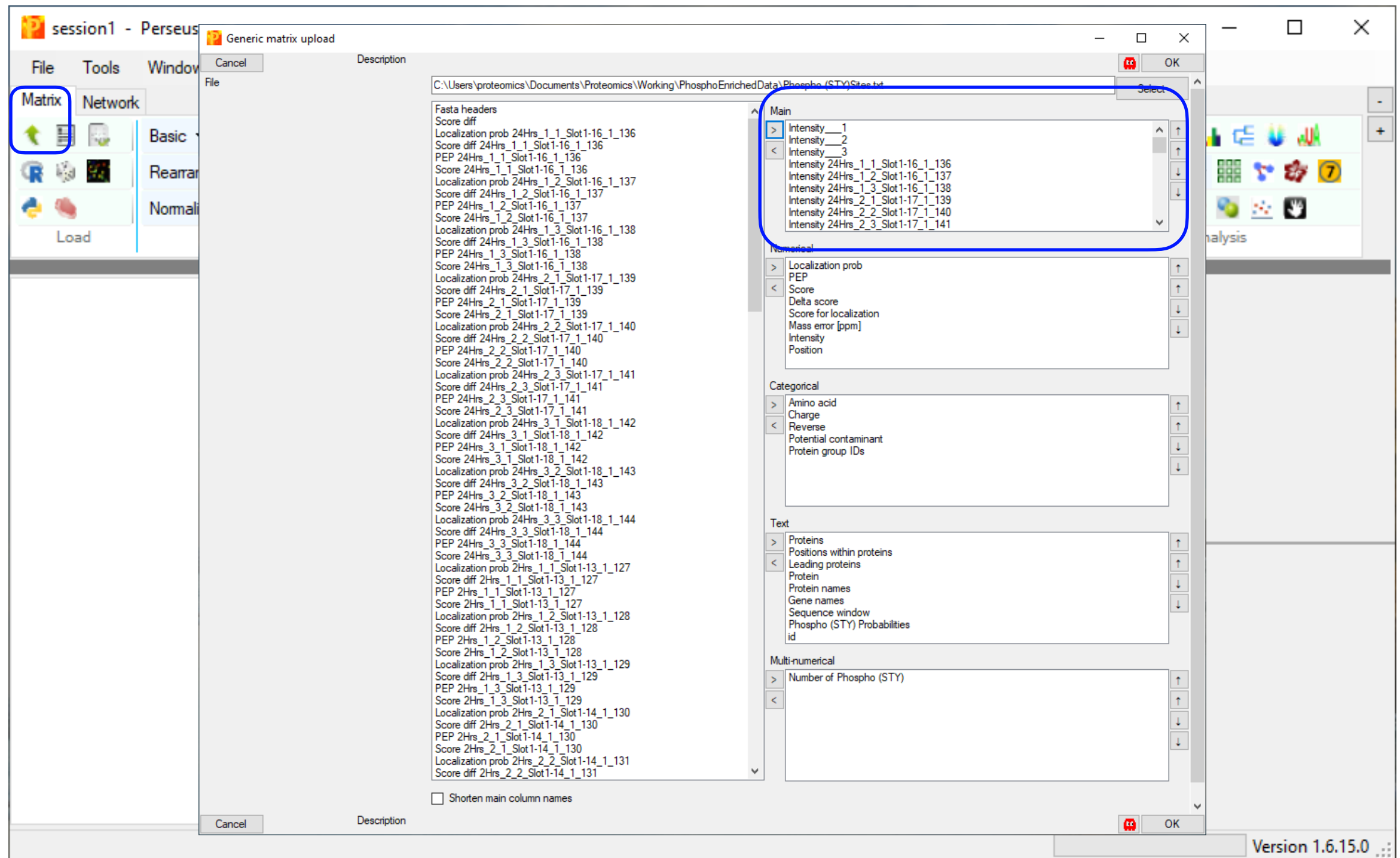
t-Test: Is the mean of two groups of measurement repeats the same ?

Comparison between several samples:

ANOVA - Test

ANOVA-Test: Are the means of several groups of measurement repeats the same ?

Start Perseus and Load Data: **Matrix > Generic matrix upload:**
Phospho (STY)Sides.txt



Start Perseus and Load Data: **Matrix > Generic matrix upload:** Phospho (STY)Sides.txt

session1 - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series DE analysis Isobaric Labeling P₂

Rearrange Filter columns Annot. rows Learning Modifications External Z 2D X P_N

Normalization Quality Tests Outliers WGCNA Clustering f(x) 1D P₂

Load Processing Analysis Multi-proc. Export

matrix1

Data

Type	ity	N: Position	T: Proteins	T: Positio... within	T: Leading proteins	T: Protein	T: Protein names	T: Phospho (STY) Probabilities	T: id	M: Number of
	ric	Numeric	Text	Text	Text	Text	Text	Text	Text	Multi-n...
1		95	P0DN3...	95;95	P0DN37	P0DN37	Peptid...	HT(0.01)GS(0.49)GILS(0.49)MANAGPNT(0.01)NGSQFFICTAK	0	
2	0	99	P0DN3...	99;99	P0DN37	P0DN37	Peptid...	HT(0.032)GS(0.265)GILS(0.669)MANAGPNT(0.034)NGSQFFICTAK	1	1
3	00	153	A5YM7...	153;27...	A5YM72	A5YM72	Carnos...	EGQET(1)LVKEEVEAFLRS(1)EALGDILQVAVK	2	2
4		112	A6NFI3	112	A6NFI3	A6NFI3	Zinc fin...	GQDAKS(1)PVLQEK	3	1
5	3	246	A6NKD9	246	A6NKD9	A6NKD9	Coiled...	S(1)LDDL SAPPHR	4	1
6		251	A6NKD9	251	A6NKD9	A6NKD9	Coiled...	AGAT(0.327)RRS(0.327)LDDL(0.346)APPHR	5	
7		12	A6NMY...	12;12;30	A6NMY...	A6NMY...	Putativ...	S(0.003)T(0.003)VHEILCKLS(0.964)LEGDHS(0.028)T(0.002)PPSAYGSVK	6	1
8	0	629	A7XYQ1	629	A7XYQ1	A7XYQ1	Sine o...	RAGS(1)PPGPPGAGGQLGFPVLQGPQDGVIDLTVGHR	7	1
9		116	Q96HA...	116;11...	Q96HA...	Q96HA...	Nuclea...	GLNS(0.004)QS(0.088)S(0.908)DDHLNKR	8	1
10	00	39	Q9961...	39;39;39	Q9961...	Q9961...	Eukary...	QPLLLS(1)EDEEDTKR	9	1
11	0	435	C9JH2...	435;641	C9JH2...	C9JH2...	Proline...	LFRLS(1)PGHAAPLLPGGWWTGPPDKEPLGSAIAR	10	1
12	0	37	CON_...	37	CON_...	CON_...		KDADEAS(0.999)MNKVELGS(0.001)R	11	1
13	0	115	CON_...	115	CON_...	CON_...		VPQLEIVPNS(1)AEER	12	1
14		35	CON_...	35	CON_...	CON_...		FQS(1)EEQQQTEDELQDK	13	1
15	00	138	CON_...	138;138	CON_...	CON_...	Alpha...	CDSSPDS(1)AEDVRK	14	1
16	2	320	CON_...	320	CON_...	CON_...		HT(0.005)FS(0.015)GVAS(0.98)VESSSGEAFHVGK	15	1
17		323	CON_...	323	CON_...	CON_...		HTFSGVASVES(0.918)S(0.081)S(0.001)GEAFHVGK	16	1
18		324	CON_...	324	CON_...	CON_...		HTFSGVAS(0.004)VES(0.463)S(0.463)S(0.069)GEAFHVGK	17	
19	0	44	O00178	44	O00178	O00178	GTP-b...	LHGGFDS(1)DCS(1)EDGEALNGEPELDLTSK	18	2
20	0	47	O00178	47	O00178	O00178	GTP-b...	LHGGFDS(1)DCS(1)EDGEALNGEPELDLTSK	19	2
21		580	O00178	580	O00178	O00178	GTP-b...	LLQTT(0.004)NNS(0.708)PMNS(0.288)KPQQIK	20	1
22		24	O00178	24	O00178	O00178	GTP-b...	SAMDSPVPAS(0.002)MFAPEPS(0.608)S(0.39)PGAAR	21	1
23	0	25	O00178	25	O00178	O00178	GTP-b...	SAMDSPVASMFAPEPS(0.072)S(0.928)PGAAR	22	1

2,444 items 1 selected

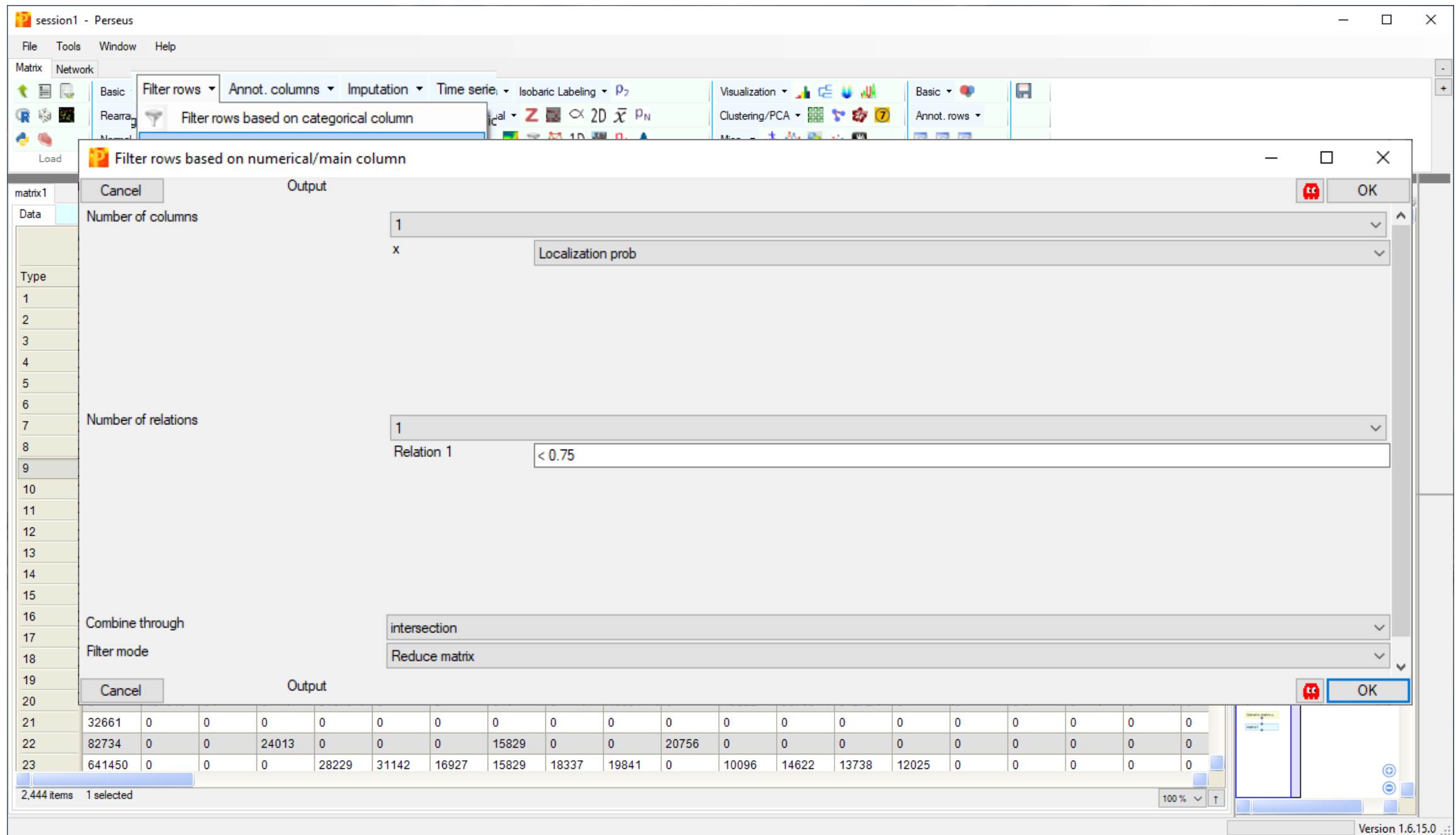
100 %

Version 1.6.15.0

Generic matrix u...
matrix1

Filter Irrelevant Data: **Processing -> Filter rows -> based on categorical column, based on numerical/main column:**

Filter out: reverse hits, contaminants, hits with a low localisation probability (< 0.75)



session1 - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series Isobaric Labeling P₂

Visualization Clustering/PCA Basic Annot. rows

Filter rows based on categorical column

Filter rows based on numerical/main column

Cancel Output OK

matrix1

Data

Number of columns 1

x Localization prob

Number of relations 1

Relation 1 < 0.75

Combine through intersection

Filter mode Reduce matrix

Cancel Output OK

21	32661	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	82734	0	0	24013	0	0	0	15829	0	0	20756	0	0	0	0	0	0	0	0
23	641450	0	0	0	28229	31142	16927	15829	18337	19841	0	10096	14622	13738	12025	0	0	0	0

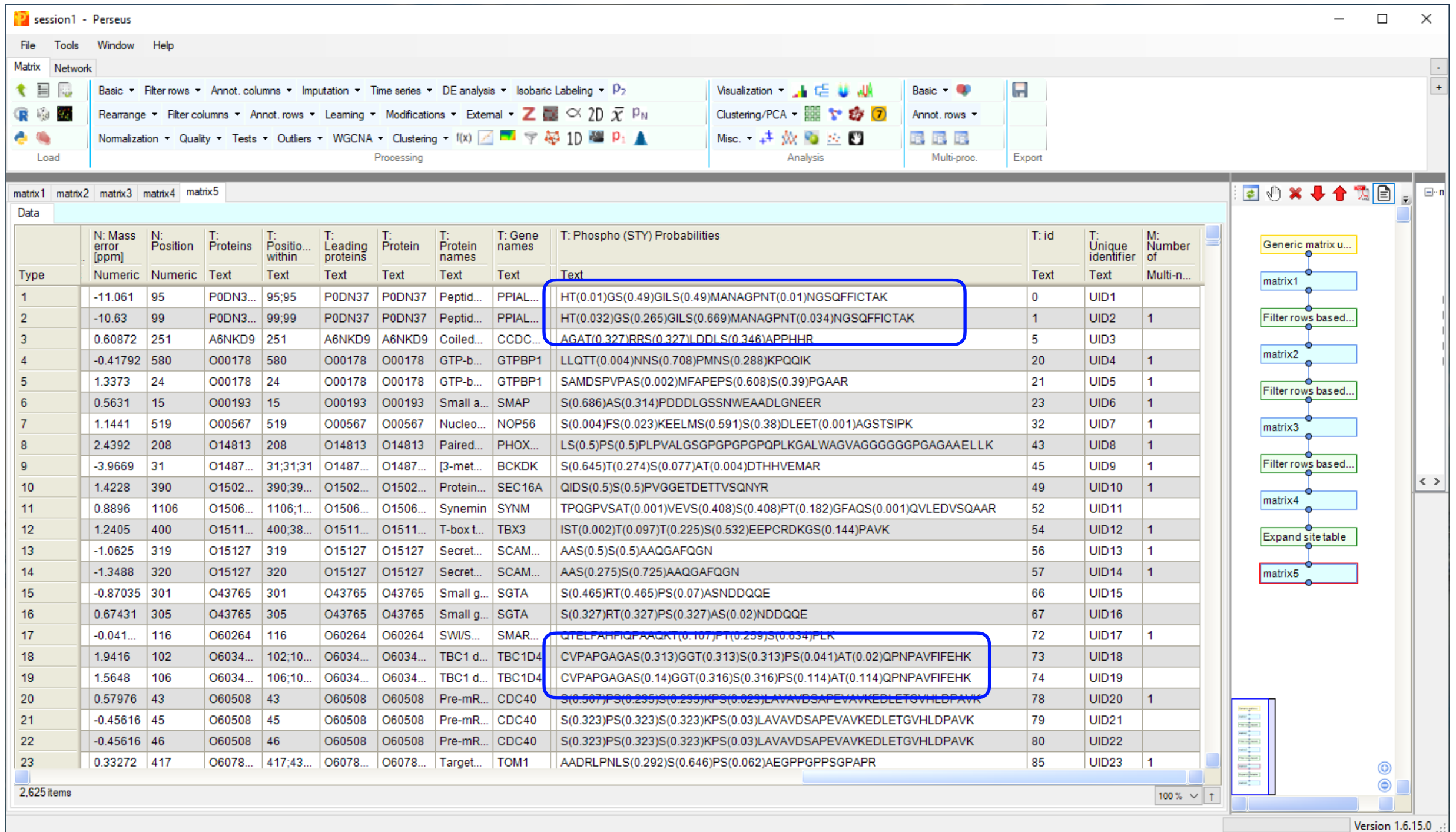
2,444 items 1 selected

100 %

Version 1.6.15.0

Reformat the Table: **Processing** -> **Modification** -> **Expand site table**:

Every single phosphorylation status (1, 2 or 3 phosphorylations within one peptide) gets its own row -> every row represents a phosphorylation state



The screenshot displays the Perseus software interface. The main window shows a data table with columns for various protein and peptide properties. The 'T: Phospho (STY) Probabilities' column is highlighted with a blue box, showing the expanded site table for each peptide. The workflow diagram on the right side of the interface shows the sequence of steps: 'Generic matrix u...' -> 'matrix1' -> 'Filter rows based...' -> 'matrix2' -> 'Filter rows based...' -> 'matrix3' -> 'Filter rows based...' -> 'matrix4' -> 'Expand site table' -> 'matrix5'.

Type	N: Mass error [ppm]	N: Position	T: Proteins	T: Positio... within	T: Leading proteins	T: Protein	T: Protein names	T: Gene names	T: Phospho (STY) Probabilities	T: id	T: Unique identifier	M: Number of
	Numeric	Numeric	Text	Text	Text	Text	Text	Text	Text	Text	Text	Multi-n...
1	-11.061	95	P0DN3...	95;95	P0DN37	P0DN37	Peptid...	PPIAL...	HT(0.01)GS(0.49)GILS(0.49)MANAGPNT(0.01)NGSQFFICTAK	0	UID1	
2	-10.63	99	P0DN3...	99;99	P0DN37	P0DN37	Peptid...	PPIAL...	HT(0.032)GS(0.265)GILS(0.669)MANAGPNT(0.034)NGSQFFICTAK	1	UID2	1
3	0.60872	251	A6NKD9	251	A6NKD9	A6NKD9	Coiled...	CCDC...	AGAT(0.327)RRS(0.327)LDLIS(0.346)APPHHR	5	UID3	
4	-0.41792	580	O00178	580	O00178	O00178	GTP-b...	GTPBP1	LLQTT(0.004)NNS(0.708)PMNS(0.288)KPQKIK	20	UID4	1
5	1.3373	24	O00178	24	O00178	O00178	GTP-b...	GTPBP1	SAMDSPVPAS(0.002)MFAPEPS(0.608)S(0.39)PGAAR	21	UID5	1
6	0.5631	15	O00193	15	O00193	O00193	Small a...	SMAP	S(0.686)AS(0.314)PDDDLGSSNWEADLGNEER	23	UID6	1
7	1.1441	519	O00567	519	O00567	O00567	Nucleo...	NOP56	S(0.004)FS(0.023)KEELMS(0.591)S(0.38)DLEET(0.001)AGSTSIPK	32	UID7	1
8	2.4392	208	O14813	208	O14813	O14813	Paired...	PHOX...	LS(0.5)PS(0.5)PLPVALGSGPGPGPGPLKALWAGVAGGGGGPGAGAAELLK	43	UID8	1
9	-3.9669	31	O1487...	31;31;31	O1487...	O1487...	[3-met...	BCKDK	S(0.645)T(0.274)S(0.077)AT(0.004)DTHHVEAR	45	UID9	1
10	1.4228	390	O1502...	390;39...	O1502...	O1502...	Protein...	SEC16A	QIDS(0.5)S(0.5)PVGGETDETTVSQNYR	49	UID10	1
11	0.8896	1106	O1506...	1106;1...	O1506...	O1506...	Synemin	SYNM	TPQGPVSAT(0.001)VEVS(0.408)S(0.408)PT(0.182)GFAQS(0.001)QVLEDVSQAAR	52	UID11	
12	1.2405	400	O1511...	400;38...	O1511...	O1511...	T-box t...	TBX3	IST(0.002)T(0.097)T(0.225)S(0.532)EEPCRDKGS(0.144)PAVK	54	UID12	1
13	-1.0625	319	O15127	319	O15127	O15127	Secret...	SCAM...	AAS(0.5)S(0.5)AAQGAFFQGN	56	UID13	1
14	-1.3488	320	O15127	320	O15127	O15127	Secret...	SCAM...	AAS(0.275)S(0.725)AAQGAFFQGN	57	UID14	1
15	-0.87035	301	O43765	301	O43765	O43765	Small g...	SGTA	S(0.465)RT(0.465)PS(0.07)ASNDQQE	66	UID15	
16	0.67431	305	O43765	305	O43765	O43765	Small g...	SGTA	S(0.327)RT(0.327)PS(0.327)AS(0.02)NDQQE	67	UID16	
17	-0.041...	116	O60264	116	O60264	O60264	SW/S...	SMAR...	QTELPAPGAGAS(0.107)P(0.259)S(0.034)PLK	72	UID17	1
18	1.9416	102	O6034...	102;10...	O6034...	O6034...	TBC1 d...	TBC1D4	CVPAPGAGAS(0.313)GGT(0.313)S(0.313)PS(0.041)AT(0.02)QNPFAVFIFEHK	73	UID18	
19	1.5648	106	O6034...	106;10...	O6034...	O6034...	TBC1 d...	TBC1D4	CVPAPGAGAS(0.14)GGT(0.316)S(0.316)PS(0.114)AT(0.114)QNPFAVFIFEHK	74	UID19	
20	0.57976	43	O60508	43	O60508	O60508	Pre-mR...	CDC40	S(0.507)PS(0.235)S(0.235)KPS(0.023)LAVAVDSAPEVAVKEDLETGVHLDPAVK	78	UID20	1
21	-0.45616	45	O60508	45	O60508	O60508	Pre-mR...	CDC40	S(0.323)PS(0.323)S(0.323)KPS(0.03)LAVAVDSAPEVAVKEDLETGVHLDPAVK	79	UID21	
22	-0.45616	46	O60508	46	O60508	O60508	Pre-mR...	CDC40	S(0.323)PS(0.323)S(0.323)KPS(0.03)LAVAVDSAPEVAVKEDLETGVHLDPAVK	80	UID22	
23	0.33272	417	O6078...	417;43...	O6078...	O6078...	Target...	TOM1	AADRLPNLS(0.292)S(0.646)PS(0.062)AEGPPGPPSGPAPR	85	UID23	1

2,625 items

Version 1.6.15.0

t-Test and ANOVA test require normal distributed data

The screenshot displays the Perseus software interface. At the top, there is a menu bar with 'File', 'Tools', 'Window', and 'Help'. Below the menu bar is a toolbar with various icons for data manipulation and analysis. The main window is divided into two panes. The left pane shows a list of data matrices (matrix1 through matrix6) and a 'Data' section with a table of values. The right pane shows a 'Transform' dialog box with a 'Transformation' field set to 'log2(x)' and a list of columns to be transformed. The dialog box has 'Cancel' and 'OK' buttons. The status bar at the bottom indicates '2,625 items 1 selected' and '100%' zoom.

