



Proteome Discoverer 2.4 overview

Software for identification and quantification of proteins in complex samples

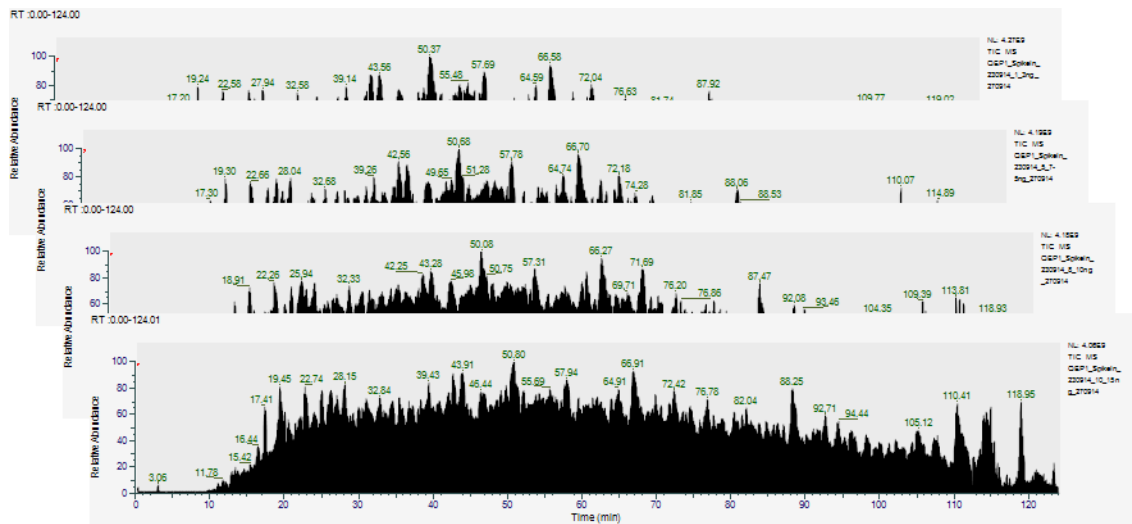


1.0 - July, 2008	Workflow-based solution for protein identification
1.1 - Nov, 2009	Automation, batch processing, reporter ion quantification
1.2 - Apr, 2011	SILAC quantification
1.3 - Mar, 2012	Validation (Percolator, PhosphoRS), biological annotation, 64-bit
1.4 - May, 2013	Deep data mining (Sequest HT, library searching)
2.0 - Mar, 2015	Architectural changes, study management, large data sets
2.1 - Oct, 2015	Improved reporter ion quantification
2.2 – July, 2017	Label Free Quantification, Statistics, Cross-Linking
2.3 – Jan, 2019	Improved library search, heat maps, PTM site tables, cross-link quantification, annotation groups, ProSightPD
2.4 – Oct, 2019	Precursor detector/chimeric spectra, scripting node, TMTpro, new licensing

What is Proteome Discoverer?

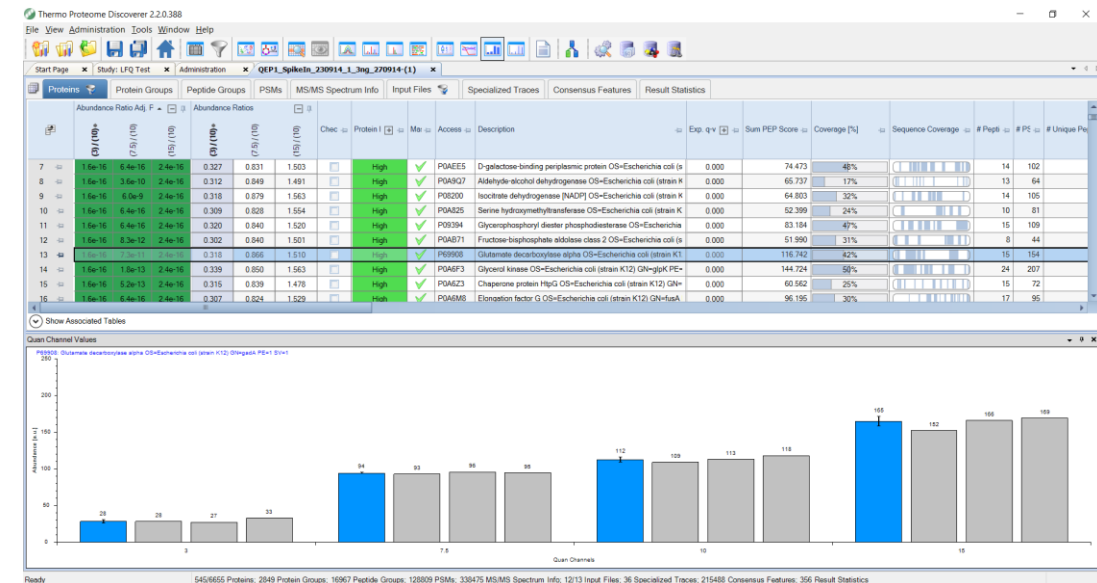
- Software for analysis of quantitative discovery proteomics data
- Supports hybrid Orbitrap, Q Exactive, and ion trap mass spectrometers
- Is software for simple conversion of raw files to protein ID lists (with quan) enough?