Transform Dialog Box:

Transformation: $\log_2(x)$

Columns:

- Intensity
- Intensity 24Hrs_1_1_Slot1-16_1_136
- Intensity 24Hrs_1_2_Slot1-16_1_137
- Intensity 24Hrs_1_3_Slot1-16_1_138
- Intensity 24Hrs_2_1_Slot1-17_1_139
- Intensity 24Hrs_2_2_Slot1-17_1_140
- Intensity 24Hrs_2_3_Slot1-17_1_141
- Intensity 24Hrs_3_1_Slot1-18_1_142
- Intensity 24Hrs_3_2_Slot1-18_1_143
- Intensity 24Hrs_3_3_Slot1-18_1_144
- Intensity 2Hrs_1_1_Slot1-13_1_127
- Intensity 2Hrs_1_2_Slot1-13_1_128

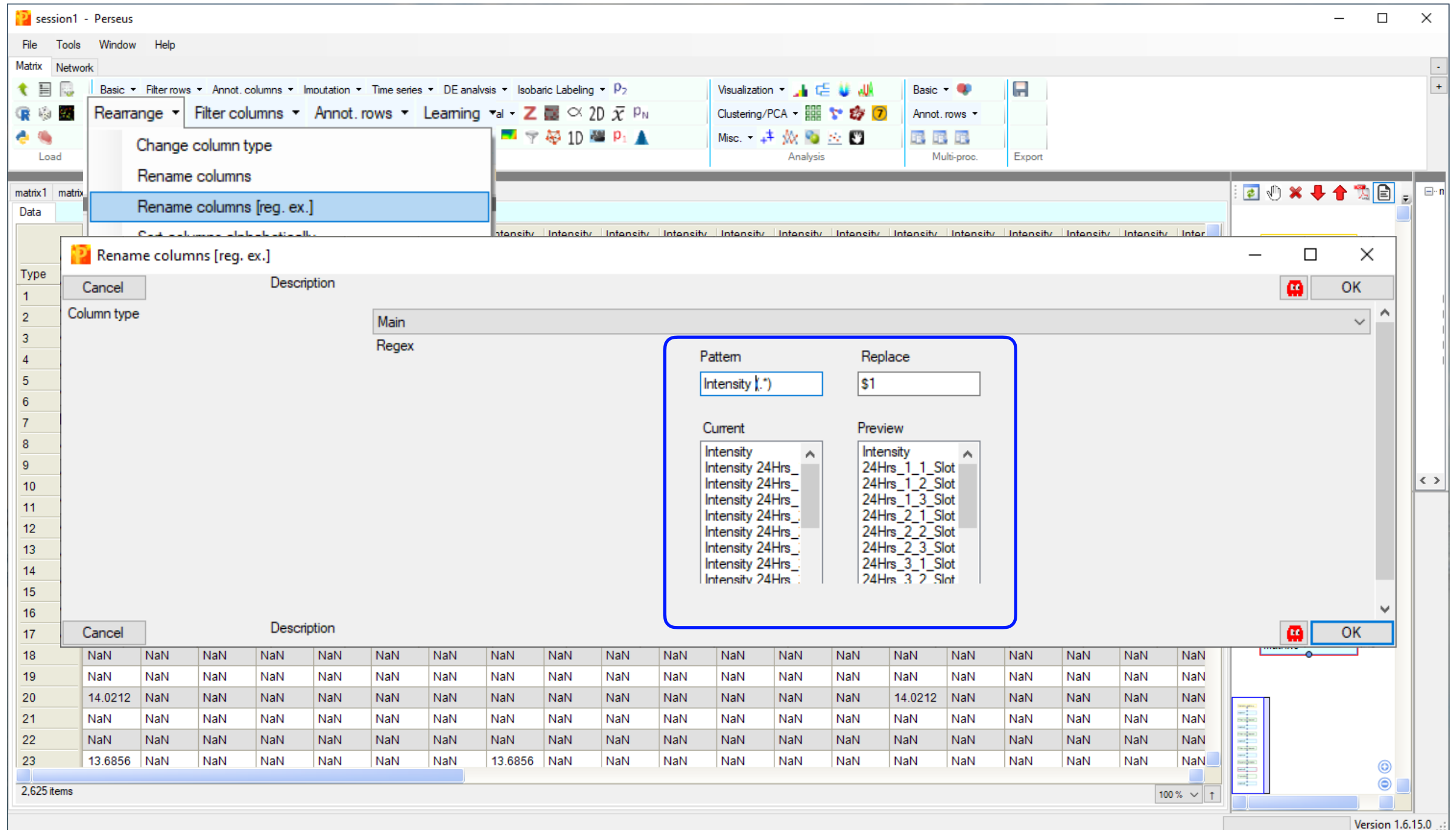
Data Table:

Type	Value	Intensity	Intensity 24Hrs_1_1_Slot1-16_1_136	Intensity 24Hrs_1_2_Slot1-16_1_137	Intensity 24Hrs_1_3_Slot1-16_1_138	Intensity 24Hrs_2_1_Slot1-17_1_139	Intensity 24Hrs_2_2_Slot1-17_1_140	Intensity 24Hrs_2_3_Slot1-17_1_141	Intensity 24Hrs_3_1_Slot1-18_1_142	Intensity 24Hrs_3_2_Slot1-18_1_143	Intensity 24Hrs_3_3_Slot1-18_1_144	Intensity 2Hrs_1_1_Slot1-13_1_127	Intensity 2Hrs_1_2_Slot1-13_1_128
1	0	32708	0	0	0	0	0	0	0	0	0	0	0
2	1	58763	0	0	0	0	0	0	0	0	0	0	26055
3	0	0	0	0	0	0	0	0	0	0	0	0	0
4	3	0	0	0	0	0	0	0	0	0	0	0	0
5	8	0	0	0	0	0	0	0	0	0	0	0	0
6	5	0	0	0	0	0	0	0	0	0	0	0	0
7	2	38467	0	0	0	0	0	0	0	0	0	0	0
8	6	0	0	0	0	0	0	0	0	0	0	0	0
9	7	0	0	0	0	0	0	0	0	0	0	0	0
10	1	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0
12	2	0	0	0	0	0	0	0	0	0	0	0	0
13	2	32708	0	0	0	0	0	0	0	0	0	0	0
14	0	58763	0	0	0	0	0	0	0	0	0	0	26055
15	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	38467	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	16626	0	0	0	0	0	0	0	0	0	0	16626
21	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	13176	0	0	0	0	0	0	0	0	0	0	13176

2,625 items 1 selected

Version 1.6.15.0

Cosmetics - Shorten the Sample Names: **Processing -> Rearrange -> Rename columns (reg.ex.)**



The screenshot shows the Perseus software interface. The 'Rename columns [reg. ex.]' dialog box is open, and the 'Pattern' field contains 'Intensity (.*)'. The 'Replace' field contains '\$1'. The 'Current' list shows 'Intensity' and 'Intensity 24Hrs_'. The 'Preview' list shows 'Intensity' and '24Hrs_1_1_Slot'. The background shows the main Perseus window with a data table and various toolbars.

Type	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	14.0212	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.0212	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	13.6856	NaN	NaN	NaN	NaN	NaN	NaN	NaN	13.6856	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

2,625 items

Version 1.6.15.0

Adding a Series Protein Annotations: **Processing -> Annot. columns -> Add annotations**

Download annotation

Annotation files for all species derived from UniProt can be downloaded from <http://annotations.perseus-framework.org>

Put the downloaded annotation files into Perseus\conf\annotations.

matrix1

Data

Type	Main	Main	Main	Main	Main	Main	Main	24Hrs_...	24Hrs_...	24Hrs_...	2Hrs_1...	2Hrs_1_2_Slot1-13_1_128	2Hrs_1...
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	17.185	NaN	NaN	16.7526	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	14.9953	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	16.3362	14.5515	NaN	NaN	NaN	13.9503	NaN	NaN	14.3412	NaN	NaN	NaN	NaN
6	19.1595	NaN	16.7004	NaN	NaN	NaN	NaN	18.7213	NaN	NaN	NaN	NaN	NaN
7	14.6467	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.6467	NaN
8	19.253	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	12.8393	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	12.8393	NaN
10	13.8556	NaN	NaN	NaN	NaN	NaN	13.8556	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12	14.5067	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.5067
13	14.9974	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.9974
14	15.8426	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.6693	NaN	14.9974
15	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
17	15.2313	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
20	14.0212	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
23	13.6856	NaN	NaN	NaN	NaN	NaN	NaN	13.6856	NaN	NaN	NaN	NaN	NaN

2,625 items

100%

Version 1.6.15.0

Generic matrix u...

matrix1

Filter rows based...

matrix2

Filter rows based...

matrix3

Filter rows based...

matrix4

Expand site table

matrix5

Transform

matrix6

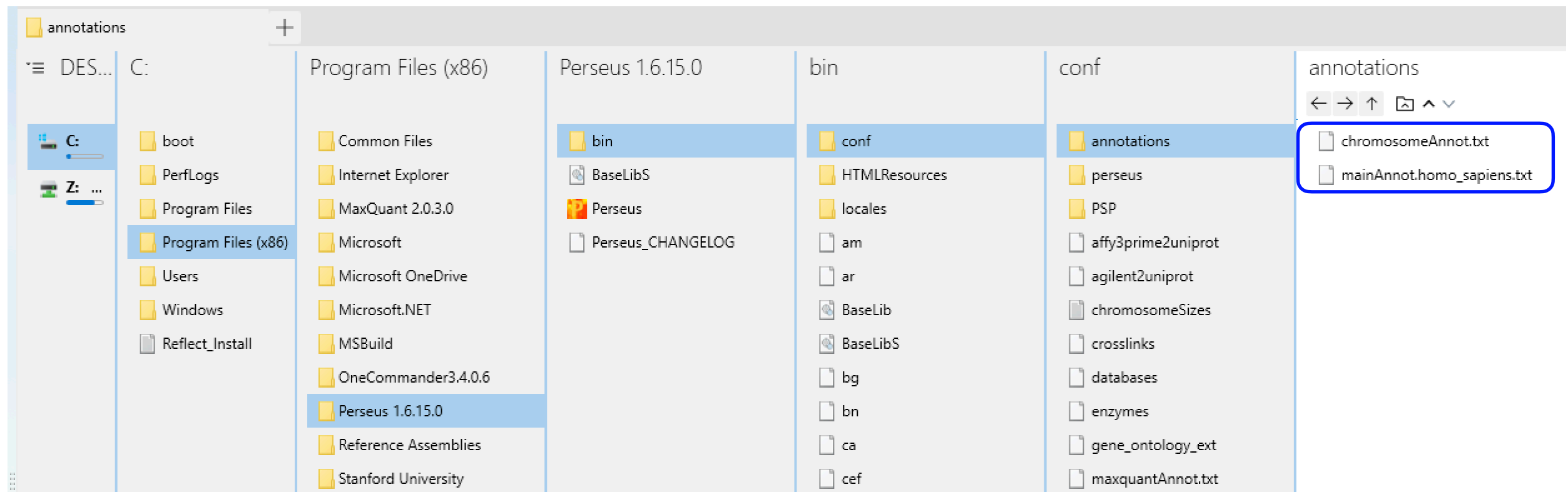
Rename column...

matrix7

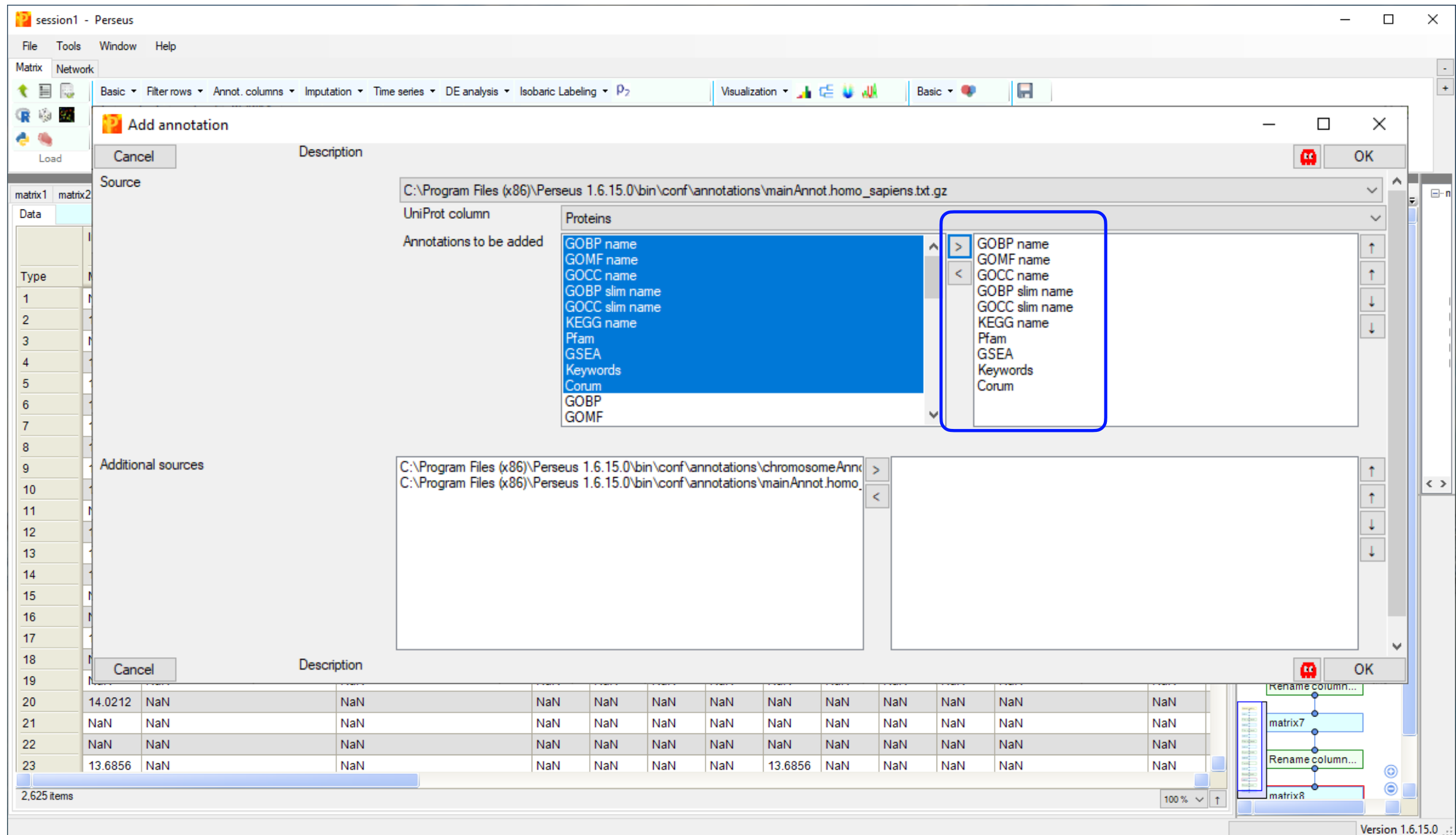
Rename column...

matrix8

Adding a Series Protein Annotations: **Processing -> Annot. columns -> Add annotations**



Adding a Series of Protein Annotations: **Processing -> Annot. columns -> Add annotations**



The screenshot displays the Perseus software interface. The 'Add annotation' dialog box is open, showing the 'Annotations to be added' section. A blue box highlights the following annotations:

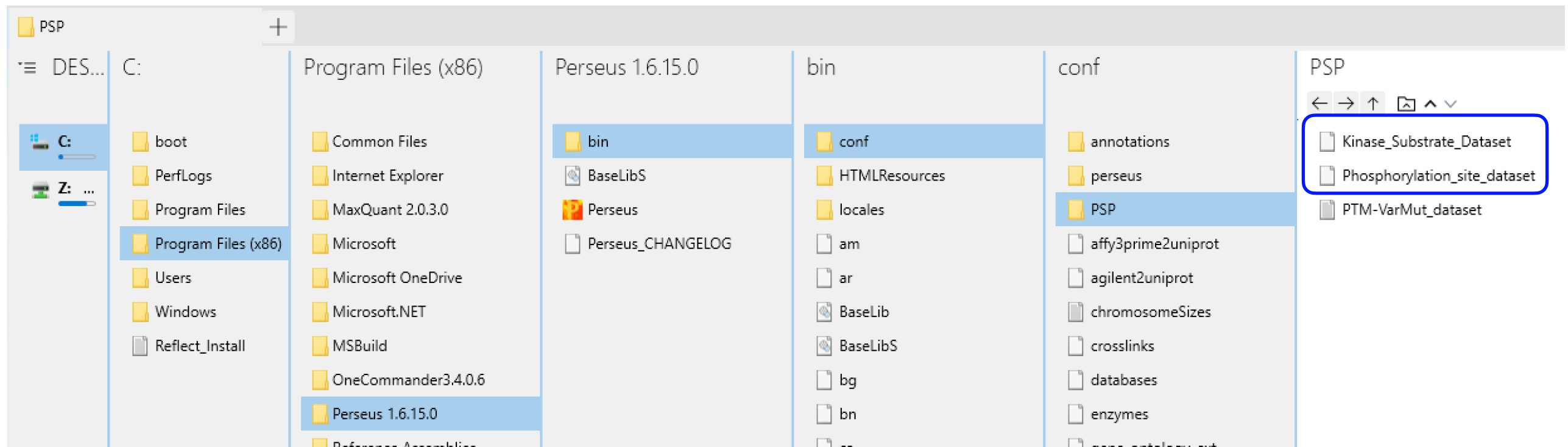
- GOBP name
- GOMF name
- GOCC name
- GOBP slim name
- GOCC slim name
- KEGG name
- Pfam
- GSEA
- Keywords
- Corum

The background shows a data matrix with 2,625 items and a network diagram. The matrix has columns labeled 'matrix1' and 'matrix2'. The network diagram shows nodes labeled 'matrix7' and 'matrix8' connected by edges.

Adding Peptide Specific Annotations: **Processing -> Modifications -> Add sequence features**



<https://www.phosphosite.org/>



Adding Peptide Specific Annotations: **Processing -> Modifications -> Add sequence features**

session1 - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series DE analysis Isobaric Labeling P₂

Rearrange Filter columns Annot. rows Learning Modifications External Z < 2D x P_N

Normalization Quality Tests Outliers WGCN

Load

Visualization Clustering/PCA Misc. Analysis Basic Annot. rows Multi-proc. Export

matrix1 matrix2 matrix3 matrix4 matrix5 matrix6 matrix7 matrix8

Data

Type	N: Score for localization	N: Mass error [ppm]	N: Intensity	N: Position	T: Protein	T: Position within protein	T: Leading proteins	T: Protein	T: Protein names	T: Phospho (STY) Probabilities
	Numeric	Numeric	Numeric	Numeric	Text	Text	Text	Text	Text	Text
1	84.071	-11.061	0	95	P0DN37	95;95	P0DN37	P0DN37	Peptid...	HT(0.01)GS(0.49)GILS(0.49)MANAGPNT(0.01)NGSQFFICTAK
2	92.15	-10.63	149000	99	P0DN37	99;99	P0DN37	P0DN37	Peptid...	HT(0.032)GS(0.265)GILS(0.669)MANAGPNT(0.034)NGSQFFICTAK
3	53.028	9.0084	2071100	153	A5YM72;A5YM72-4;A5YM72-5;A5YM72-3	153;27...	A5YM72	A5YM72	Carnos...	EGQET(1)LVKEEVEAFLRS(1)EALGDILQVAVK

Expand site table
Add linear motifs
Add known sites
Add modification counts
Kinase-substrate relations
Add sequence features
Add regulatory sites
Shorten motif length

Site-spe

Add sequence features

Cancel Description OK

Proteins Proteins

Positions within proteins Positions within proteins

Add status column ☐

Cancel Description OK

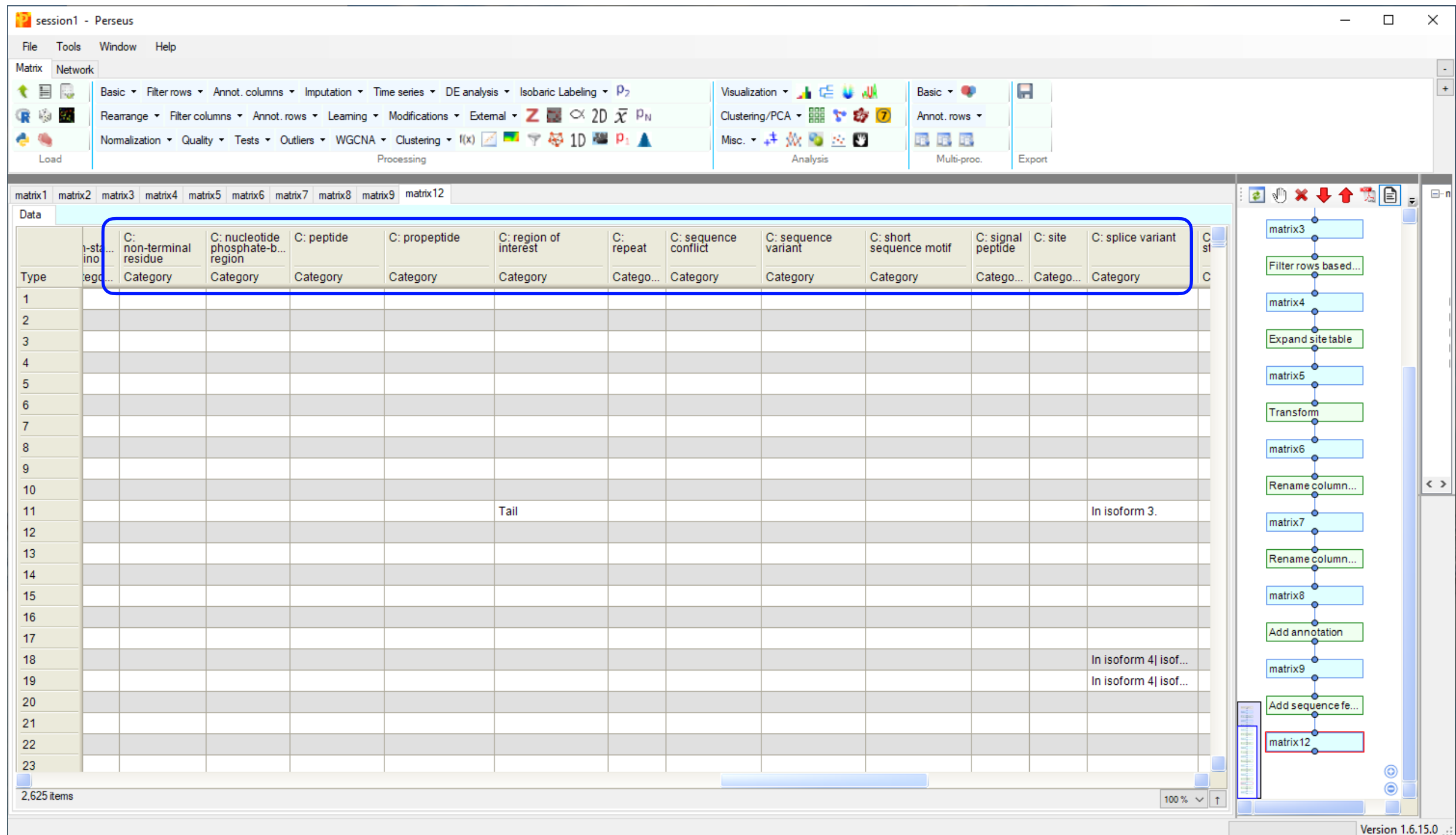
	N: Score for localization	N: Mass error [ppm]	N: Intensity	N: Position	T: Protein	T: Position within protein	T: Leading proteins	T: Protein	T: Protein names	T: Phospho (STY) Probabilities
	Numeric	Numeric	Numeric	Numeric	Text	Text	Text	Text	Text	Text
13	73.532	2.0266	714850	115	CON__P02662	115	CON__...	CON__...		VPQLEIVPNS(1)AEER
14	86.997	-0.72796	10735	35	CON__P02666	35	CON__...	CON__...		FQS(1)EEQQQTEDELQDK
15	214.32	-1.8056	1222600	138	CON__P12763;P02765	138;138	CON__...	CON__...	Alpha-...	CDSSPDS(1)AEDVRK
16	90.806	1.7791	9227.2	320	CON__P12763	320	CON__...	CON__...		HT(0.005)FS(0.015)GVAS(0.98)VESSSGEAFHVGK
17	159.44	1.6203	58430	323	CON__P12763	323	CON__...	CON__...		HTFSGVASVES(0.918)S(0.081)S(0.001)GEAFHVGK
18	63.355	-0.94033	0	324	CON__P12763	324	CON__...	CON__...		HTFSGVAS(0.004)VES(0.463)S(0.463)S(0.069)GEAFHVGK
19	66.41	2.1304	197040	44	O00178	44	O00178	O00178	GTP-b...	LHGGFDS(1)DCS(1)EDGEALNGEPELDLTSK
20	66.41	2.1304	197040	47	O00178	47	O00178	O00178	GTP-b...	LHGGFDS(1)DCS(1)EDGEALNGEPELDLTSK
21	51.829	-0.41792	32661	580	O00178	580	O00178	O00178	GTP-b...	LLQTT(0.004)NNS(0.708)PMNS(0.288)KPQIQ
22	45.826	1.3373	82734	24	O00178	24	O00178	O00178	GTP-b...	SAMDSPVPAS(0.002)MFAPEPS(0.608)S(0.39)PGAAR
23	91.1	0.94941	641450	25	O00178	25	O00178	O00178	GTP-b...	SAMDSPVPASMFAPPEPS(0.072)S(0.928)PGAAR

2,444 items 1 selected

100% ↑

Version 1.6.15.0

Adding Peptide Specific Annotations: **Processing** -> **Modifications** -> **Add sequence features**



The screenshot displays the Perseus software interface. The top menu bar includes File, Tools, Window, and Help. Below it, the 'Matrix' and 'Network' tabs are visible. The 'Processing' workflow is active, showing a sequence of steps: matrix1, matrix2, matrix3, matrix4, matrix5, matrix6, matrix7, matrix8, matrix9, and matrix12. The 'Add sequence features' step is highlighted in red.

The main data table shows peptide annotations. The columns are categorized as follows:

- C: non-terminal residue** (Category)
- C: nucleotide phosphate-b... region** (Category)
- C: peptide** (Category)
- C: propeptide** (Category)
- C: region of interest** (Category)
- C: repeat** (Category)
- C: sequence conflict** (Category)
- C: sequence variant** (Category)
- C: short sequence motif** (Category)
- C: signal peptide** (Category)
- C: site** (Category)
- C: splice variant** (Category)

The table contains 23 rows of data. Row 11 shows 'Tail' under 'C: region of interest' and 'In isoform 3.' under 'C: splice variant'. Row 18 shows 'In isoform 4| isof...' under 'C: splice variant'. Row 19 shows 'In isoform 4| isof...' under 'C: splice variant'. The status bar at the bottom indicates '2,625 items' and 'Version 1.6.15.0'.

Adding Peptide Specific Annotations: **Processing -> Modifications -> Add linear motifs, Add known sites, Kinase-substrate relations**

Phospho - Analysis - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series DE analysis Isobaric Labeling P₂

Rearrange Filter columns Annot. rows Learning Modifications External Z 2D X P_N

Normalization Quality Tests Outliers WGCNA Clustering f(x) 1D P₁

Load Processing Analysis Multi-proc. Export

matrix1 matrix2 matrix3 matrix4 matrix5 matrix6 matrix7 matrix8 matrix9 matrix12 matrix13 matrix14 matrix15

Data

Type	C: splice variant	C: strand	C: topolog... domain	C: transit peptide	C: transmem... region	C: turn	C: unsure residue	C: zinc finger region	C: Motifs	C: Known site	C: Origin	N: Localiz... prob	N: PEP	N: Score	N: Delta score
	Catego...	Catego...	Catego...	Catego...	Catego...	Catego...	Catego...	Category	Category	Category	Catego...	Numeric	Numeric	Numeric	Numeric
1									G protein-coupled receptor kinase 1 subs...	+	HTP	0.4897...	1.4862...	84.071	75.583
2									G protein-coupled receptor kinase 1 subs...	+	HTP	0.6689	4.0908...	92.15	83.662
3									Casein kinase I substrate motif			0.34568	0.0081...	45.826	36.67
4									ERK1,2 kinase substrate motif;G protein...	+	CST;H...	0.7083...	0.0151...	51.829	36.301
5									DNA dependent Protein kinase substrate...	+	CST;H...	0.6078...	1.0756...	61.503	56.552
6									Casein kinase II substrate motif;PKA kina...	+	CST;H...	0.6861...	2.1224...	159.84	154.5
7									Casein kinase I substrate motif	+	CST;H...	0.5908...	4.4763...	72.311	63.546
8									DNA dependent Protein kinase substrate...	+	HTP;L...	0.4999...	4.9631...	165.61	164.06
9									14-3-3 domain binding motif;Calmodulin...	+	HTP	0.6445...	0.0015...	65.12	62.819
10									b-Adrenergic Receptor kinase substrate m...	+	HTP	0.4999...	0.0012...	66.509	63.435
11	In isofo...								Casein kinase II substrate motif			0.4079...	7.8047...	65.942	63.692
12									Casein kinase II substrate motif;G protein...	+	CST;H...	0.5320...	0.0012...	41.87	21.979
13			Cytopl...						14-3-3 domain binding motif;Calmodulin...	+	CST;H...	0.5	5.2750...	178.52	132.46
14			Cytopl...						MDC1 BRCT domain binding motif;Plk1 P...	+	CST;H...	0.7250...	5.2750...	178.52	132.46
15									Casein kinase II substrate motif;G protein...	+	CST;H...	0.4649...	0.0003...	86.952	44.891
16									14-3-3 domain binding motif;Akt kinase s...	+	CST;H...	0.3268...	0.0041...	73.305	41.255
17									Cdc2 kinase substrate motif;CDK kinase s...	+	CST;H...	0.6343...	2.0786...	89.636	85.316
18	In isofo...								Casein kinase II substrate motif;G protein...	+	HTP	0.3128...	0.0012...	42.888	41.515
19	In isofo...								Casein kinase II substrate motif;ERK1,2 k...			0.3155...	0.0011...	49.872	45.631
20									Casein kinase II substrate motif;ERK1,2 k...	+	HTP	0.5067...	3.3131...	87.097	85.936
21									Calmodulin-dependent protein kinase II s...	+	HTP	0.3234...	2.1850...	75.522	72.866
22									Casein kinase II substrate motif;MDC1 B...	+	HTP	0.3234...	2.1850...	75.522	72.866
23									Casein kinase II substrate motif;ERK1,2 k...	+	CST;H...	0.64569	3.8228...	55.637	55.195

2,625 items 1 selected

100% ↑

Version 1.6.15.0

matrix3

Filter rows based...

matrix4

Expand site table

matrix5

Transform

matrix6

Rename column...

matrix7

Rename column...

matrix8

Add annotation

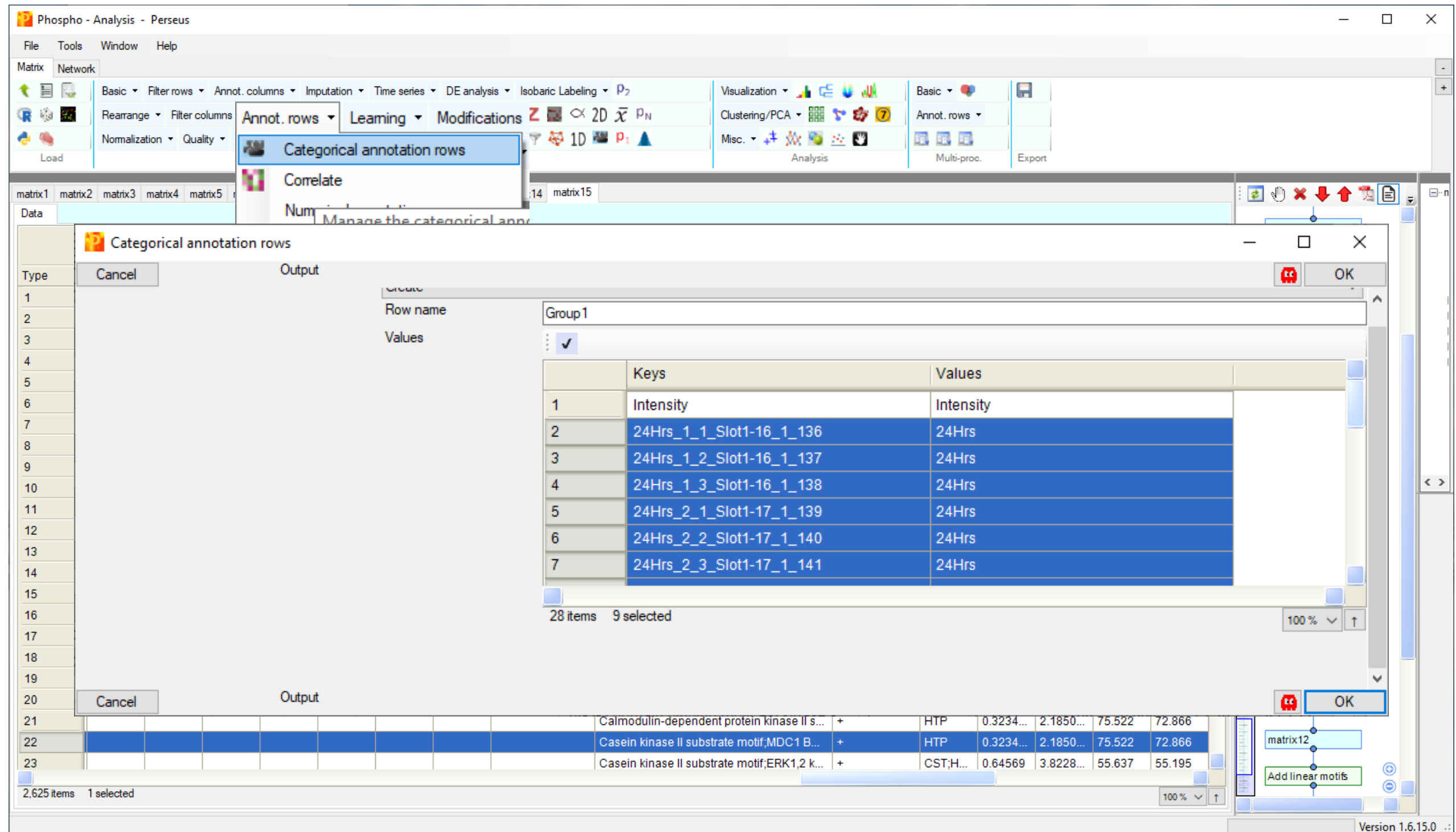
matrix9

Add sequence fe...

matrix12

Add linear motifs

Sample Grouping (technical and biological replicates !): **Annot. Rows > Categorical Annotation Row**



Phospho - Analysis - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series DE analysis Isobaric Labeling P₂

Rearrange Filter columns Annot. rows Learning Modifications Z α 2D x̄ P_N

Normalization Quality

Load

Visualization Clustering/PCA Misc. Analysis Multi-proc. Export

matrix1 matrix2 matrix3 matrix4 matrix5

Data

Correlate

Num

Manage the categorical annotation rows

Categorical annotation rows

Cancel Output

Group1

✓

	Keys	Values
1	Intensity	Intensity
2	24Hrs_1_1_Slot1-16_1_136	24Hrs
3	24Hrs_1_2_Slot1-16_1_137	24Hrs
4	24Hrs_1_3_Slot1-16_1_138	24Hrs
5	24Hrs_2_1_Slot1-17_1_139	24Hrs
6	24Hrs_2_2_Slot1-17_1_140	24Hrs
7	24Hrs_2_3_Slot1-17_1_141	24Hrs

28 items 9 selected

100 %

Cancel Output

21	Calmodulin-dependent protein kinase II s...	+	HTP	0.3234...	2.1850...	75.522	72.866
22	Casein kinase II substrate motif;MDC1 B...	+	HTP	0.3234...	2.1850...	75.522	72.866
23	Casein kinase II substrate motif;ERK1,2 k...	+	CST;H...	0.64569	3.8228...	55.637	55.195

2,625 items 1 selected

100 %

matrix12

Add linear motifs

Version 1.6.15.0

Data Quality: **Filter Rows > Filter rows based on valid numbers**

Phospho - Analysis - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series DE analysis Isobaric Labeling P₂

Rearrange Filter columns Annot. rows Learning Modifications External Z 2D X P_N

Normalization Quality Tests Outliers WGCNA Clustering f(x) 1D P₁

Load Processing

Visualization Clustering/PCA Misc. Analysis

Basic Annot. rows Multi-proc. Export

matrix1 matrix2 matrix3 matrix4 matrix5 matrix6 matrix7 matrix8 matrix9 matrix12 matrix13 matrix14 matrix15 matrix16 matrix18

Data

Type	Intensity	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	2Hrs_1...	2Hrs_1...	2Hrs_1...	2Hrs_2...	2Hrs_2...	2Hrs_2...	2Hrs_3...	2Hrs_3...	2Hrs_3...
Group1	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main
1	16.3362	14.5515	NaN	NaN	NaN	13.9503	NaN	NaN	14.3412	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	19.1595	NaN	16.7004	NaN	NaN	NaN	NaN	18.7213	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	19.253	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	15.8426	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.6693	NaN	14.9974	NaN	NaN	NaN	NaN	NaN	NaN
5	15.6651	NaN	14.096	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	13.8921	13.8594	NaN	NaN	NaN	12.0971	NaN	NaN
6	13.225	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7	13.225	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
8	17.721	14.3159	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	15.4932	15.4915	NaN
9	14.4723	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10	15.4102	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11	16.7913	NaN	NaN	NaN	NaN	NaN	14.6923	NaN	14.2517	14.3274	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12	16.9228	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
13	16.3669	NaN	NaN	NaN	NaN	NaN	NaN	13.3486	13.3574	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
14	16.7568	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	15.6584	NaN	15.849	NaN	NaN	NaN	NaN	NaN	NaN
15	15.9293	NaN	NaN	NaN	NaN	NaN	NaN	15.0488	NaN	14.799	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
16	19.6875	NaN	NaN	NaN	NaN	NaN	17.0285	17.9613	NaN	18.0875	17.4329	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
17	14.3207	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	13.4086	13.227	NaN	NaN	NaN	NaN	NaN	NaN
18	15.4789	NaN	NaN	NaN	NaN	NaN	14.1445	NaN	NaN	14.7502	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
19	19.525	NaN	18.0916	NaN	NaN	17.8682	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
20	16.0793	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	10.685	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
21	15.535	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.4274	14.6352	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
22	14.7733	NaN	NaN	12.9706	12.1929	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	12.8113	NaN	NaN	NaN	NaN

106 items

100% ↑

Version 1.6.15.0

matrix8

Add annotation

matrix9

Add sequence fe...

matrix12

Add linear motifs

matrix13

Add known sites

matrix14

Kinase-substrate...

matrix15

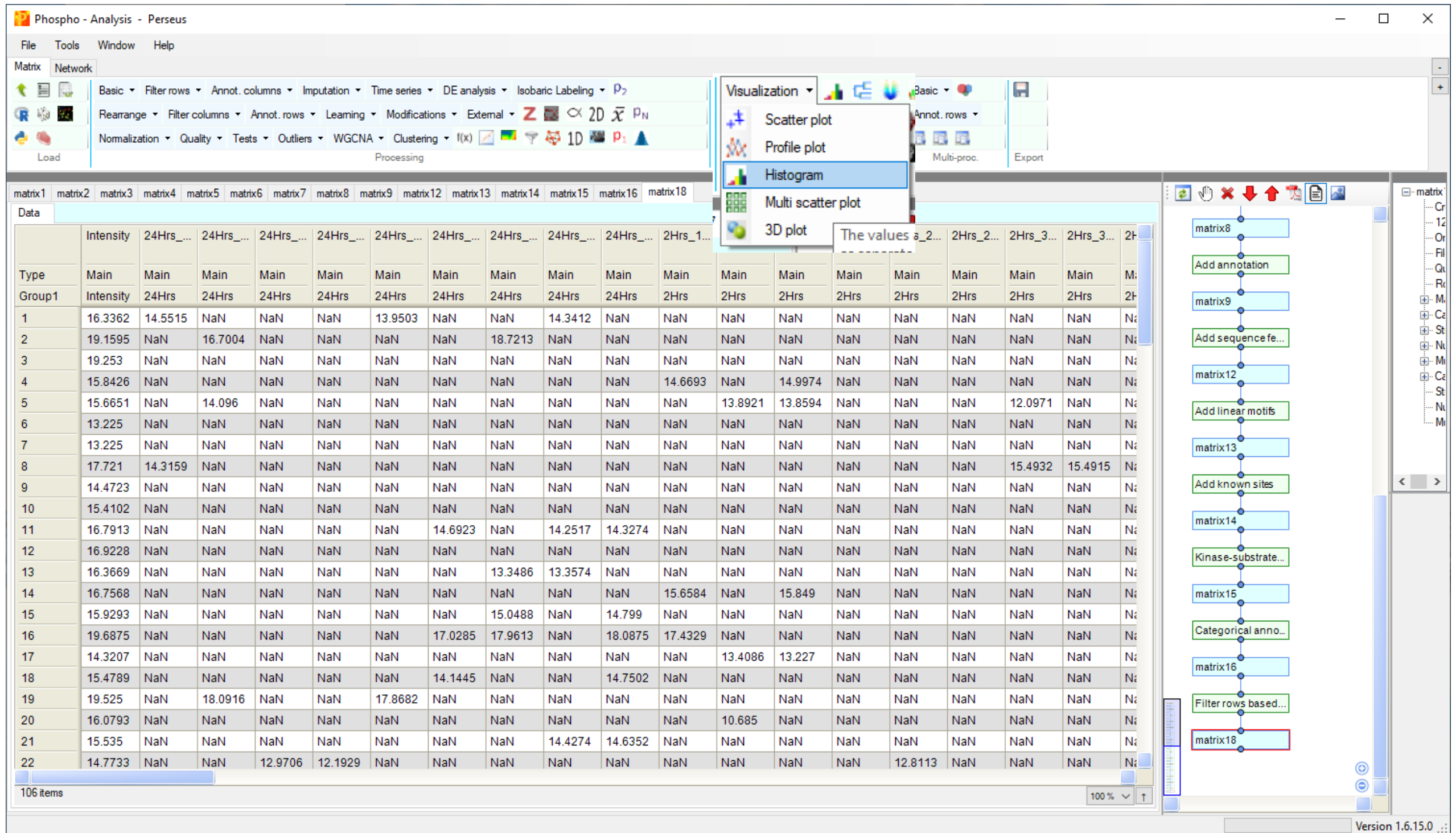
Categorical anno...

matrix16

Filter rows based...

matrix18

Data Quality - Normal distributed values ? : **Visualisation - Histogram**



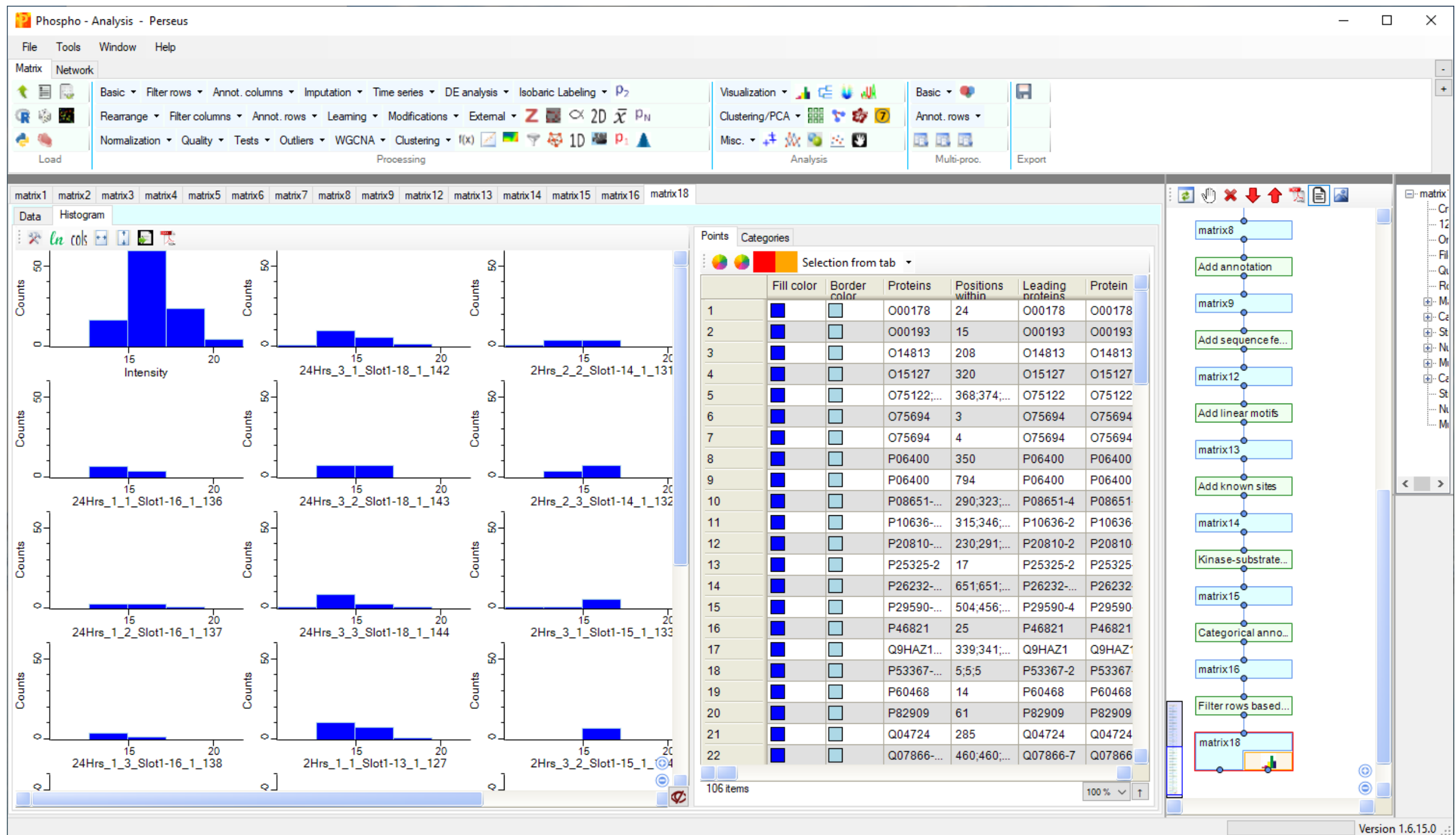
The screenshot displays the Perseus software interface, specifically the 'Phospho - Analysis - Perseus' window. The main area shows a data matrix with 22 rows and 18 columns. The columns are labeled 'Intensity', '24Hrs_...', '24Hrs_...', '24Hrs_...', '24Hrs_...', '24Hrs_...', '24Hrs_...', '24Hrs_...', '24Hrs_...', '24Hrs_...', '24Hrs_...', '24Hrs_1...', '24Hrs_1...', '24Hrs_1...', '24Hrs_1...', '24Hrs_1...', '24Hrs_1...', and '24Hrs_1...'. The rows are numbered 1 to 22. The matrix contains numerical values, with many cells showing 'NaN'.

A 'Visualization' menu is open, showing options: 'Scatter plot', 'Profile plot', 'Histogram' (highlighted), 'Multi scatter plot', and '3D plot'. The 'Histogram' option is selected, and a tooltip for 'The values a_2...' is visible.

On the right side, a workflow diagram is visible, showing a sequence of steps: 'matrix8', 'Add annotation', 'matrix9', 'Add sequence fe...', 'matrix12', 'Add linear motifs', 'matrix13', 'Add known sites', 'matrix14', 'Kinase-substrate...', 'matrix15', 'Categorical anno...', 'matrix16', 'Filter rows based...', and 'matrix18'.

The bottom status bar indicates '106 items' and 'Version 1.6.15.0'.