RAW files



?

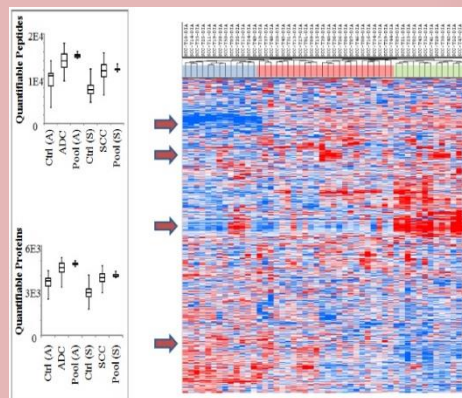
Identified proteins and quantitative results



Challenges for analysis of quantitative proteomics data

Complex studies/Large datasets

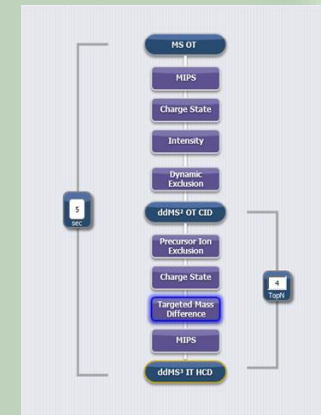
- 100's of raw files
- Results need to be presented by sample, not raw file
- Statistics and proper study design are required



Requires study management

Complex acquisition methods

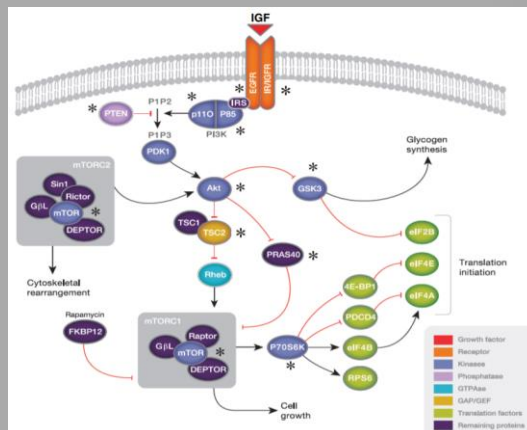
- TMT – SPS MS³
- Glycopeptides – HCD-triggered-> CID->ET_hcD
- Cross-linking – MS2/MS2/MS3
- Top down – CID, ETD, HCD, ET_hcD, UVPD



Requires customizable workflows

Biological complexity

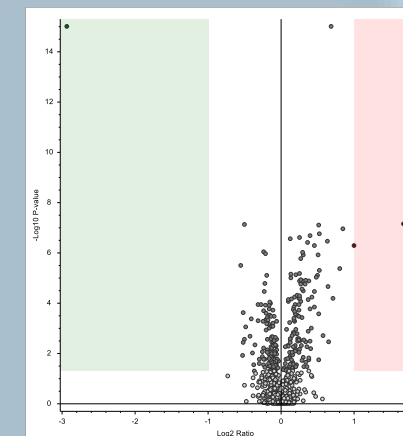
- >10,000 proteins
- PTMs
- Proteoforms
- Metaproteomics
- Pathway analysis
- Protein structure



Requires links to bioinformatic databases

Results interpretation

- How to denote significantly changing proteins/peptides?
- What is already known about proteins of interest?
- How do we make biological conclusions?



Requires statistics and visualization

What is Proteome Discoverer?