Data Quality - Normal distributed values ? : **Visualisation - Histogram**



Data Quality: **Visualisation - Multi scatter plot**

Phospho - Analysis - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series DE analysis Isobaric Labeling P₂

Rearrange Filter columns Annot. rows Learning Modifications External Z 2D x̄ P_N

Normalization Quality Tests Outliers WGCNA Clustering f(x) 1D P₁

Load Processing

Visualization Scatter plot Profile plot Histogram

Basic Annot. rows Multi-proc. Export

matrix1 matrix2 matrix3 matrix4 matrix5 matrix6 matrix7 matrix8 matrix9 matrix12 matrix13 matrix14 matrix15 matrix16 matrix18

Multi scatter plot

Cancel Description OK

Rows

Columns

Intensity

24Hrs_1_1_Slot1-16_1_136

24Hrs_1_2_Slot1-16_1_137

24Hrs_1_3_Slot1-16_1_138

24Hrs_2_1_Slot1-17_1_139

24Hrs_2_2_Slot1-17_1_140

24Hrs_2_3_Slot1-17_1_141

24Hrs_3_1_Slot1-18_1_142

24Hrs_3_2_Slot1-18_1_143

24Hrs_3_3_Slot1-18_1_144

2Hrs_1_1_Slot1-13_1_127

2Hrs_1_2_Slot1-13_1_128

Intensity

24Hrs_1_1_Slot1-16_1_136

24Hrs_1_2_Slot1-16_1_137

24Hrs_1_3_Slot1-16_1_138

24Hrs_2_1_Slot1-17_1_139

24Hrs_2_2_Slot1-17_1_140

24Hrs_2_3_Slot1-17_1_141

24Hrs_3_1_Slot1-18_1_142

24Hrs_3_2_Slot1-18_1_143

24Hrs_3_3_Slot1-18_1_144

2Hrs_1_1_Slot1-13_1_127

2Hrs_1_2_Slot1-13_1_128

Cancel Description OK

19 19.525 NaN 18.0916 NaN NaN 17.8682 NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN

20 16.0793 NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN 10.685 NaN NaN NaN NaN NaN NaN

21 15.535 NaN NaN NaN NaN NaN NaN NaN NaN 14.4274 14.6352 NaN NaN NaN NaN NaN NaN

22 14.7733 NaN NaN 12.9706 12.1929 NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN NaN

106 items

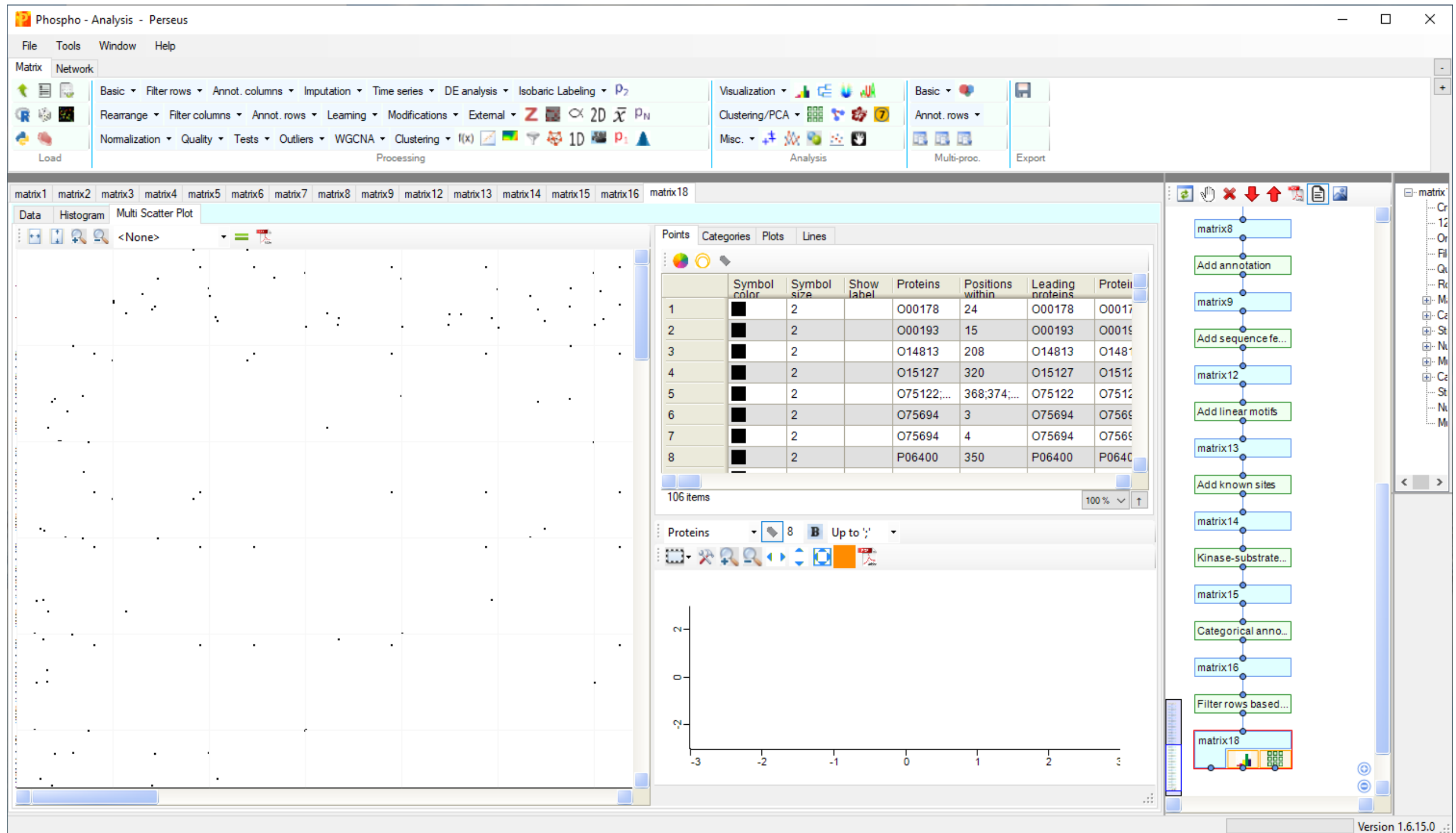
100 %

Version 1.6.15.0

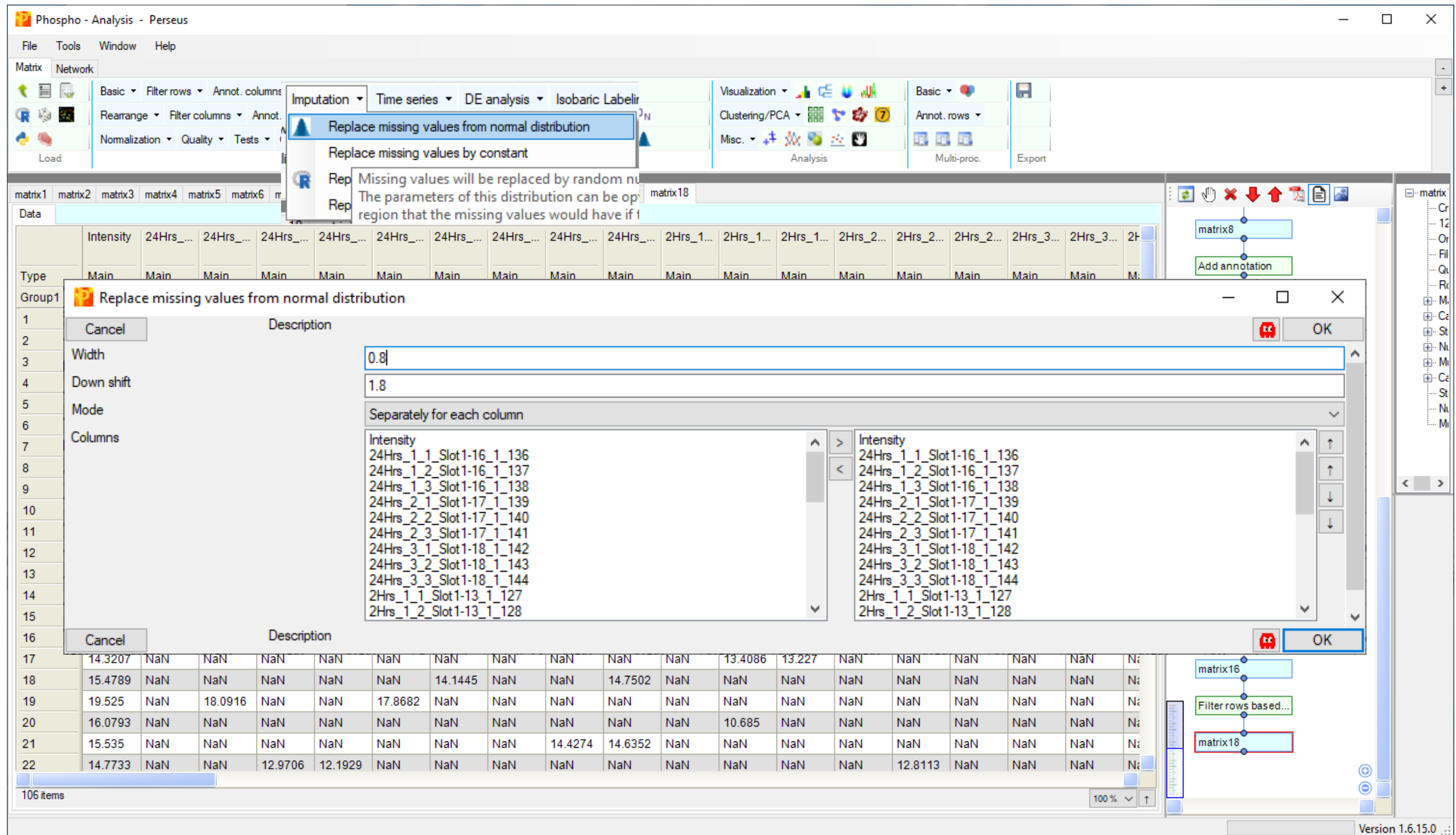
Filter rows based...

matrix18

Data Quality: **Visualisation - Multi scatter plot**



Data Quality - Fill in missing values : **Imputation > Replace missing values from normal distribution**



The screenshot shows the Perseus software interface with the 'Imputation' menu open and the 'Replace missing values from normal distribution' dialog box displayed. The dialog box has the following fields and options:

- Width:** 0.8
- Down shift:** 1.8
- Mode:** Separately for each column
- Columns:** A list of columns including Intensity, 24Hrs_1_1_Slot1-16_1_136, 24Hrs_1_2_Slot1-16_1_137, 24Hrs_1_3_Slot1-16_1_138, 24Hrs_2_1_Slot1-17_1_139, 24Hrs_2_2_Slot1-17_1_140, 24Hrs_2_3_Slot1-17_1_141, 24Hrs_3_1_Slot1-18_1_142, 24Hrs_3_2_Slot1-18_1_143, 24Hrs_3_3_Slot1-18_1_144, 2Hrs_1_1_Slot1-13_1_127, and 2Hrs_1_2_Slot1-13_1_128.

The background shows a data table with columns labeled 'Intensity', '24Hrs_...', and '2Hrs_...'. The table contains numerical values and 'NaN' entries. The status bar at the bottom indicates '106 items' and 'Version 1.6.15.0'.

Processing > Tests > Multiple-sample tests

Phospho - Analysis - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series DE analysis Isobaric Labeling P₂

Rearrange Filter columns

Normalization Quality Test

Load

Visualizations

Clustering/PCA

Misc.

Analysis

Basic

Annot. rows

Multi-proc.

Export

matrix1 matrix2 matrix3 matrix4 matrix5 matrix6 matrix7 matrix8 matrix9 matrix10 matrix11 matrix12 matrix13 matrix14 matrix15 matrix16 matrix17 matrix18 matrix19

24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_...

Intensity 24Hrs_...

Main Main Main Main Main Main Main Main Main Main Main Main Main Main Main Main Main Main Main Main

Group1 Intensity 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_... 24Hrs_...

1 19

2 19

3 20

4 16

5 19

6 15

7 15

8 15

9 16

10 16

11 15

12 16

13 15

14 15

15 15

16 16

17 17

18 16

19 14

20 14

21 15.4102 13.2555 11.4385 13.3443 11.3483 12.7971 13.2011 9.57164 13.2838 13.7657 12.4571 10.5595 11.4486 12.2673 9.36997 13.7042 13.2317 15.2439 13.1481

22 18.1849 15.9343 11.9368 13.7326 12.9803 16.2116 13.9156 10.7019 15.8074 13.725 11.2435 10.1573 12.3782 14.8243 14.4916 13.0312 11.4811 16.1581 14.1481

106 items 1 selected

100 %

Version 1.6.15.0

Multiple-sample tests

Cancel

Output

Group1

ANOVA

S0

0

Permutation-based FDR

FDR

0.05

Report q value

☒

Number of randomizations

250

Preserve grouping in randomizations

<None>

-Log10

☒

Suffix

Cancel

Output

OK

matrix18

Replace missing



Statistical Data Analysis:

Phospho - Analysis - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series DE analysis Isobaric Labeling P₂

Rearrange Filter columns Annot. rows Learning Modifications External Z 2D X P_N

Normalization Quality Tests Outliers WGCNA Clustering f(x) 1D P₁

Visualization Clustering/PCA Misc. Basic Annot. rows

Load Processing Analysis Multi-proc. Export

matrix1 matrix2 matrix3 matrix4 matrix5 matrix6 matrix7 matrix8 matrix9 matrix12 matrix13 matrix14 matrix15 matrix16 matrix18 matrix19 matrix20

Data

Type	go...	C: unsure residue	C: zinc finger region	C: Motifs	C: Known site	C: Origin	C: ANOVA Significant	N: Localiz... pro	N: PEP	N: Score	N: Delta score	N: Score for localiza...	N: Mass error [ppm]	N: Position	N: -Log ANOVA p value	N: ANOVA q-value	T: Proteins
Group1		Catego...	Catego...	Catego...	Catego...	Catego...	Category	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Text
1				DNA d...	+	CST;H...		0.6078...	1.0756...	61.503	56.552	45.826	1.3373	24	0.9993...	0.3652...	O00178
2				b-Adre...	+	HTP		0.7113...	7.0314...	86.997	78.509	86.997	1.7872	979	0.1665...	0.7525...	Q9Y490
3				BARD1...	+	HTP		0.7279...	0.0003...	116.95	93.17	116.95	2.0009	11	1.80569	0.347	Q9Y3B9
4				Cdc2 k...	+	CST;H...		0.6513...	1.8319...	127.07	96.195	92.707	-1.7029	188	1.05011	0.3516...	Q9Y388
5				ATM k...	+	CST;H...		0.7245...	0.0002...	97.949	75.226	84.58	0.89152	406	0.7574...	0.4584...	Q9Y2W...
6				MDC1...				0.7366...	8.0542...	92.707	83.193	92.707	-0.094...	689	0.5923...	0.5258...	Q9Y2J...
7				Casein...				0.6532...	1.0697...	104.15	85.148	95.263	-1.1235	147	2.23524	0.232	Q9UPY...
8				Casein...	+	CST;H...		0.6969...	3.2324...	69.215	56.75	69.215	-1.0723	604	0.2218	0.7333...	Q9UM...
9				BARD1...	+	HTP		0.5612...	2.4838...	79.653	60.94	79.653	-0.2012	898	0.1381...	0.7821...	Q9UKV...
10				14-3-3...	+	CST;H...		0.67465	0.0010...	83.077	77.513	59.357	3.3449	89	0.6407...	0.51	Q9UKI2
11				Cdc2 k...	+	CST;H...		0.6607...	0.0066...	71.912	65.884	71.912	0.12245	1335	0.1871...	0.7626...	Q9UG...
12				Casein...	+	CST;H...		0.7473...	7.3163...	103.23	92.745	103.23	-0.54301	510	1.72583	0.3392	Q9Y6G9
13				Casein...	+	CST;H...		0.7002...	0.0002...	87.232	80.377	87.232	0.17994	50	0.5158...	0.5326...	Q9NX63
14				Casein...	+	CST;H...		0.7172...	8.9183...	70.552	65.21	70.552	1.0025	598	1.32437	0.27175	Q9H2Y...
15				Cdc2 k...	+	HTP		0.7330...	3.6595...	152.59	139.93	78.75	0.67129	151	0.1148...	0.8024	Q9H2U...
16				MDC1...	+	HTP		0.7275...	1.8715...	172.47	149.09	100.6	1.2404	150	0.2853...	0.7188...	Q9H2U...
17				Cdc2 k...	+	CST;H...		0.5698...	4.7809...	99.909	85.905	99.909	0.76708	94	0.2550...	0.7511...	Q9GZR...
18				Casein...	+	HTP		0.7187...	1.7134...	122.2	107.83	122.2	-0.57429	462	0.2016...	0.7630...	Q9BZZ...
19				AMP-a...	+	CST;H...		0.5	4.5211...	69.94	68.208	69.94	0.0526...	127	0.0936...	0.8372...	Q9BY7...
20				G prote...	+	CST;H...		0.6066...	4.5211...	69.94	68.208	61.003	0.56065	126	1.21124	0.3348...	Q9BY7...
21				AMP-a...	+	CST;H...		0.7310...	9.6356...	142.02	128.02	142.02	0.27257	26	0.39826	0.6103...	Q9BRL...
22				DNA d...	+	HTP;L...		0.6957...	1.3632...	102.1	97.663	57.053	-0.51608	380	0.3328...	0.6834...	Q99638

106 items

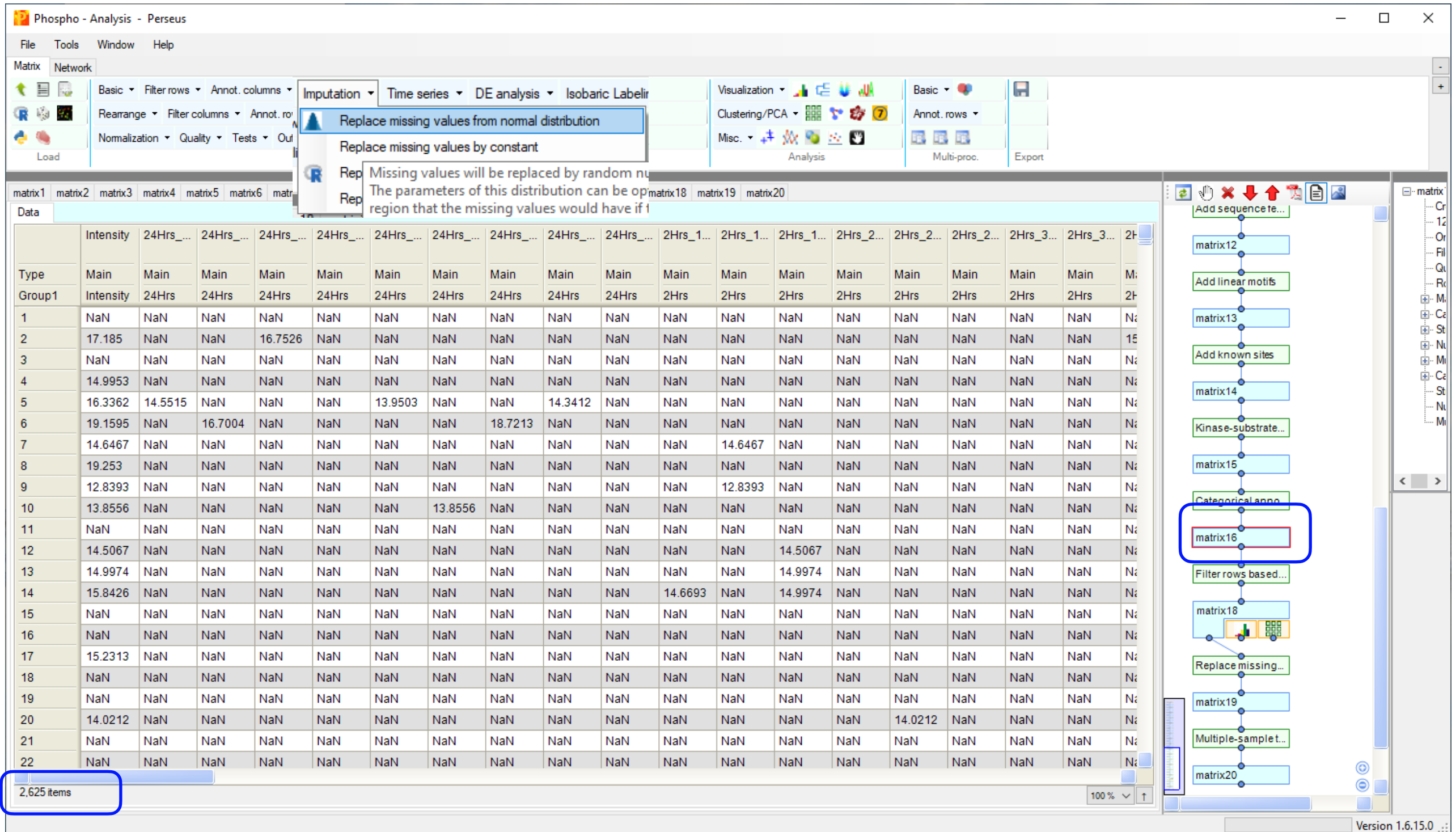
100 %

Version 1.6.15.0

matrix12 matrix13 matrix14 matrix15 matrix16 matrix18 matrix19 matrix20

Add sequence te...
matrix12
Add linear motifs
matrix13
Add known sites
matrix14
Kinase-substrate...
matrix15
Categorical anno...
matrix16
Filter rows based...
matrix18
Replace missing...
matrix19
Multiple-sampl...
matrix20

Data Quality - Fill in missing values : **Imputation > Replace missing values from normal distribution**



The screenshot displays the Perseus software interface. The 'Imputation' menu is open, showing the option 'Replace missing values from normal distribution'. The main data matrix is visible, showing 22 rows and 20 columns. The first column is labeled 'Intensity' and the subsequent columns are labeled '24Hrs_...' and '2Hrs_...'. The matrix contains numerical values and many 'NaN' entries. The status bar at the bottom left indicates '2,625 items'.

Imputation Menu Options:

- Replace missing values from normal distribution
- Replace missing values by constant
- Missing values will be replaced by random numbers. The parameters of this distribution can be optimized for the region that the missing values would have if it were not missing.

Data Matrix (Sample Rows):

Type	Intensity	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	2Hrs_1...	2Hrs_1...	2Hrs_1...	2Hrs_2...	2Hrs_2...	2Hrs_2...	2Hrs_3...	2Hrs_3...	2Hrs_4...
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	17.185	NaN	NaN	16.7526	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	15
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	14.9953	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	16.3362	14.5515	NaN	NaN	NaN	13.9503	NaN	NaN	14.3412	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	19.1595	NaN	16.7004	NaN	NaN	NaN	NaN	18.7213	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7	14.6467	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.6467	NaN	NaN	NaN	NaN	NaN	NaN	NaN
8	19.253	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	12.8393	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	12.8393	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10	13.8556	NaN	NaN	NaN	NaN	NaN	13.8556	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12	14.5067	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.5067	NaN	NaN	NaN	NaN	NaN	NaN	NaN
13	14.9974	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.9974	NaN	NaN	NaN	NaN	NaN	NaN	NaN
14	15.8426	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.6693	NaN	14.9974	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
17	15.2313	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
20	14.0212	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.0212	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Workflow Diagram (Right Panel):

- Add sequence re...
- matrix12
- Add linear motifs
- matrix13
- Add known sites
- matrix14
- Kinase-substrate...
- matrix15
- Categorical anno...
- matrix16
- Filter rows based...
- matrix18
- Replace missing...
- matrix19
- Multiple-sampl...
- matrix20

Status Bar: 2,625 items

Version: 1.6.15.0

Data Quality - Normal distributed values ? : **Visualisation - Histogram**

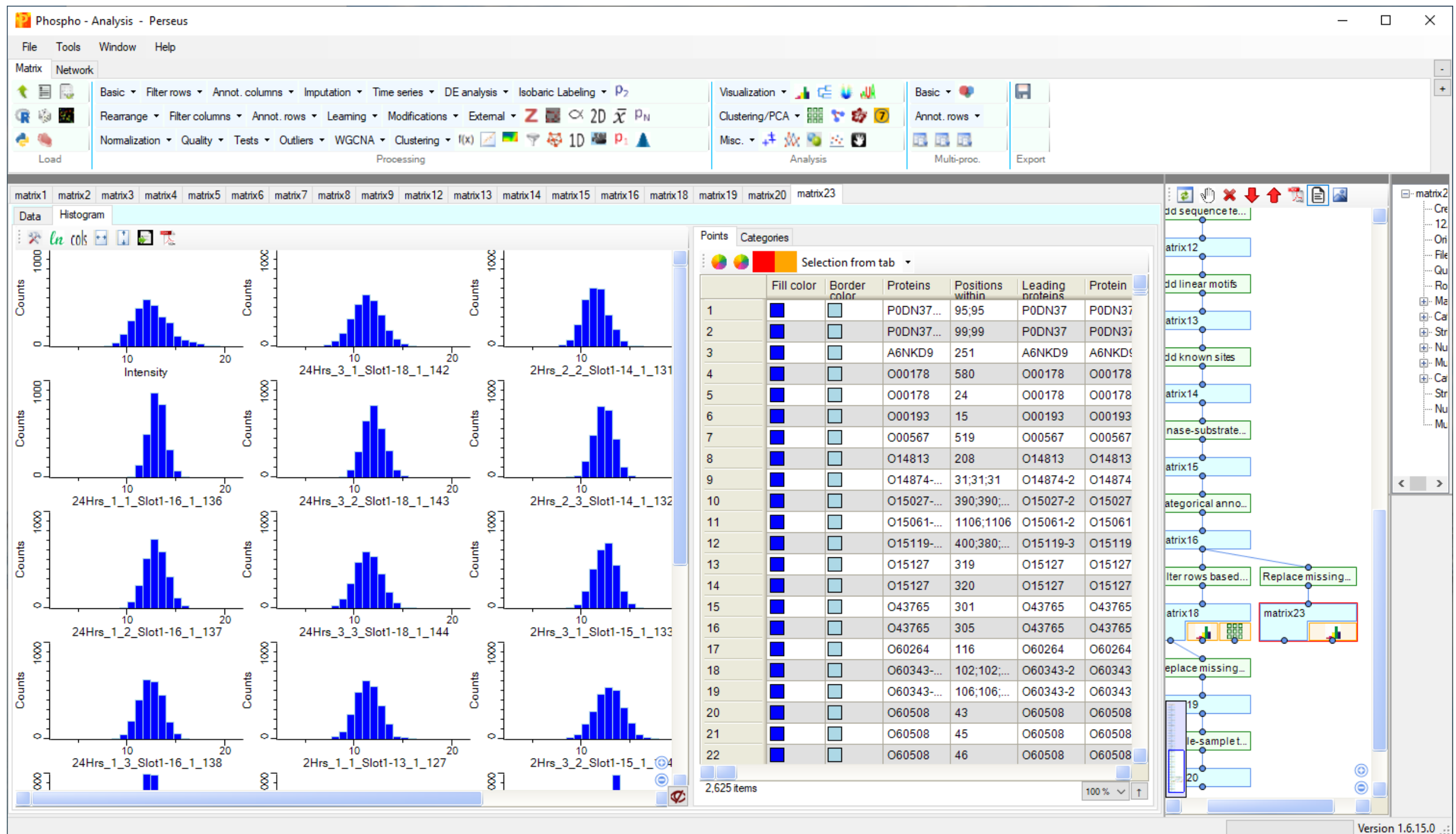
The screenshot displays the Perseus software interface. The main window shows a data matrix with columns labeled 'Intensity', '24Hrs_...', and '2Hrs_1...'. A 'Visualization' menu is open, highlighting the 'Histogram' option. The right sidebar contains a workflow diagram with steps like 'Add sequence te...', 'matrix12', 'Add linear motifs', 'matrix13', 'Add known sites', 'matrix14', 'Kinase-substrate...', 'matrix15', 'Categorical anno...', 'matrix16', 'Filter rows based...', 'Replace missing...', 'matrix18', 'matrix23', 'Replace missing...', 'matrix19', 'Multiple-sampl...', and 'matrix20'.

Type	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Mi
Group1	Intensity	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	2Hrs	2Hrs	2Hrs	2Hrs	2Hrs	2Hrs	2Hrs	2H
1	13.3647	13.8437	12.178	14.482	11.5502	11.4771	12.8985	12.3835	12.3824	15.9646	10.1868	12.636	12.5138	12.4971	10.9895	13.2194	11.8407	12.9186	13
2	17.185	12.6896	12.3739	16.7526	12.3848	10.4419	13.1452	10.5409	12.6606	9.54516	12.2183	11.4907	11.9252	13.7497	11.6987	12.703	13.2687	12.2004	15
3	11.5992	11.1253	10.9093	12.1276	12.2231	10.9818	14.4092	11.9274	13.1248	13.6075	10.8171	8.4056	10.9302	13.8803	11.7329	10.6513	13.3823	14.0117	12
4	14.9953	13.4819	13.099	13.6877	11.7429	12.4998	13.3938	10.027	11.6213	10.4981	10.6074	9.708	12.5335	12.7737	10.7986	11.6831	12.329	12.4312	12
5	16.3362	14.5515	13.6945	13.1385	13.4325	13.9503	13.5793	11.767	14.3412	12.0964	9.39797	10.6668	13.8407	13.4643	10.488	13.0886	13.1338	13.2554	13
6	19.1595	11.7927	16.7004	12.2251	11.2502	11.6518	12.3782	18.7213	11.4796	12.6681	10.6459	10.0707	10.8312	13.9526	9.89737	12.8857	13.1683	17.1702	12
7	14.6467	14.6108	13.4513	11.8726	12.2796	12.4787	11.999	8.15486	10.5455	14.2451	11.8851	14.6467	12.4259	11.7363	11.9291	10.9548	12.1763	13.0053	14
8	19.253	13.9454	13.3582	13.2217	12.0033	9.83273	12.7425	12.7105	12.3687	11.3307	11.0211	12.2608	11.6375	13.6697	10.7533	11.4656	13.0167	12.8278	12
9	12.8393	12.0907	11.6965	11.6007	12.3542	12.761	12.168	11.3266	12.5945	9.58097	14.1624	12.8393	11.873	11.7728	10.7149	12.3561	14.7253	14.4285	13
10	13.8556	14.0111	14.7812	13.3721	10.9781	10.901	13.8556	13.1143	10.7214	11.1531	12.1255	11.165	12.5386	13.5965	12.931	12.1652	10.98	12.0417	13
11	11.4319	13.9621	12.8074	12.2766	12.68	12.4253	13.9683	11.5081	12.8741	12.8463	8.67526	13.9086	12.1131	11.7949	12.6112	10.6515	12.3532	12.9516	12
12	14.5067	13.6579	13.0994	9.7121	13.2295	15.1016	12.3098	11.3758	12.5238	12.225	11.0836	12.0366	14.5067	13.973	10.583	11.781	11.2393	13.5659	12
13	14.9974	11.7546	11.9	14.1259	11.8035	11.4857	13.5251	13.7541	10.3012	13.7539	9.87544	13.388	14.9974	13.4927	10.5421	13.0525	14.5743	13.4238	14
14	15.8426	12.1121	13.264	10.3956	13.7324	12.8566	12.4854	10.5511	11.4453	12.1425	14.6693	9.90486	14.9974	13.3236	11.0139	10.1887	11.5898	10.1022	14
15	12.2337	12.5745	14.4194	12.3722	12.0187	12.5994	13.5733	8.11496	12.7258	10.476	13.5561	10.8617	11.1085	13.0176	13.1041	11.8794	12.5619	11.6096	15
16	14.0464	14.4539	12.714	11.2255	11.594	10.9785	13.8133	10.8904	11.2313	14.4483	11.7069	11.4795	8.75025	13.3244	10.2302	12.3384	11.5439	10.591	13
17	15.2313	12.4483	14.0806	12.5214	11.2906	13.2147	13.749	9.52596	10.0648	9.90714	10.8063	10.1308	13.201	13.6463	11.9169	13.0861	12.9071	11.9359	13
18	10.4866	12.5946	12.5789	10.8391	11.6916	11.19	13.1749	11.1597	12.1417	12.5609	12.5626	10.6238	11.4325	13.463	11.774	14.09	11.474	11.8924	13
19	12.2437	13.2192	12.1395	9.59071	12.5494	8.76201	11.5109	10.4165	10.3822	11.3295	11.5924	10.3826	11.943	11.9539	11.5704	13.6836	12.2566	9.57355	13
20	14.0212	12.6428	13.8134	15.6171	11.9125	13.4816	12.6547	10.6436	12.355	11.2993	10.9023	11.9022	12.1047	10.8355	14.0212	12.6018	12.967	12.8085	13
21	12.2366	12.89	11.2275	12.1009	11.6629	12.7105	13.0553	11.5279	13.8888	11.6796	12.5579	12.5361	13.4223	11.415	12.1534	13.0929	13.1143	13.9347	13
22	13.0614	10.6836	11.7201	9.55783	13.2711	11.815	12.2729	10.4841	10.3131	9.57104	10.2125	13.2102	11.3967	12.4033	11.7988	11.9984	10.0194	12.5312	14

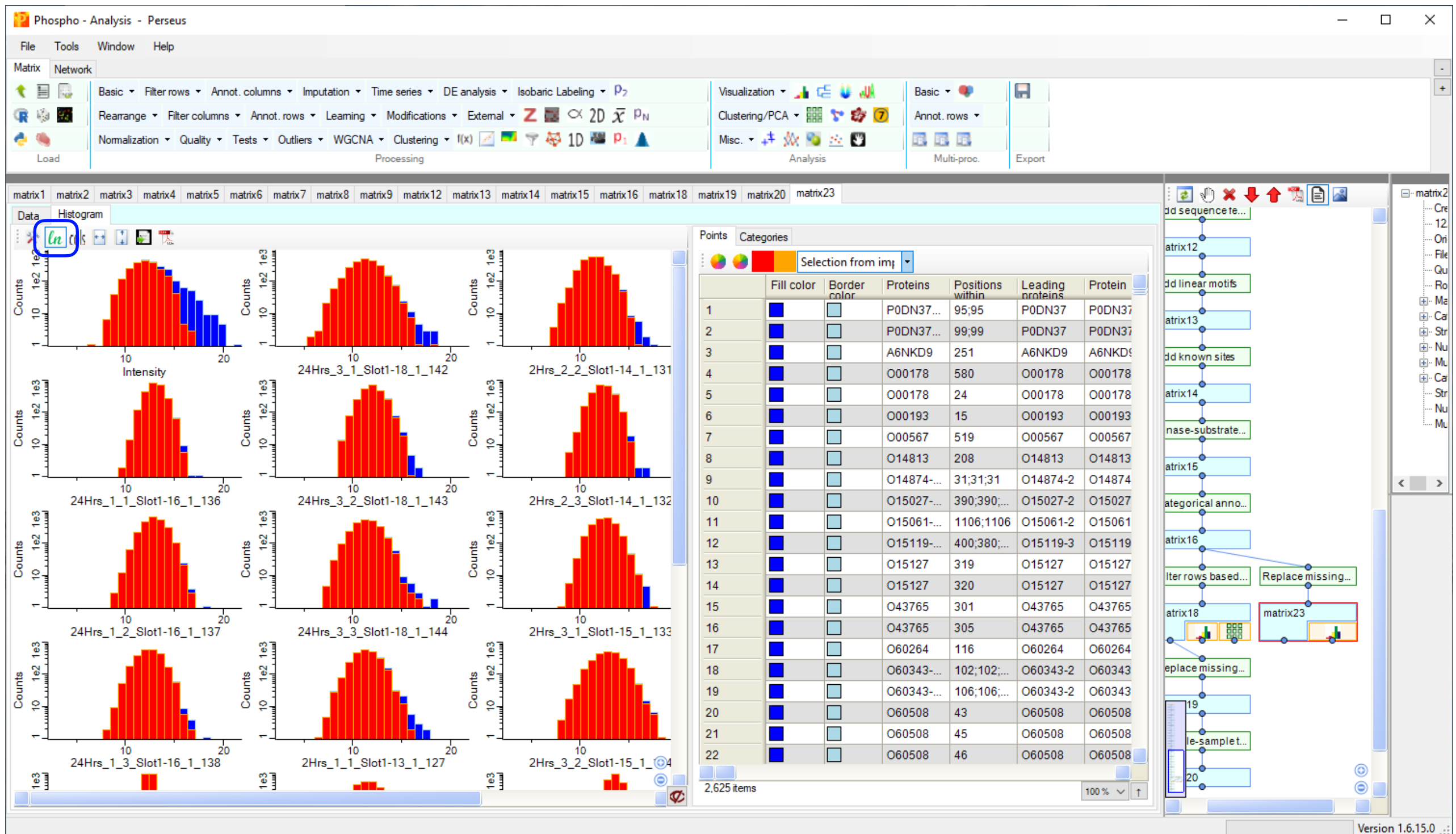
2,625 items

Version 1.6.15.0

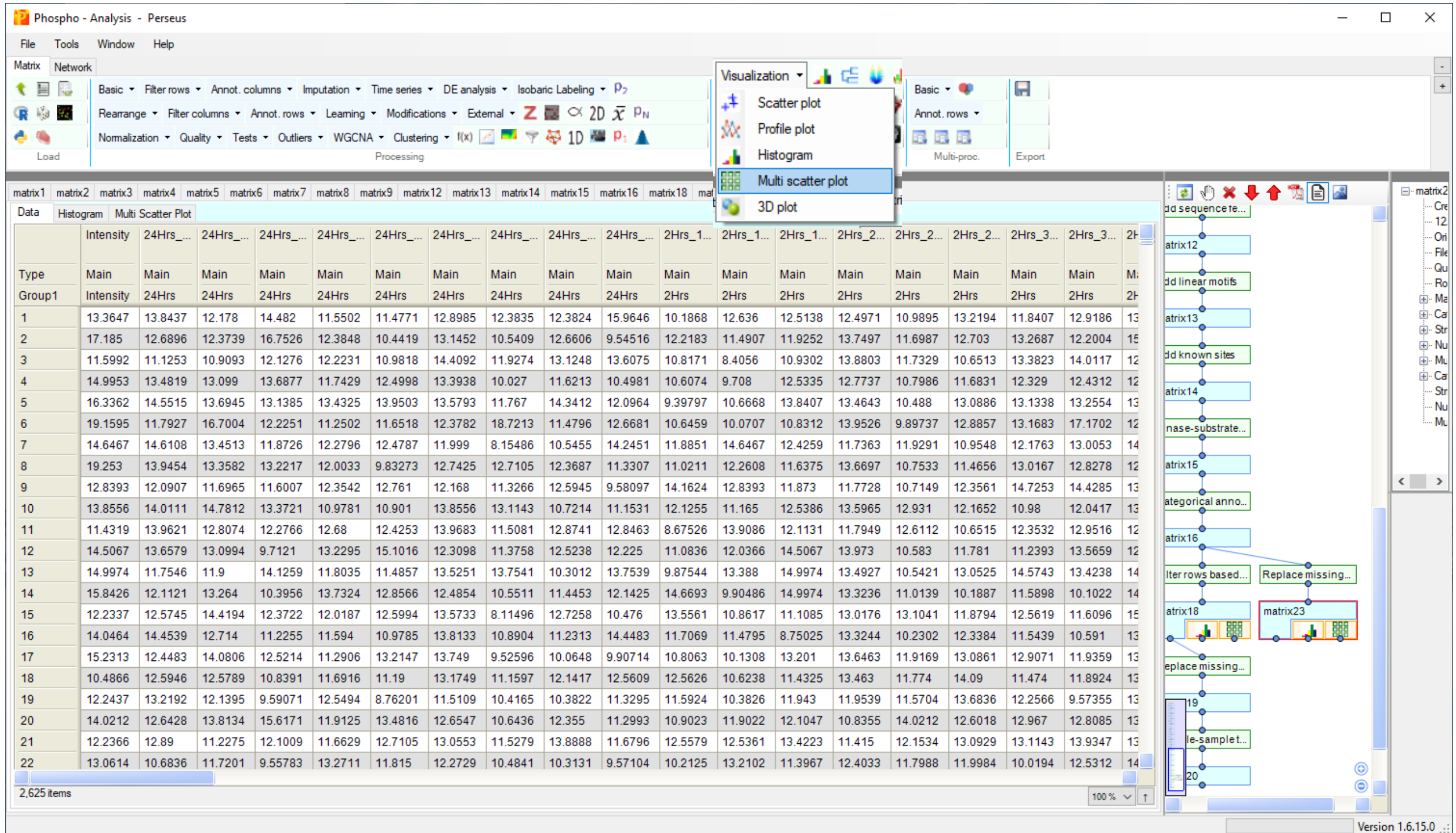
Data Quality - Normal distributed values ? : **Visualisation - Histogram**



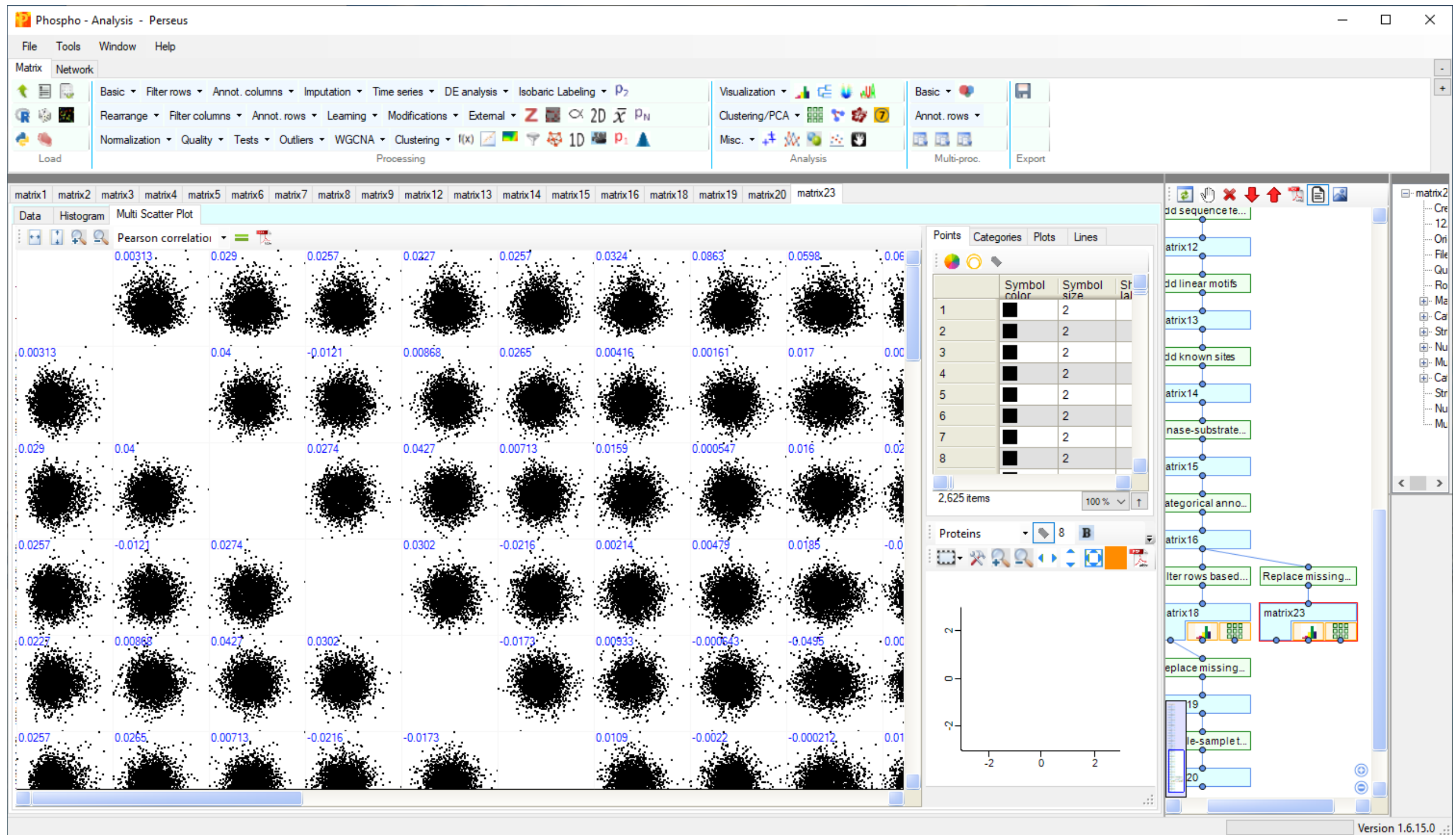
Data Quality - Normal distributed values ? : **Visualisation - Histogram**



Data Quality: **Visualisation - Multi scatter plot**



Data Quality: **Visualisation - Multi scatter plot**



Data Quality - Too many missing values = no successful repeats on phosphorylation sites

Phospho - Analysis - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series DE analysis Isobaric Labeling P₂

Rearrange Filter columns Annot. rows Learning Modifications External Z 2D X P_N

Normalization Quality Tests Outliers WGCNA Clustering f(x) 1D P₁

Visualization Clustering/PCA Misc. Basic Annot. rows

Load Processing Analysis Multi-proc. Export

matrix1 matrix2 matrix3 matrix4 matrix5 matrix6 matrix7 matrix8 matrix9 matrix12 matrix13 matrix14 matrix15 matrix16 matrix18 matrix19 matrix20

Data

Type	Intensity	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	24Hrs_...	2Hrs_1...	2Hrs_1...	2Hrs_1...	2Hrs_2...	2Hrs_2...	2Hrs_2...	2Hrs_3...	2Hrs_3...	2Hrs_3...
Group1	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main
	Intensity	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	24Hrs	2Hrs	2Hrs	2Hrs	2Hrs	2Hrs	2Hrs	2Hrs	2Hrs	2Hrs
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	17.185	NaN	NaN	16.7526	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	15
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	14.9953	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5	16.3362	14.5515	NaN	NaN	NaN	13.9503	NaN	NaN	14.3412	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	19.1595	NaN	16.7004	NaN	NaN	NaN	NaN	18.7213	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7	14.6467	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.6467	NaN	NaN	NaN	NaN	NaN	NaN	NaN
8	19.253	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	12.8393	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	12.8393	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10	13.8556	NaN	NaN	NaN	NaN	NaN	13.8556	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12	14.5067	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.5067	NaN	NaN	NaN	NaN	NaN	NaN	NaN
13	14.9974	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.9974	NaN	NaN	NaN	NaN	NaN	NaN	NaN
14	15.8426	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.6693	NaN	14.9974	NaN	NaN	NaN	NaN	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
16	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
17	15.2313	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
18	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
19	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
20	14.0212	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	14.0212	NaN	NaN	NaN	NaN	NaN
21	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
22	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

2,625 items

100 %

Version 1.6.15.0

matrix12 matrix13 matrix14 matrix15 matrix16 matrix18 matrix19 matrix20

Add sequence re...

matrix12

Add linear motifs

matrix13

Add known sites

matrix14

Kinase-substrate...

matrix15

Categorical anno...

matrix16

Filter rows based...

matrix18

Replace missing...

matrix19

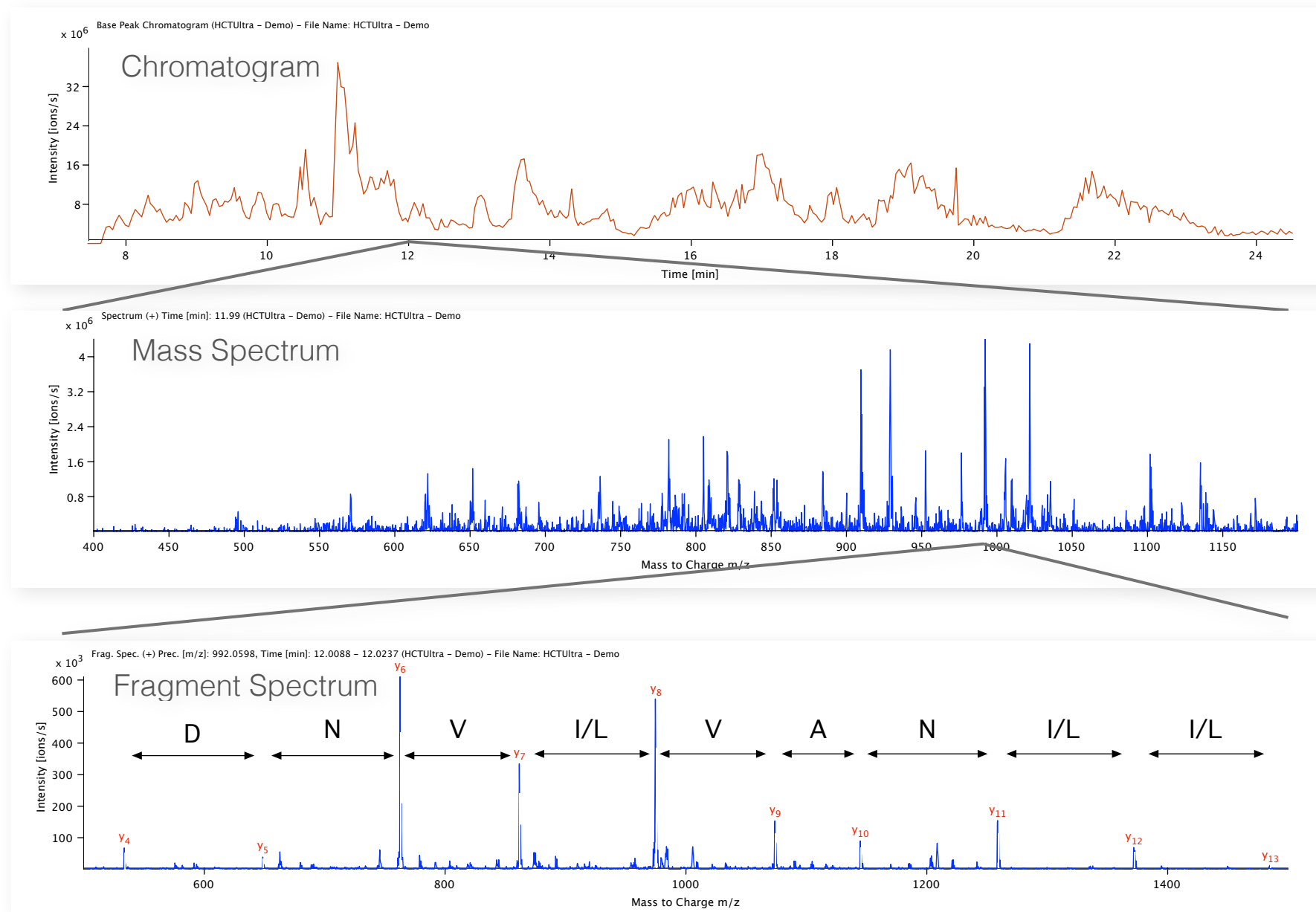
Multiple-sampl...

matrix20

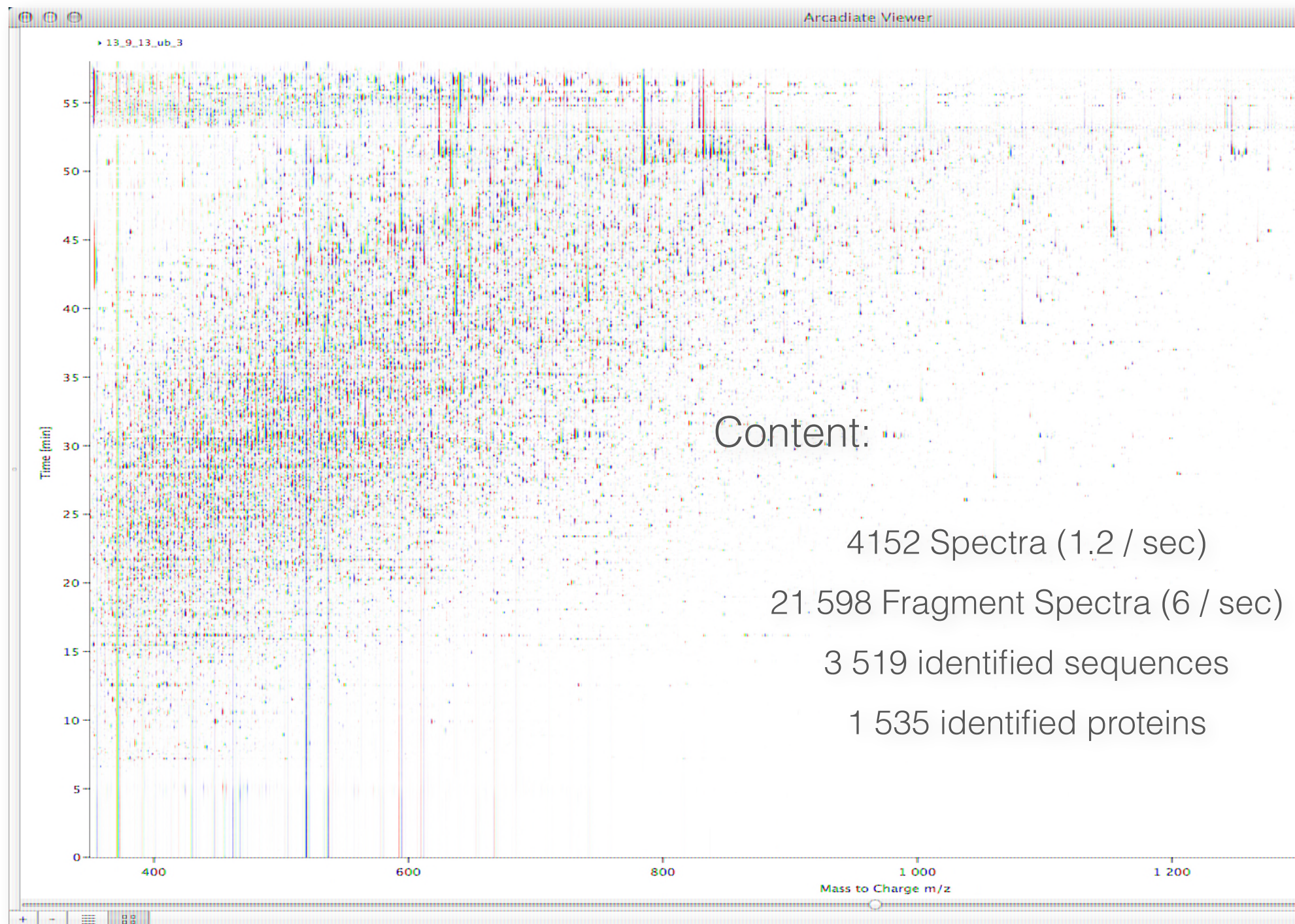
Special Mass Spectrometric Acquisition Regimes