Study Management

- Maps “study factors” to quantification channels
 - Set up replicates, statistical analysis
 - Manage files and search results
-

Client/server based workflow processing system

- Customizable data analysis pipelines for complex acquisition methods
 - Extensible framework allows faster deployment of new algorithms
 - Support for large datasets
-

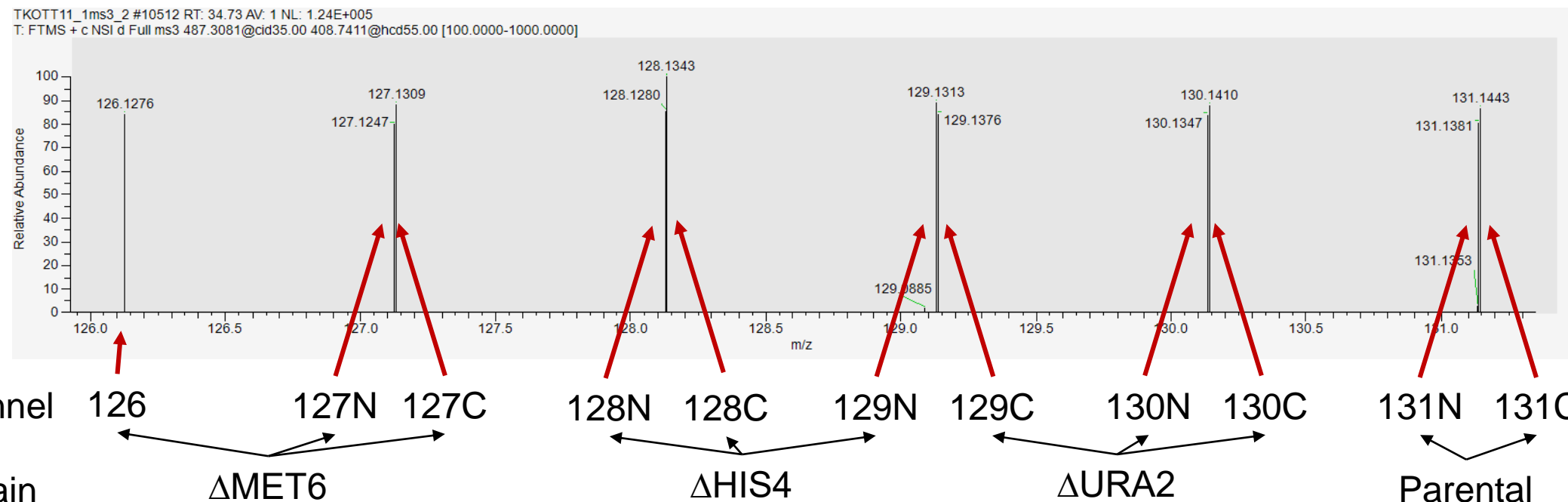
Biological Annotation

- Pathway, GO term, protein family annotation
 - ProteinCard for summary of known information of selected proteins
 - Links to KEGG, Wikipathway, Reactome maps
-

Data Interpretation

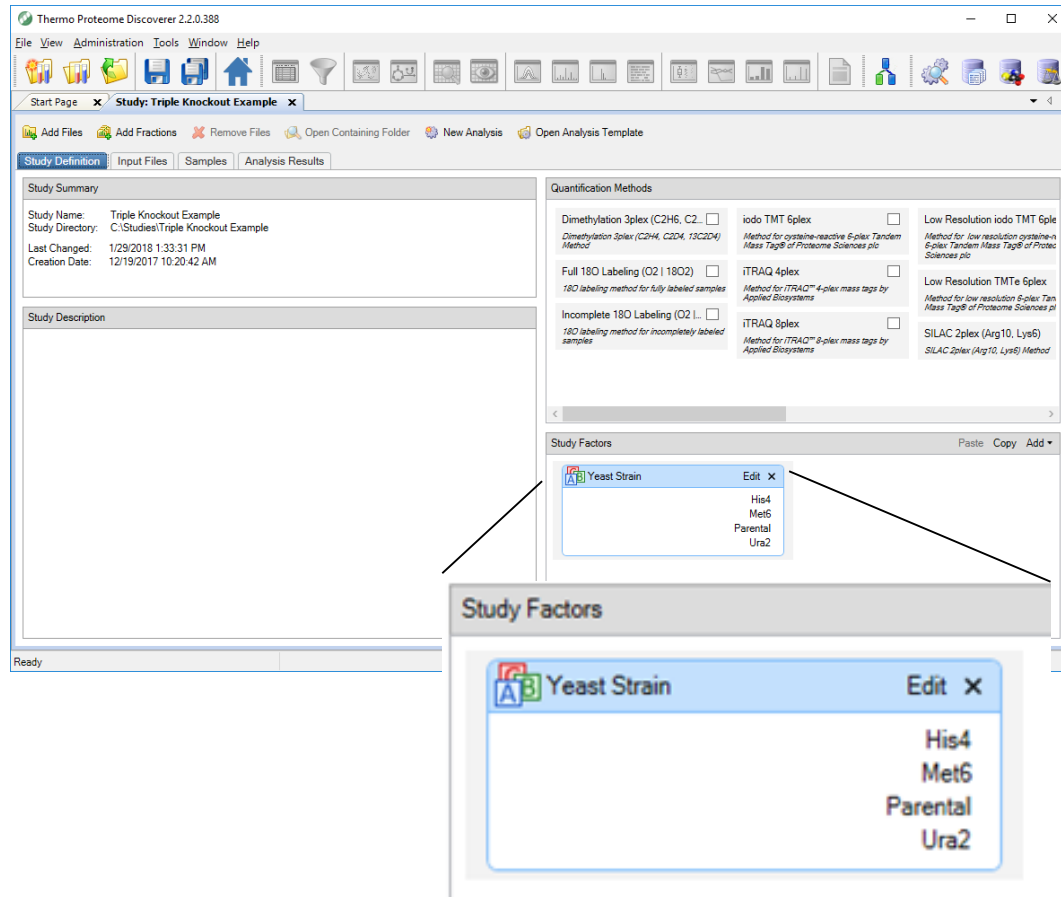
- Hierarchical views with links between proteins, peptides, PSMs
- Interactive graphical views for statistical analysis