The Problem of Missing Repeats

Origin: Data Dependent Acquisition and Under-Sampling

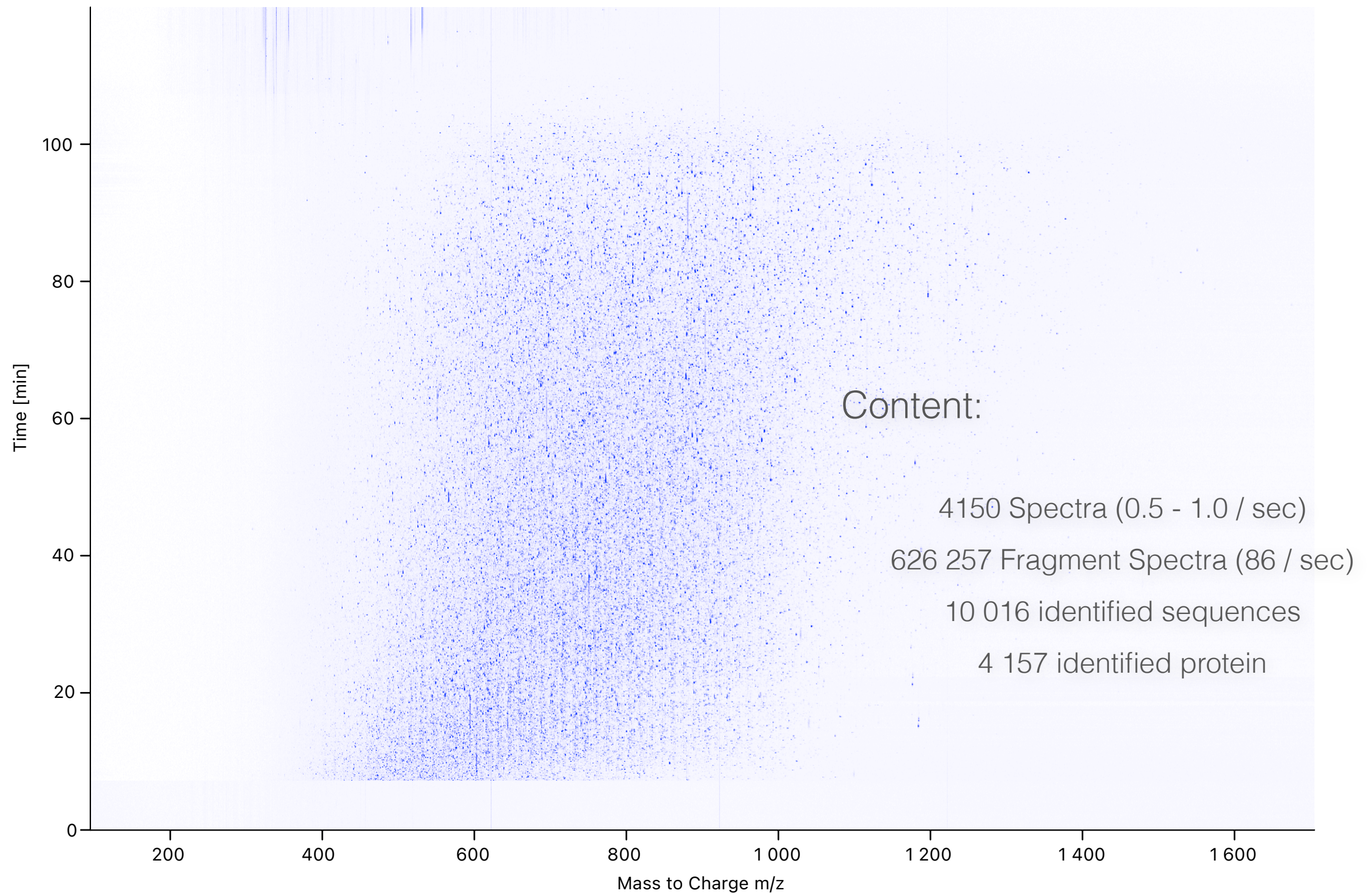


QExactive Standard Single Run

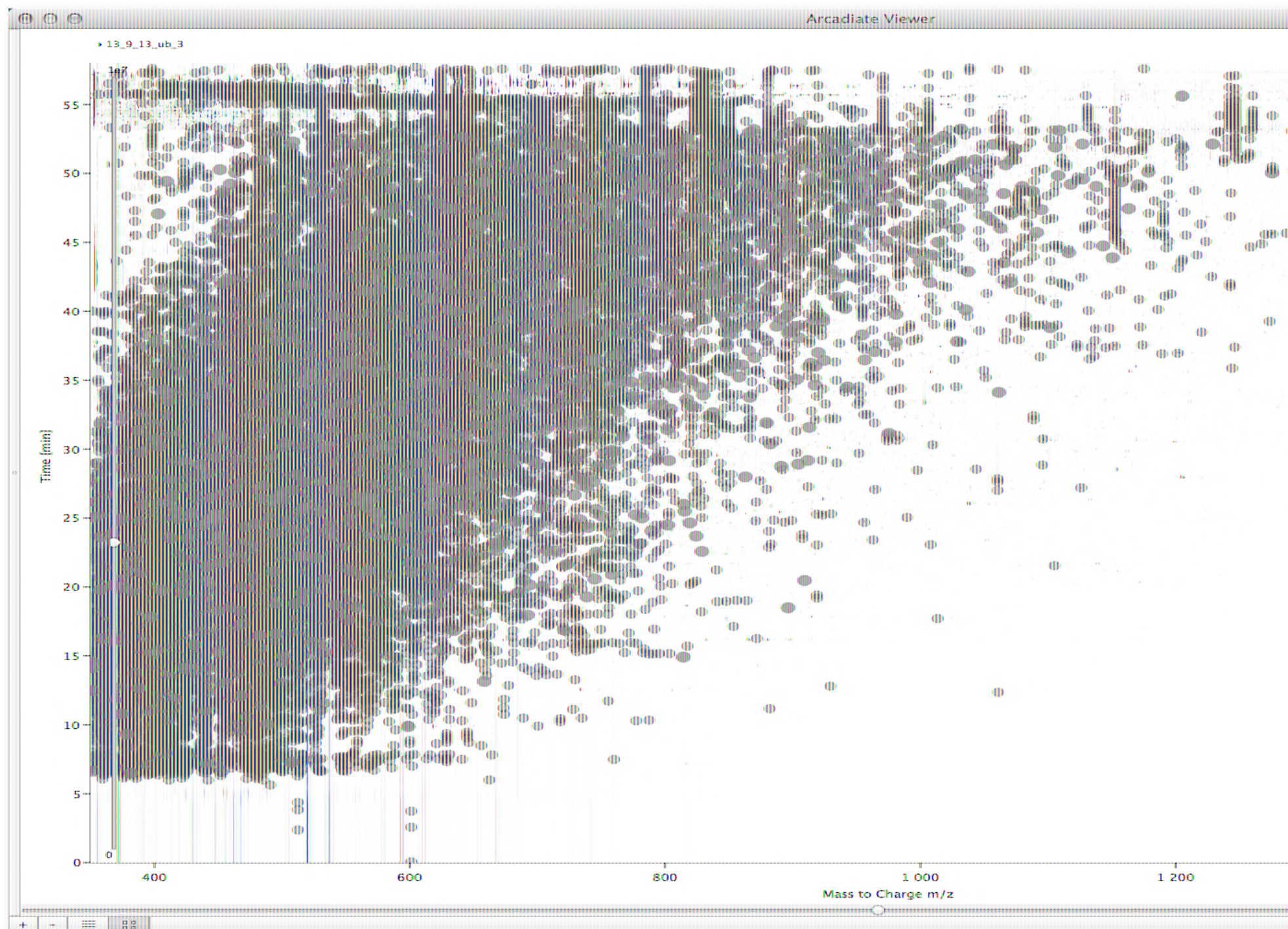


timsTOF Standard Single Run

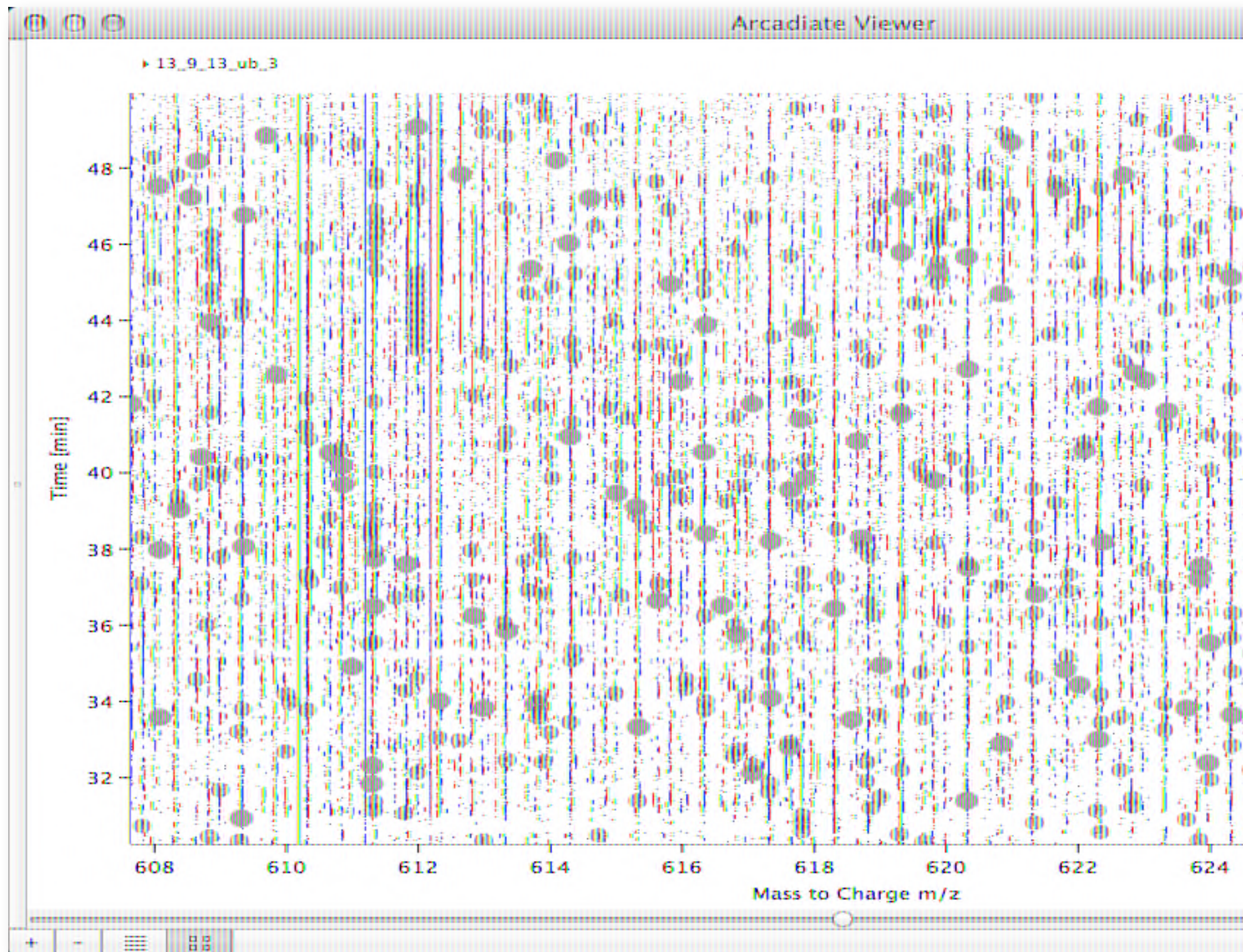
▸ timsTof Large



QExactive Standard Single Run



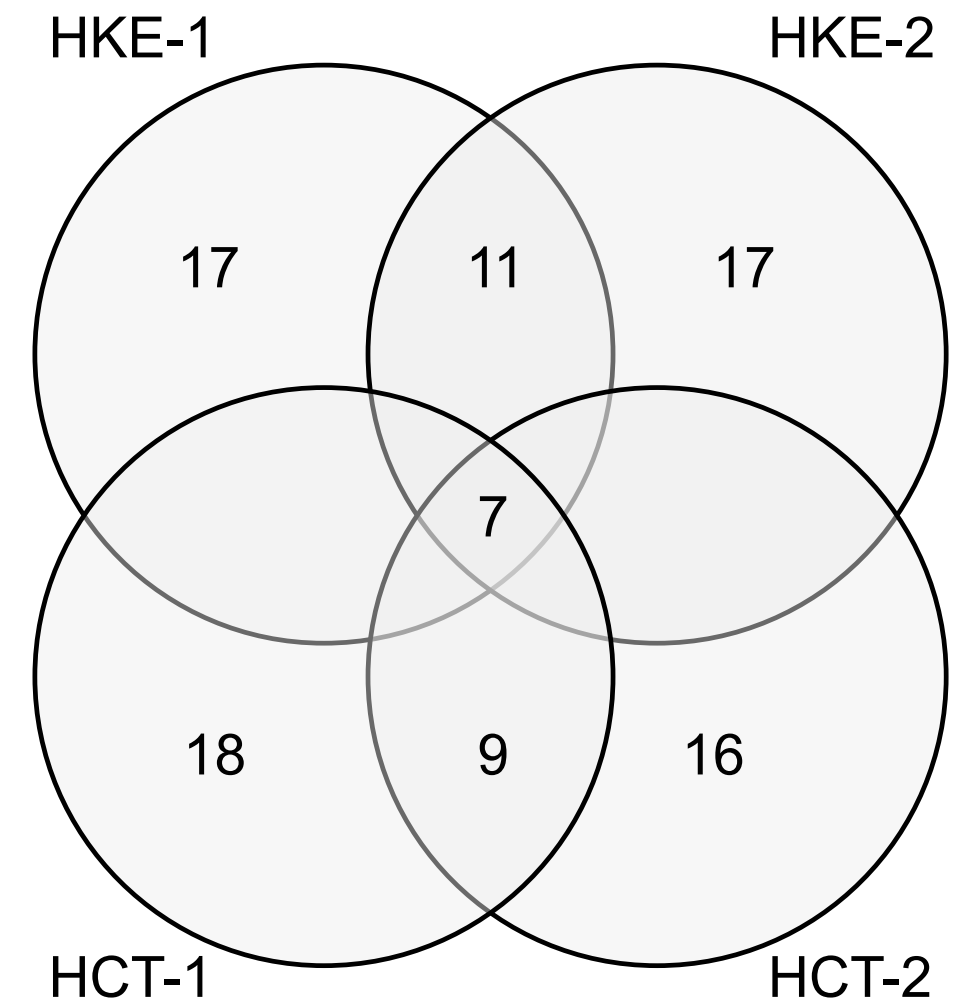
QExactive Standard Single Run



Reproducibility

Re-Sequencing Internal Standard Peptides

Sample	<i>Number of Protein Derived Peptides Identified (SILAC labelled)</i>	<i>Number of Synthetic Peptides Identified of 52 Previously Sequenced (unlabelled)</i>
HKE - 1	20 519	17
HKE - 2	19 358	17
Common	12 548	11
HCT - 1	20 045	18
HCT - 2	16 229	16
Common	10 353	9
Common	6 826	7



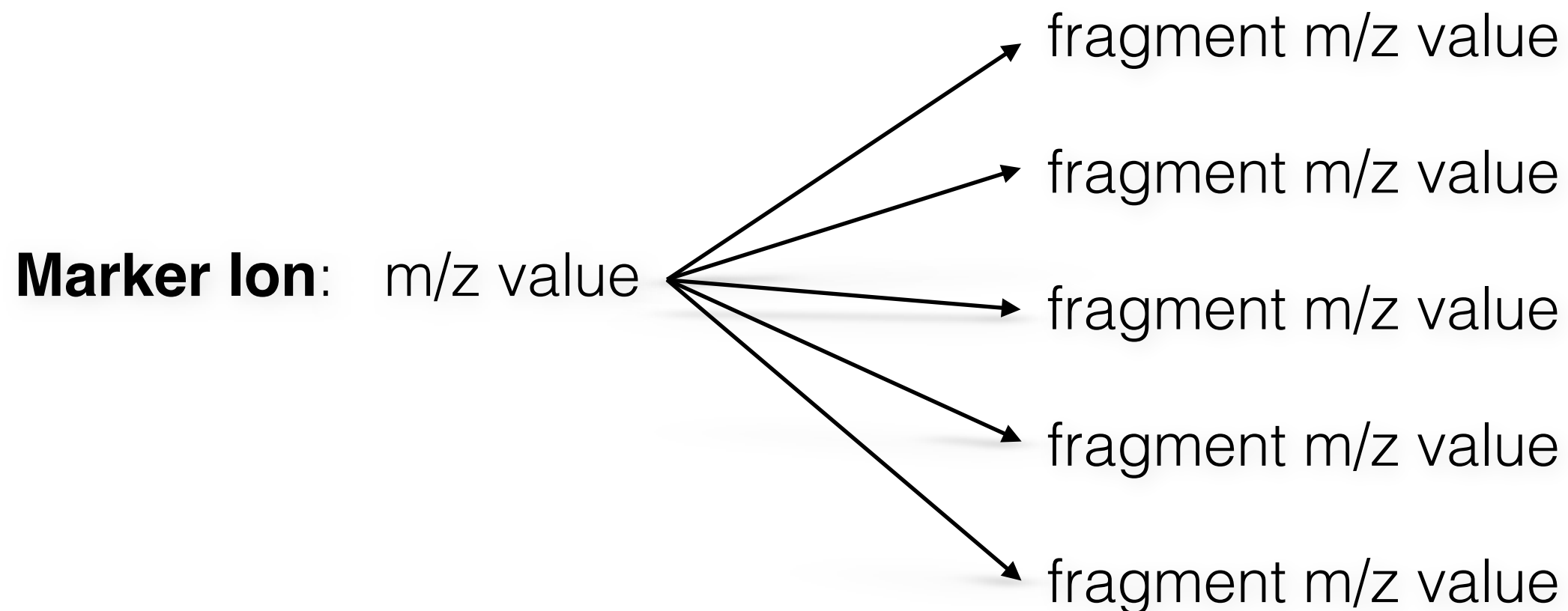
Multiple Reaction Monitoring (MRM) Scans

- **Purpose:** Detection and quantification of well known molecules on very low levels - even below the spectral noise level
- **Application:** Widely used in the pharmaceutical industry to detect drugs and their metabolites.
Used in forensic pathology and for athletes screening
- **Operation Principle:** The mass spectrometer limits itself to the detection of specific fragment ions generated by molecules with specific masses

Multiple Reaction Monitoring (MRM) Scans

Data Analysis: Marker Ions based

Question: Has a specific marker ion been seen ?

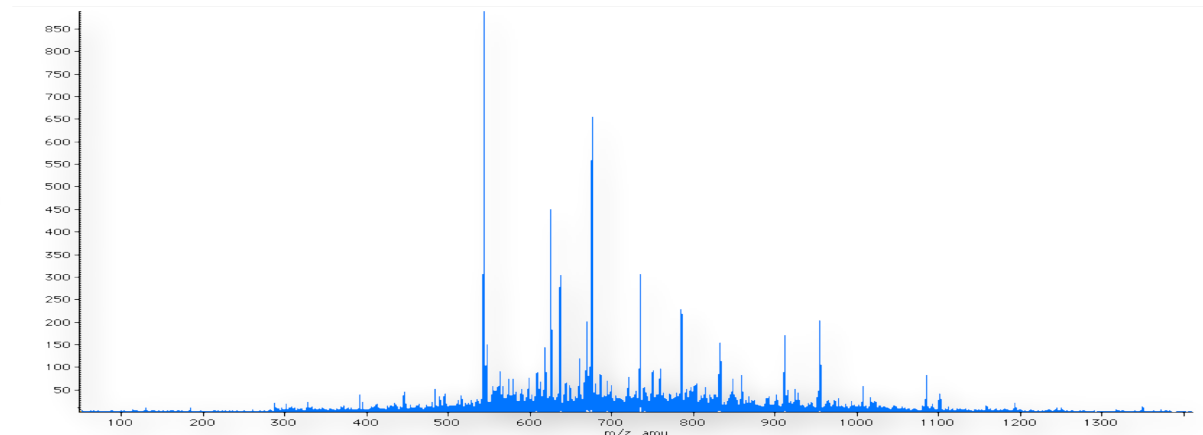


Global Proteomic Analysis

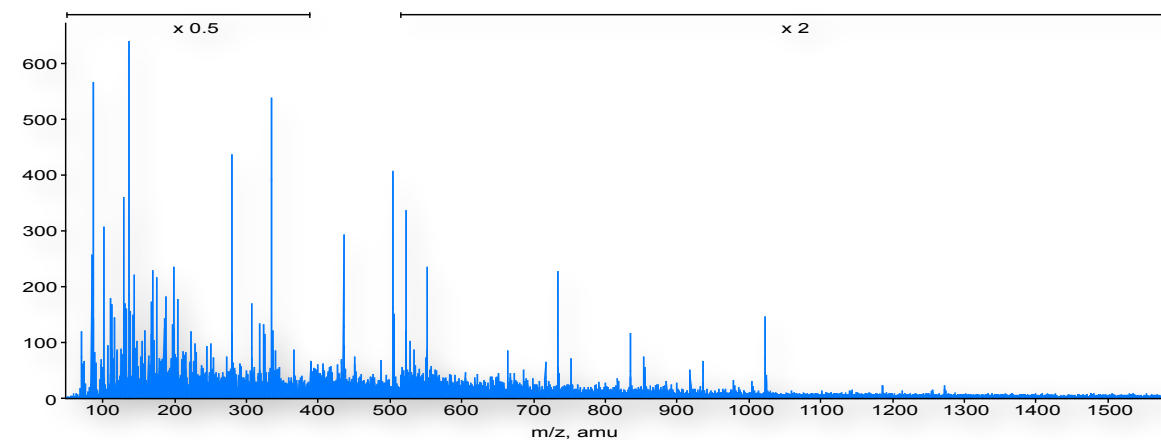
Approach: Data Independent Acquisition

MSE^E - Acquisition or All Ion Fragmentation

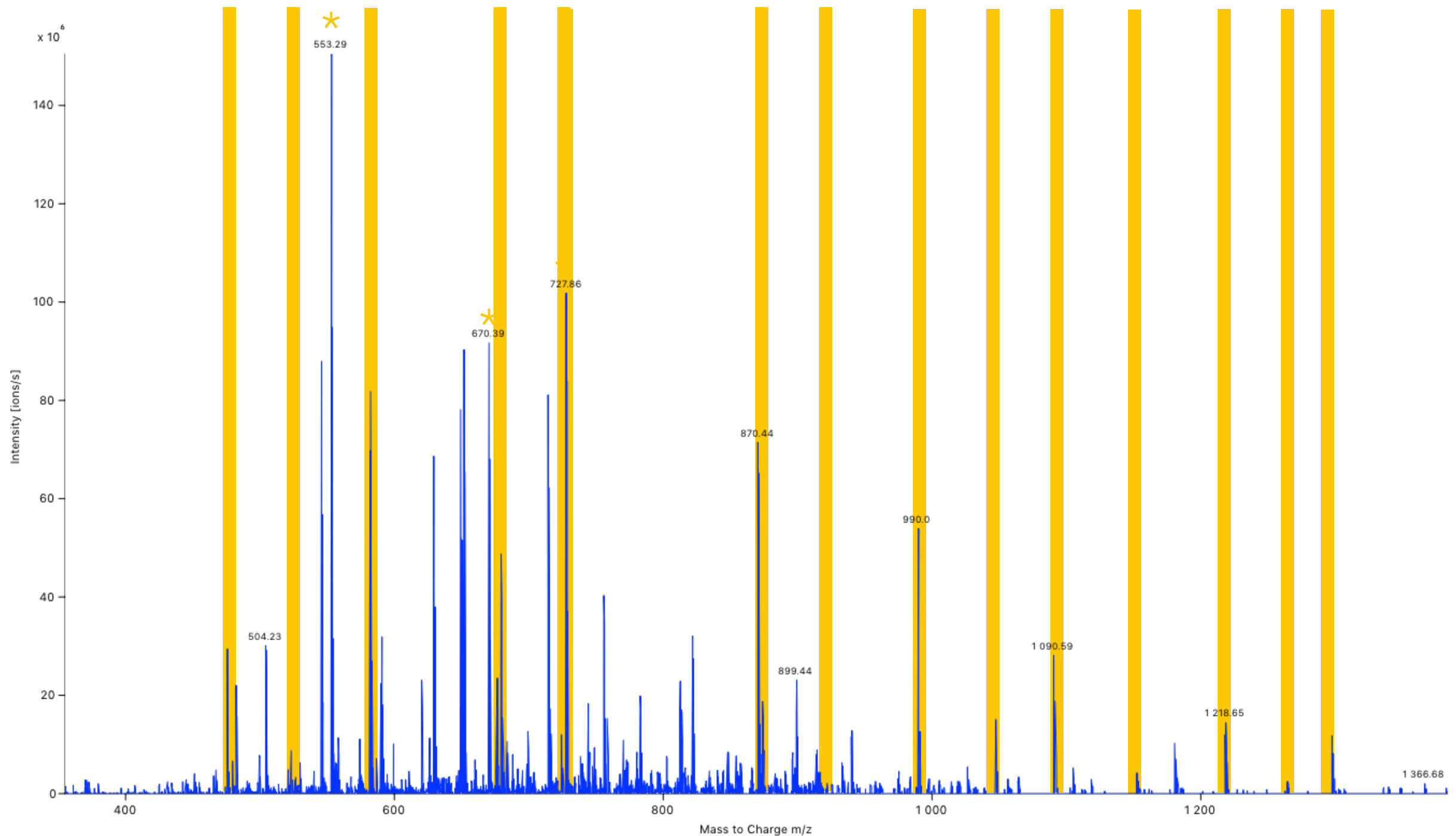
Mass Spectrum



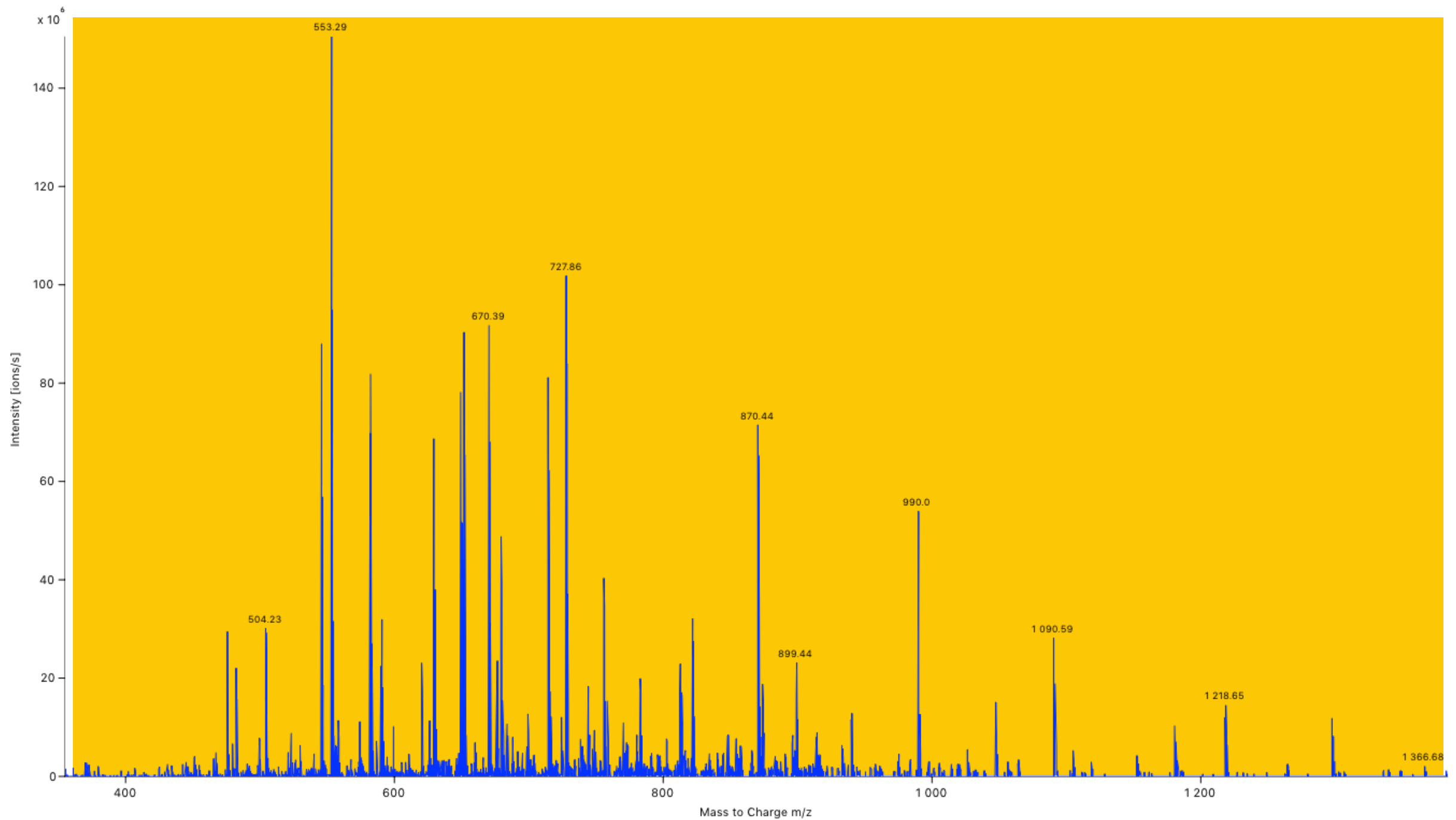
Fragment Spectrum



Data Independent Acquisition = MS^E = SWATH Acquisition

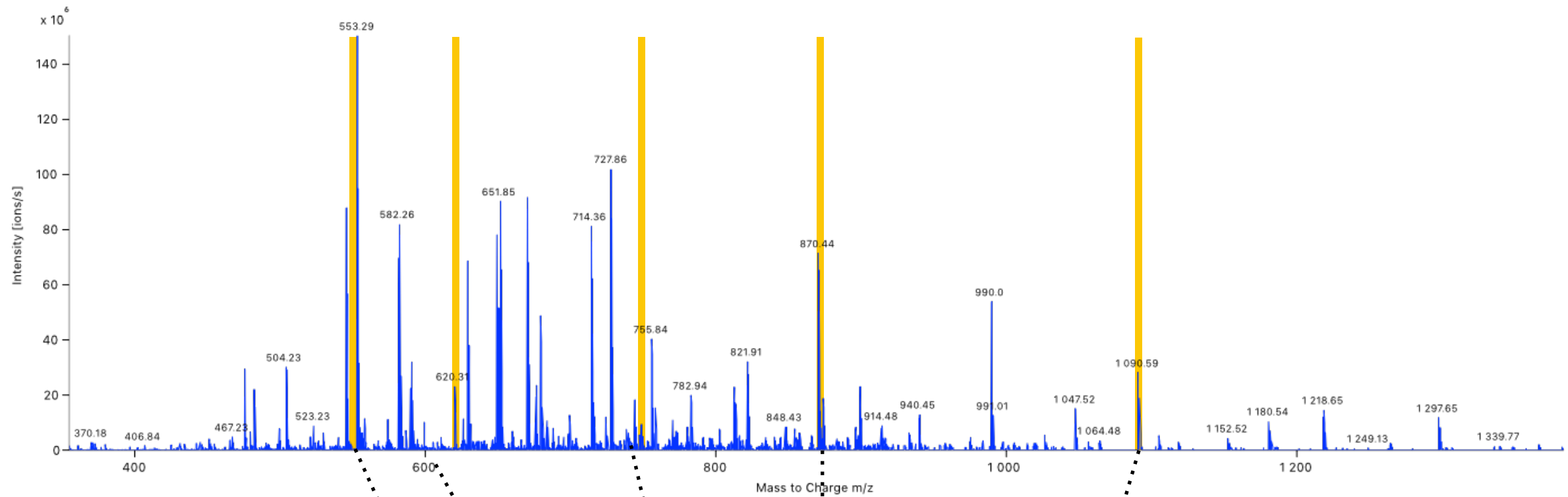


Data Independent Acquisition = MS^E = SWATH Acquisition

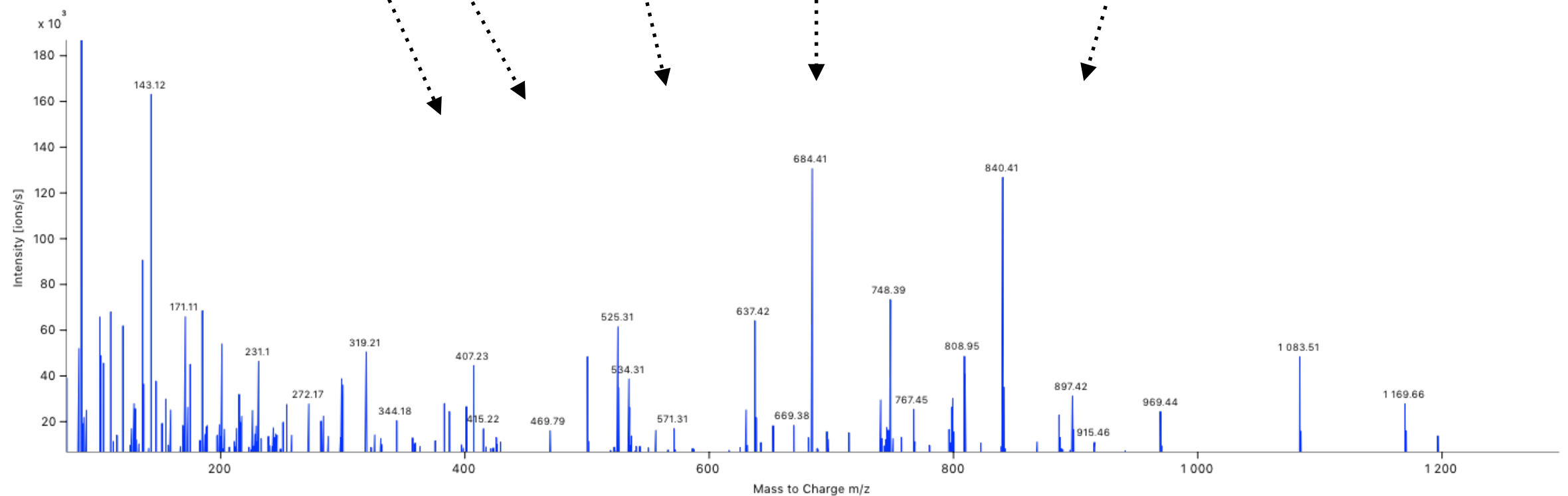


Data Independent Acquisition = MS^E = SWATH Acquisition

Intact Ion Spectrum

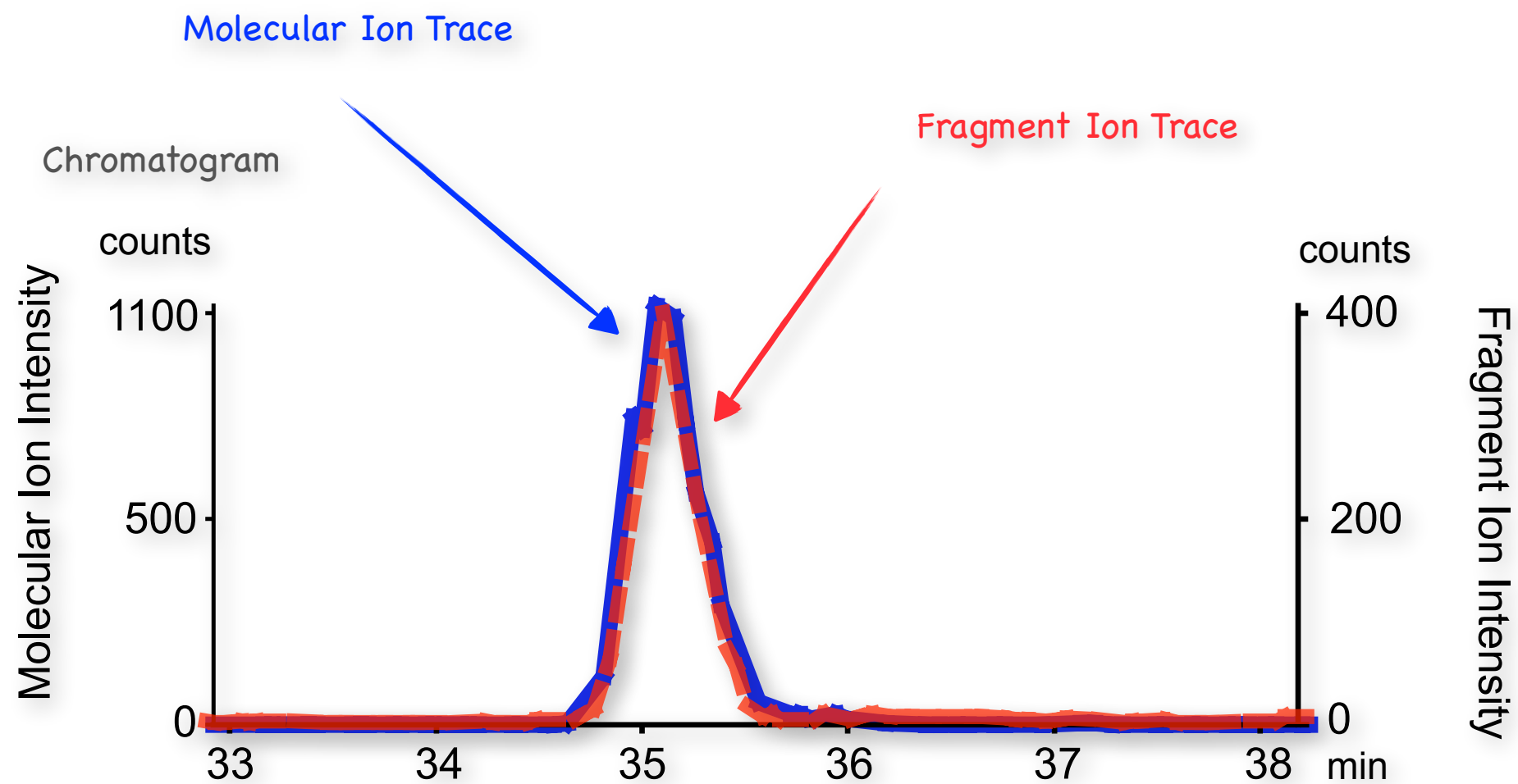


Fragment Spectrum



Data Independent Acquisition = MS^E = SWATH Acquisition

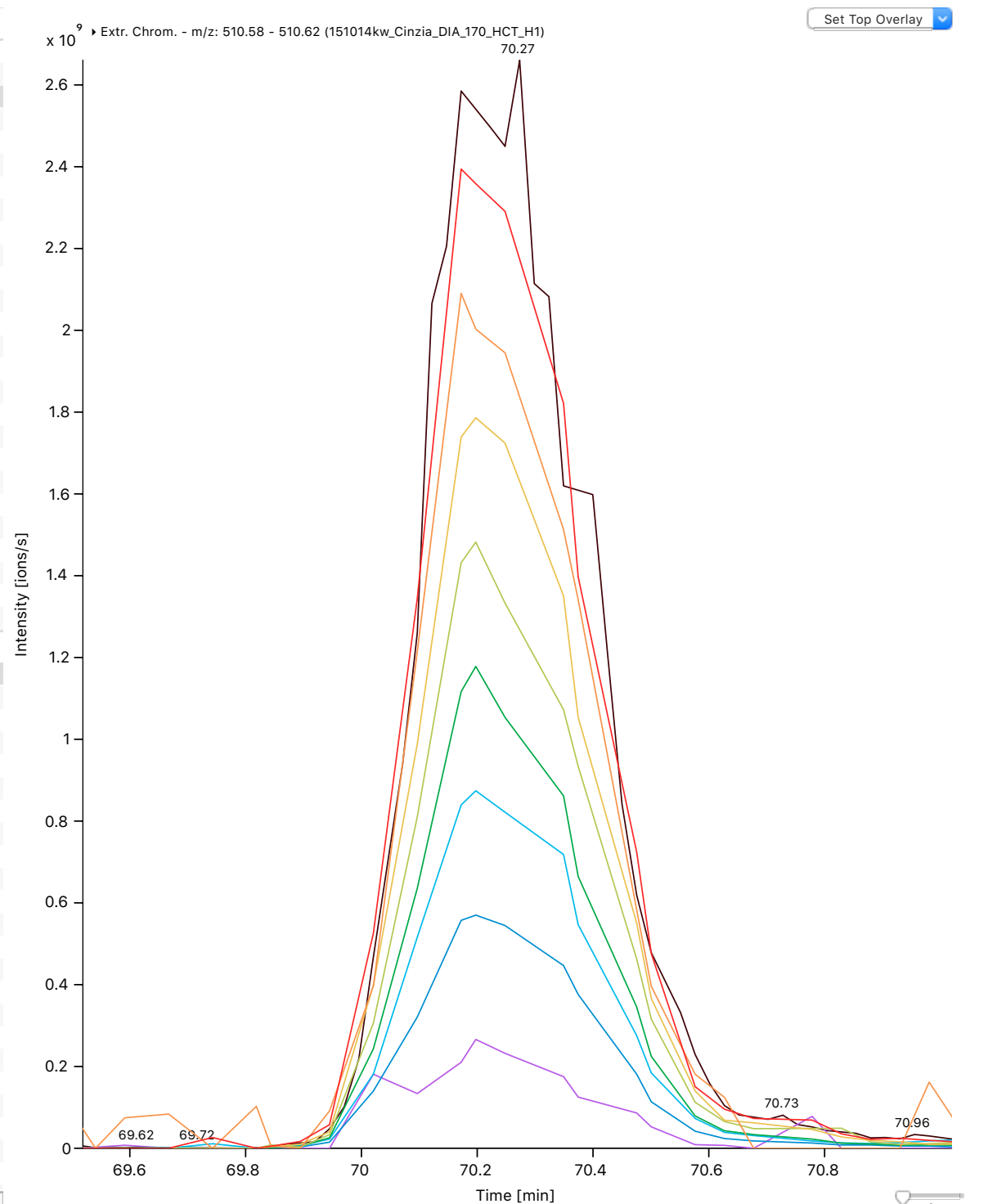
Fragment - Precursor association by time profile correlation



Data Independent Acquisition = MSE^{E} = SWATH Acquisition

Name	m/z	Retention Time	No o...	Retention Time (151014k...	151014kw_Cinzia_DIA_170_HCT...
ITDFGLARLLEGDEK	559....	138.5958	14	140.418	492 550
HTFSGVASVESSSGEAFHVGK	707....	162.3692	11	156.916	8 626
KLEELELDEQQR	510....	68.5794	8	70.269	4 241 069
LFSANDVENIFSR	756....	144.9625	8	143.333	3 128 798
RISGASELGPFS DPR	530....	82.2804	8	77.72	246 338
DAFLGSFLYEYSR	784....	167.0433	7	160.882	197 208
IGDFGLATVK	510....	95.0955	7	90.615	92 851
VLGLLGALDPYK	630....	149.8718	7	148.504	45 351
AQEALDFYGEVR	699....	146.1449	6	140.732	63 576
EFDPGLPLPPGW EK	791....	137.6905	6	137.325	951 896
RESGPGIAPGP EPHGLTNKK	511.2...	42.9602	6	46.234	17 622
EIDWVLL EQK	636....	129.876	5	123.247	99 887
GPEQTADDADDAAGHK	533....	40.3745	5	35.445	296 882
GPQNAFFLVK	560....	99.0171	5	98.041	36 308
GVPQGQVDFYAR	604....	124.2611	5	118.885	18 415
IWHHTFYNELR	758....	73.2697	5	73.359	292 746
LGLFHSGTAK	515.79	56.4629	5	50.375	15 932
LITAQQILK	514....	94.5173	5	94.303	180 906
TYQGSYGFR	540....	67.3737	5	63.912	39 902
VPVSAPSSSDPWTGR	771....	84.6305	5	81.101	4 675 309
GESIEPLDPSEK	650....	70.969	4	68.426	4 755 789
HLVDEPQNLIK	653....	78.0865	4	76.691	20 180 242
HTLNQIDSVK	577....	46.0258	4	39.884	26 961
QASTDAGTAGALTPQHVR	594....	53.5777	4	52.687	1 184 976
SLAIELLDK	501....	108.8008	4	107.983	467 113
EAEAAIYHLQLFEELR	644....	139.3877	7	140.418	7 655

Fragment Ion m/z	151014kw_Cinzia_DIA_170_HCT...	Detection (151014kw_Cinzia_DIA_170_HCT_...
431.2356	4 532 022	●
500.272	2 577 101	●
560.2781	6 217 827	●
613.3561	3 385 008	●
675.3051	9 540 321	●
742.3987	4 252 526	●
788.3891	2 870 498	●
917.4317	553 246	●



Data Independent Acquisition = MS^E = SWATH Acquisition

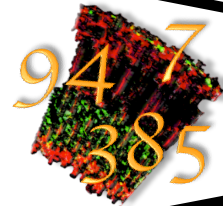
Peptide Detection Reproducibility in Proteomic Samples

Name	Detection (160... ▾	Detect...	Detection (16022...	Detection (...	Detection...	Detection...	Detection (...	Dete...	m/z	Retention Ti...	No of
KIILSDEGK	●	●	●	●	●	●	●	●	501.7951	19.3818	5
KIILSDEGK	●	●	●	●	●	●	●	●	501.7951	19.3818	5
FQVDNNNR	●	●	●	●	●	●	●	●	503.7387	15.6921	3
AVPLAGFGYGLPISR	●	●	●	●	●	●	●	●	506.6209	55.5616	4
VGGHAAEYGAEALER	●	●	●	●	●	●	●	●	510.5824	28.3806	7
KLEELDEQQR	●	●	●	●	●	●	●	●	510.598	30.0462	7
ATVVYQGER	●	●	●	●	●	●	●	●	511.7668	20.3375	4
LGFLHSGTAK	●	●	●	●	●	●	●	●	515.789	24.3068	4
LSSPATLNSR	●	●	●	●	●	●	●	●	523.2868	24.3341	6
DTGILDSIGR	●	●	●	●	●	●	●	●	523.7773	40.9677	4
GLQDEDEGYR	●	●	●	●	●	●	●	●	526.7357	20.3515	4
TYQGSYGFRR	●	●	●	●	●	●	●	●	539.7504	26.8606	4
IDQNVEELK	●	●	●	●	●	●	●	●	544.282	25.9312	4
YVLPNFEVK	●	●	●	●	●	●	●	●	554.8067	42.7296	6
AAISGENAGLVR	●	●	●	●	●	●	●	●	579.317	27.4778	5
LEEAVWRPY	●	●	●	●	●	●	●	●	581.7965	38.6712	6
ADVGRLSPEVK	●	●	●	●	●	●	●	●	585.8277	21.2501	4
TGLLLSDPDK	●	●	●	●	●	●	●	●	586.3308	44.9321	4
HTLNQIDEVK	●	●	●	●	●	●	●	●	598.8172	21.2443	6
EYTDASFTNR	●	●	●	●	●	●	●	●	602.2665	25.8031	5
GVPQGQVDFYAR	●	●	●	●	●	●	●	●	604.8063	35.8425	4
LPPNVVEESAR	●	●	●	●	●	●	●	●	605.8233	27.1724	5
SNSFNNPLGRR	●	●	●	●	●	●	●	●	610.2933	27.3168	6
NFPSPVDAAFR	●	●	●	●	●	●	●	●	610.8075	42.9589	4
QVSSHQVLAR	●	●	●	●	●	●	●	●	619.3524	26.6196	4
QVSSHQVLAR	●	●	●	●	●	●	●	●	619.3524	26.6196	4
VGQALLQGNTERR	●	●	●	●	●	●	●	●	643.3461	19.9504	5
FYNQVSTPLLR	●	●	●	●	●	●	●	●	669.3601	39.2361	5
AGNILLNTEGHAK	●	●	●	●	●	●	●	●	669.3611	24.8295	5
GANILLTDNGHVK	●	●	●	●	●	●	●	●	676.3667	27.7314	3
QGVDDAFYTLVR	●	●	●	●	●	●	●	●	692.3517	50.5798	6
SSSLEGFHNHFK	●	●	●	●	●	●	●	●	695.3305	24.4947	3
AQEALDFYGEVR	●	●	●	●	●	●	●	●	699.3383	41.5673	6
LTQEHIEALLDK	●	●	●	●	●	●	●	●	705.3835	36.751	10
LVEFSAFLEQQR	●	●	●	●	●	●	●	●	733.8875	54.7772	6
GSHQISLDNPDYQQDFFPK	●	●	●	●	●	●	●	●	746.0128	45.3207	3
QVPVESDLQEIK	●	●	●	●	●	●	●	●	749.4121	46.4353	5
FNSLNELVDYHRR	●	●	●	●	●	●	●	●	753.8666	39.2106	6
VPVSAPSSSDPWTR	●	●	●	●	●	●	●	●	771.8808	34.8465	4
IQPSGGTNINEALLR	●	●	●	●	●	●	●	●	791.9312	36.6708	5
IPEEHDLESQIRK	●	●	●	●	●	●	●	●	797.4147	25.5336	7
AEQWNVNYVETSAK	●	●	●	●	●	●	●	●	819.8893	35.2123	6
ITDFGLARLLGDEK	●	●	●	●	●	●	●	●	838.9457	56.3549	3
EAPAKLESQAGQVSR	●	●	●	●	●	●	●	●	849.9418	24.6916	5
IQNTGDYYDLYGGEK	●	●	●	●	●	●	●	●	868.3915	38.6125	7
DTVIKPLLVEPEGLEK	●	●	●	●	●	●	●	●	594.0079	47.8738	8
SYENVDSGDK	●	●	●	●	●	●	●	●	557.2383	13.9919	4
LEESYYWLTG	●	●	●	●	●	●	●	●	666.3301	19.5777	3
NHLDYRPVALLFHK	●	●	●	●	●	●	●	●	861.9745	39.9643	4
QASTDAGTAGALTPQHVR	●	●	●	●	●	●	●	●	890.9491	23.8563	8
TLTEDEIATILQSTLK	●	●	●	●	●	●	●	●	888.4867	67.6154	10
VLSLAQEQVGGSPK	●	●	●	●	●	●	●	●	771.4117	33.1987	6
TQVNTQAEQLR	●	●	●	●	●	●	●	●	644.3354	22.9372	5
ILEEAHELSEDHYK	●	●	●	●	●	●	●	●	856.9082	31.2746	10
EHAVEGDCDFQLLK	●	●	●	●	●	●	●	●	554.2603	36.4915	5
FVAHGLMILAR	●	●	●	●	●	●	●	●	614.3561	47.4626	5
AEFQDALEK	●	●	●	●	●	●	●	●	525.7588	28.9019	4

Data Independent Acquisition = MSE = SWATH Acquisition

How to analyse DIA Data ?

Arcadiate ?



- Pros:**
- It is a Mac programme 😊
 - Graphical display of chromatograms -> Data quality
- Cons:**
- It is old
 - It requires mzML/mzXML data (up to 50 GByte/run)

Skyline ?



- Pros:**
- Graphical display of chromatograms -> Data quality
 - Very powerful beyond DIA
- Cons:**
- It requires mzML/mzXML data (up to 50 GByte/run)
 - It might crash because of these very large data files

DIA-NN ?



- Pros:**
- Can read native MS Data
 - Artificial intelligence based algorithms
 - Very good memory management
 - fast
- Cons:**
- No graphical interface for MS data

Data Independent Acquisition



DIA-NN

06/10/2022 14:47:05 x +

Experiment name: 06/10/2022 14:47:05

Reset settings Clear log Save log

Input

Raw diaPASEF .d Clear list Convert to .dia

Output

☐ Use existing .quant files when available

Main output C:\Program Files\DIA-NN\v1-8\report

Temp/.dia dir

☐ Generate spectral library ☒ Quantities matrices

Output library

☐ Generate Proteo input from FASTA or spectral library

Precursor FDR (%) 1.0 Threads 6

☒ Generate PDF report Log level 1

Additional options

Spectral library

Add FASTA

Clear list

☐ Reannotate

DIA-NN exe diann.exe

Run Not started 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 0

☒ N-term M excision ☒ C carbamidomethylation

☐ Ox(M) ☐ Ac(N-term) ☐ Phospho ☐ K-GG

Peptide length range 7 - 30

Precursor charge range 1 - 4

Precursor m/z range 300 - 1800

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 Single-pass mode

Quantification strategy Robust LC (high precision)

Cross-run normalisation RT-dependent

Library generation Smart profiling

Speed and RAM usage Optimal results

Copy Paste

Clear pipeline Execute Abort pipeline

Step 1

Pipeline	Active	Status
----------	--------	--------

I

msms.txt

The screenshot shows the DIA-NN software interface. The 'Generate spectral library' and 'Quantities matrices' options are checked in the 'Output' section. The 'Spectral library' path is set to 'C:\Users\proteomics\Documents\DI...'. The 'msms.txt' file is highlighted in the log window.

Log Window Output:

```

0 files will be processed
[0:00] Loading spectral library C:\Users\proteomics\Documents\DIA-NN\MaxQuant
Result\txt\msms.txt
WARNING: support for MaxQuant search results is experimental; assuming all
cysteines are carbamidomethylated; MaxQuant does not include fixed modifications
in modified peptide names, so all other fixed modifications will need to be manually
annotated
[0:00] Finding proteotypic peptides (assuming that the list of UniProt ids provided for
each peptide is complete)
[0:00] Spectral library loaded: 4309 protein isoforms, 2264 protein groups and 6280
precursors in 5533 elution groups.
[0:00] Initialising library
[0:00] Saving spectral library to C:\Users\proteomics\Documents\DIA-NN
\Sequence Libraries\maxquant-library.tsv
[0:02] 6280 precursors saved
[0:02] Saving the library to C:\Users\proteomics\Documents\DIA-NN\MaxQuant
Result\txt\msms.txt.speclib
    
```


msms.txt

06/10/2022 14:47:05

+

Experiment name: 06/10/2022 14:47:05

Reset settings

Clear log

Save log

Input

Raw

diaPASEF.d

Clear list

Convert to .dia

Spectral library

C:\Users\proteomics\Documents\DI

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

0

N-term M excision

C carbamidomethylation

Ox(M)

Ac(N-term)

Phospho

K-GG

Peptide length range

7

-

30

Precursor charge range

1

-

4

Precursor m/z range

300

-

1800

Fragment ion m/z range

200

-

1800

Output

Use existing .quant files when available

Main output

C:\Program Files\DIA-NN\v1-8\report

Temp/.dia dir

Generate spectral library

Quantities matrices

Output library

C:\Users\proteomics\Documents\DI

Generate Proxit input from FASTA or spectral library

Precursor FDR (%)

1.0

Threads

6

Generate PDF report

Log level

1

Additional options

Run

Finished

Stop

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

Single-pass mode

Quantification strategy

Robust LC (high precision)

Cross-run normalisation

RT-dependent

Library generation

Smart profiling

Speed and RAM usage

Optimal results

Add to pipeline

Step 1

Remove step

Update step

Copy

Paste

Clear pipeline

Execute

Abort pipeline

Pipeline

Active

Status

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

WARNING: MBR turned off, two or more raw files are required

0 files will be processed

[0:00] Loading spectral library C:\Users\proteomics\Documents\DI

Result\txt\msms.txt

WARNING: support for MaxQuant search results is experimental; assuming all cysteines are carbamidomethylated; MaxQuant does not include fixed modifications in modified peptide names, so all other fixed modifications will need to be manually annotated

[0:00] Finding proteotypic peptides (assuming that the list of UniProt ids provided for each peptide is complete)

[0:00] Spectral library loaded: 4309 protein isoforms, 2264 protein groups and 6280 precursors in 5533 elution groups.

[0:00] Initialising library

[0:00] Saving spectral library to C:\Users\proteomics\Documents\DI

Sequence Libraries\maxquant-library.tsv

[0:02] 6280 precursors saved

[0:02] Saving the library to C:\Users\proteomics\Documents\DI

NN\MaxQuant Result\txt\msms.txt.speclib

[0:02] Log saved to C:\Users\proteomics\Documents\DI

NN\Sequence Libraries\maxquant-library.log.txt

Finished

DIA-NN exited

c1cc[nH]1

Experiment name:

06/10/2022 14:47:05

Reset settings

Clear log

Save log

Input

Raw

diaPASEF .d

Clear list

Convert to .dia

C:\Users\proteomics\Documents\DIANN\MS Data\220504kw_DIA_S7_2_S1-E3_1_9207.d

C:\Users\proteomics\Documents\DIANN\MS Data\220504kw_DIA_S7_3_S1-E4_1_9208.d

Spectral library

C:\Users\proteomics\Documents\DIANN\Sequence Libraries\uniprot_proteome.fasta

Add FASTA

Clear list

Reannotate

DIA-NN exe

diann.exe

Output

Use existing .quant files when available

Main output

C:\Users\proteomics\Documents\DIANN\Quantification\220504kw_DIA_S7_2_S1-E3_1_9207.d

Temp/.dia dir

Generate spectral library

Quantities matrices

Output library

Generate Prosit input from FASTA or spectral library

Precursor FDR (%)

1.0

Threads

6

Generate PDF report

Log level

1

Additional options

Run

Finished

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

1

-

4

Precursor m/z range

300

-

1800

Fragment ion m/z range

200

-

1800

Algorithm

Mass accuracy

15.0

Unrelated runs

MS1 accuracy

15.0

Use isotopologues

Scan window

5

MBR

Heuristic protein inference

No shared spectra

Protein inference

Genes

Neural network classifier

Single-pass mode

Quantification strategy

Robust LC (high precision)

Cross-run normalisation

RT-dependent

Library generation

Smart profiling

Speed and RAM usage

Optimal results

Add to pipeline

Remove step

Update step

Copy

Paste

Step 1

Pipeline	Active	Status

Clear pipeline

Execute

Abort pipeline

06/10/2022 14:47:05

+

Experiment name: 06/10/2022 14:47:05

Reset settings

Clear log

Save log

Input

Raw

diaPASEF .d

Clear list

Convert to .dia

C:\Users\proteomics\Documents\DIA-NN\MS Data\220504kw_DIA_S7_2_S1-E3_1_9207.d

C:\Users\proteomics\Documents\DIA-NN\MS Data\220504kw_DIA_S7_3_S1-E4_1_9208.d

Output

☐ Use existing .quant files when available

Main output

C:\Users\proteomics\Documents\DI\

Temp/.dia dir

☐ Generate spectral library

☒ Quantities matrices

Output library

☐ Generate Prosit input from FASTA or spectral library

Precursor FDR (%)

1.0

Threads

6

☒ Generate PDF report

Log level

1

Additional options

Spectral library

C:\Users\proteomics\Documents\DI\

Add FASTA

C:\Users\proteomics\Documents\DIA-NN\Sequence Libraries\uniprot_proteome.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

1

-

4

Precursor m/z range

300

-

1800

Fragment ion m/z range

200

-

1800

Add to pipeline

Remove step

Update step

Copy

Paste

Clear pipeline

Execute

Abort pipeline

Step 1

Pipeline	Active	Status

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

Mass accuracy will be fixed to 1.5e-05 (MS2) and 1.5e-05 (MS1)

The following variable modifications will be scored: UniMod:1 UniMod:21

2 files will be processed

[0:00] Loading spectral library C:\Users\proteomics\Documents\DIA-NN\Sequence Libraries\maxquant-library.tsv

[0:00] Spectral library loaded: 4309 protein isoforms, 2264 protein groups and 6280 precursors in 5533 elution groups.

[0:00] Loading protein annotations from FASTA C:\Users\proteomics\Documents\DIA-NN\Sequence Libraries\uniprot_proteome.fasta

[0:03] Annotating library proteins with information from the FASTA database

[0:03] Protein names missing for some isoforms

[0:03] Gene names missing for some isoforms

[0:03] Library contains 1957 proteins, and 1954 genes

[0:03] Initialising library

[0:03] Saving the library to C:\Users\proteomics\Documents\DIA-NN\Sequence Libraries\maxquant-library.tsv.speclib

Results

Name	
 report.gg_matrix	
 report.log	
 report	
 report.pg_matrix	
 report.pr_matrix	
 report.stats	
 report	
 report.unique_genes_matrix	
 report-first-pass.gg_matrix	
 report-first-pass.pg_matrix	
 report-first-pass.pr_matrix	
 report-first-pass.stats	
 report-first-pass	
 report-first-pass.unique_genes_matrix	
 report-lib.predicted.speclib	

Statistical Analysis of Quantitative Results

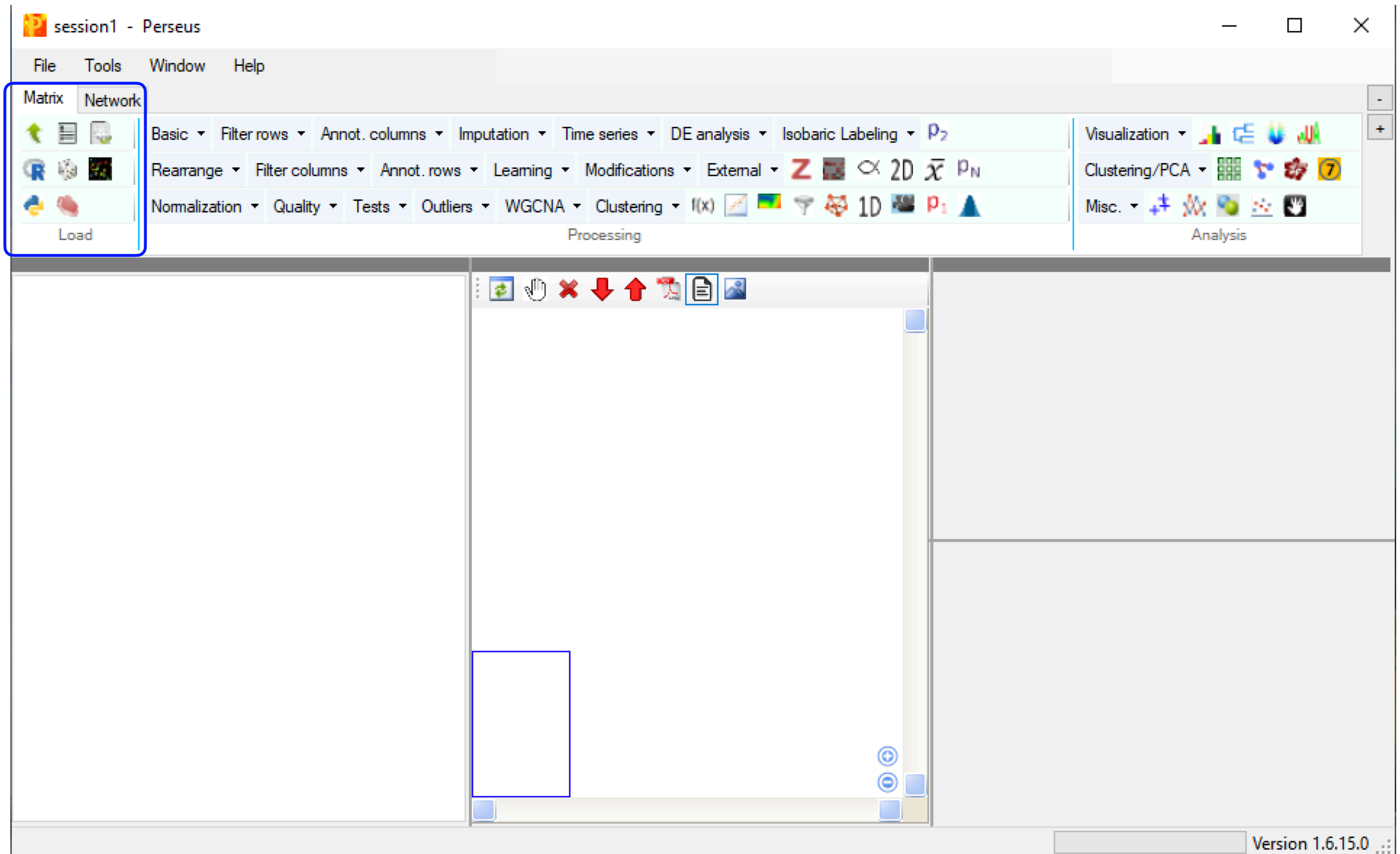


report.pg-matrix.tsv

AAAGLNRFCLEHMAIS (ECAAAGDFTLTPVW)GENEYS

Results

report.pg-matrix.tsv



The screenshot displays the Perseus software interface for a session named 'session1'. The main window is divided into several sections:

- Menu Bar:** File, Tools, Window, Help.
- Tab Bar:** Matrix (selected), Network.
- Processing Tools:** A central panel with multiple tabs for data processing:
 - Basic
 - Filter rows
 - Annot. columns
 - Imputation
 - Time series
 - DE analysis
 - Isobaric Labeling
 - Rearrange
 - Filter columns
 - Annot. rows
 - Learning
 - Modifications
 - External
 - Normalization
 - Quality
 - Tests
 - Outliers
 - WGCNA
 - Clustering
 - f(x)
- Analysis Tools:** A right-hand panel with tabs for:
 - Visualization
 - Clustering/PCA
 - Misc.
- Main Canvas:** A large white area for data visualization. It features a toolbar with icons for zooming (hand, arrows), a document icon, and a person icon. A small blue box is visible in the bottom-left corner of the canvas.
- Status Bar:** Located at the bottom right, showing 'Version 1.6.15.0'.

Results

report.pg-matrix.tsv


Generic matrix upload

Cancel

Description

File

C:\Users\proteomics\Documents\DIANA\Results\Example\report.pg_matrix.tsv

Select

Protein.Group

Protein.Ids

Protein.Names

Genes

First.Protein.Description

D:\Experiment_1+1_T-A5_1_14498.d

D:\Experiment_1+1_T-A6_1_14499.d

D:\Experiment_1+2_T-A7_1_14500.d

D:\Experiment_1+2_T-A8_1_14501.d

D:\Experiment_1-1_T-A1_1_14494.d

D:\Experiment_1-1_T-A2_1_14495.d

D:\Experiment_1-2_T-A3_1_14496.d

D:\Experiment_1-2_T-A4_1_14497.d

D:\Experiment_2+1_T-B1_1_14506.d

D:\Experiment_2+1_T-B2_1_14507.d

D:\Experiment_2+2_T-B3_1_14508.d

D:\Experiment_2+2_T-B4_1_14509.d

D:\Experiment_2-1_T-A9_1_14502.d

D:\Experiment_2-1_T-A10_1_14503.d

D:\Experiment_2-2_T-A11_1_14504.d

D:\Experiment_2-2_T-A12_1_14505.d

D:\Experiment_3+1_T-B9_1_14514.d

D:\Experiment_3+1_T-B10_1_14515.d

D:\Experiment_3+2_T-B11_1_14516.d

D:\Experiment_3+2_T-B12_1_14517.d

D:\Experiment_3-1_T-B5_1_14510.d

D:\Experiment_3-1_T-B6_1_14511.d

D:\Experiment_3-2_T-B7_1_14512.d

D:\Experiment_3-2_T-B8_1_14513.d

Main

>

<

↑

↑

↓

↓

Numerical

>

<

↑

↑

↓

↓

Categorical

>

<

↑

↑

↓

↓

Text

>

<

↑

↑

↓

↓

Multi-numerical

Cancel

Description


OK

Results

report.pg-matrix.tsv

Generic matrix upload

Cancel Description

OK

Main

- > D:\Experiment_1+1_T-A5_1_14498.d
- > D:\Experiment_1+1_T-A6_1_14499.d
- < D:\Experiment_1+2_T-A7_1_14500.d
- < D:\Experiment_1+2_T-A8_1_14501.d
- < D:\Experiment_1-1_T-A1_1_14494.d
- < D:\Experiment_1-1_T-A2_1_14495.d
- < D:\Experiment_1-2_T-A3_1_14496.d
- < D:\Experiment_1-2_T-A4_1_14497.d
- D:\Experiment_2+1_T-B1_1_14506.d

Numerical

Categorical

Text

- > Protein.Group
- > Protein.Ids
- < Protein.Names
- < Genes
- < First.Protein.Description

Multi-numerical

Cancel Description

OK

Results

report.pg-matrix.tsv

session1 - Perseus

File Tools Window Help

Matrix Network

Basic Filter rows Annot. columns Imputation Time series DE analysis Isobaric Labeling P_2

Rearrange Filter columns Annot. rows Learning Modifications External Z α $2D$ $\bar{x}$ P_N

Normalization Quality Tests Outliers WGCNA Clustering $f(x)$ $1D$ P_1

Load Processing

Visualization Clustering/PCA Misc. Analysis

Basic Annot. rows Multi-proc. Export

matrix1

Data

Type	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main	Main
1	16362.7	NaN	14611.7	15942.2	16206.9	19380.8	13917	NaN	16321.1	20755.7	22787.3	23013.7	21884.9	16227.
2	13689	12559.2	14525.7	12783.6	12972.9	12472.5	12485.1	13294.5	19402.2	16063.5	16929.9	18703	19171.7	16704
3	49362.7	48493.6	51178.2	51033.7	47102	49008.9	49306.1	51603.5	22620.7	21289.8	22758.2	19751.3	24597	23397.
4	3124.83	NaN	5468.58	4751.57	3109.21	3329.75	4774.23	4302.72	5576.75	4275.63	6540.52	7755.79	3969.76	4931.2
5	4366.98	4230.57	3465.4	3762.76	4165.3	3968.61	3086.56	3487.11	4603.6	4997.01	5253.84	5208.8	5174.88	4227.8
6	11746.3	11615.3	9610.2	9359.2	11201.5	10335.9	11792.5	9959.76	4777.7	4568.58	4766.66	4683.47	7543.84	5100.3
7	2535.6	3214.95	3960.04	3340.1	2594.93	3202.26	3434.95	3063.48	1259.63	NaN	NaN	1371.35	NaN	1226.0
8	NaN	4586.5	NaN	NaN	5468	NaN	NaN	NaN	NaN	6580.74	NaN	NaN	NaN	NaN
9	21190.4	22018.8	22844.8	25957.6	24215.6	23076.6	27526.5	21344.1	19578.1	19310.9	19727.2	19143.5	20046.7	23635.
10	6360.15	10470.8	7700.89	8829.86	4757.46	9109.97	7669.72	10011.3	10298.7	8407.45	6366.81	6963.98	11683.4	14938.
11	NaN	78844.5	66307.3	72354	NaN	NaN	135560	110130	56685.2	154158	131940	88389.6	62610	88349.
12	1303.65	1312.85	6070.01	5075.85	5153.61	5518.65	6088.21	2896.87	3244.89	3517.4	3964.78	1926.73	2906.44	4301.7
13	3512.78	1594.17	4223.33	4425.62	1341.98	2348.65	NaN	2201.88	3370.12	2450.01	NaN	NaN	3579.03	3941.4
14	7688.18	7395.88	7717.69	8333.08	8393.32	8391.34	7413.47	8208.67	4107.65	4843.59	4696.72	4844.23	5067.4	3817.3
15	3743.18	3411.68	NaN	3958.55	5583.55	5008.46	NaN	NaN	NaN	4780.18	4890.92	4994.33	4414.55	NaN
16	10522.9	10883.6	9110.73	10434.7	11010.9	10132.8	8983.97	8433.22	7934.97	11729.5	11092.6	9822.29	12398	12306.
17	23604.5	26180.1	29143.7	30325.2	28956.5	26457.9	33088.9	29291.7	46236.5	45584.4	34149.4	34558.7	46340.4	41331
18	10702.9	8387.9	8282.15	10084.4	9059.49	10810.5	11456.9	9568.66	3723.88	4662.29	5422.84	4695.59	4474.15	3837.7
19	2934.38	3251.93	1740.31	2946.23	2448.53	2652.91	2355.62	2162.71	2353.66	2085.1	2757.46	3239.87	2006.68	2798.7
20	4452.94	3674.98	5578.44	4284.49	5210.37	5817.32	4400.1	5734.18	10082.5	9343.07	6613.73	6319.92	5125.8	4473.8

4,050 items

100 %

Version 1.6.15.0

matrix1

- Creator: proteomics
- 05/22/2023 11:13:58
- Origin: C:\Users\proteomics\
- File: report.pg_matrix.tsv
- Quality: (small values are goo
- Rows (4,050)
- Main columns (24)
- Categorical columns (0)
- String columns (5)
- Numerical columns (0)
- Multi-numerical columns (0)
- Categorical rows (0)
- String rows (0)
- Numerical rows (0)
- Multi-numerical rows (0)

Conclusions

- Perseus is the tool of choice to analyse phosphorylation data
- Because of under sampling of proteomic samples special methods must be used to be able to re-measure phosphorylated peptides
- This can be multi-dimensional chromatography or Data Independent Acquisition methods on the mass spectrometer (using optionally an experimental library of fragmented, phosphorylated peptides)
- DIA Data is best analysed using DIA-NN
- Perseus is used to statistically evaluate DIA result files



End