Enables users to assign biological meaning (“study factors”) to quan channels:



Study Management in PD – Yeast Triple Knockout

- Create the 4 yeast strains as a study factor



- Assign study factors to the quan channels

Study Definition			
Input Files			
Samples			
Analysis Results			
Sample	Sample Identifier	Sample Type	Yeast Strain
S1	TKOTT11_1ms3_1 - [126]	Control	Met6
S2	TKOTT11_1ms3_1 - [127N]	Sample	Met6
S3	TKOTT11_1ms3_1 - [127C]	Sample	Met6
S4	TKOTT11_1ms3_1 - [128N]	Sample	His4
S5	TKOTT11_1ms3_1 - [128C]	Sample	His4
S6	TKOTT11_1ms3_1 - [129N]	Sample	His4
S7	TKOTT11_1ms3_1 - [129C]	Sample	Ura2
S8	TKOTT11_1ms3_1 - [130N]	Sample	Ura2
S9	TKOTT11_1ms3_1 - [130C]	Sample	Ura2
S10	TKOTT11_1ms3_1 - [131N]	Sample	Parental
S11	TKOTT11_1ms3_1 - [131C]	Sample	Parental

Study Management in PD – Yeast Triple Knockout

- Create quantitative ratios based on the Yeast Strain study factor:

Thermo Proteome Discoverer 2.2.0.388

File View Administration Tools Window Help

Start Page x Study: Triple Knockout Example x

Add Files Add Fractions Remove Files Open Containing Folder New Analysis Open Analysis Template

Study Definition Input Files Samples Analysis Results Workflows Grouping & Quantification

Sample Group and Quan Ratio Specification

Study Variables

- ☐ File
- ☐ Quan Channel
- ☒ Yeast Strain
- ☐ Sample Type

Variables printed in italics contain only a single value.

Manual Ratio Generation

Numerator: Add Ratio

Denominator:

Bulk Ratio Generation

Denominators to be used:

- ☐ Yeast Strain : Met6
- ☐ Yeast Strain : His4
- ☐ Yeast Strain : Ura2
- ☒ Yeast Strain : Parental

Generated Sample Groups

Met6

126	Control	Met6	F1: TKOTT11_1ms3_1
127N	Sample	Met6	F1: TKOTT11_1ms3_1
127C	Sample	Met6	F1: TKOTT11_1ms3_1

His4

128N	Sample	His4	F1: TKOTT11_1ms3_1
128C	Sample	His4	F1: TKOTT11_1ms3_1
129N	Sample	His4	F1: TKOTT11_1ms3_1

Ura2

Generated Ratios

- X Met6 / Parental
- X His4 / Parental
- X Ura2 / Parental

Analysis

Consensus Step

Workflow: CWF_Comprehensive_Enhanced Annotation_Reporter_Quan
Result File: TKOTT11_1ms3_1.pdResult

Child Steps: (1)

Processing Step

Workflow: PWF_Fusion_TMT_Quan_SPS_MS3_SequestHT_Percolator
Result File: TKOTT11_1ms3_1.msf

Files for Analysis: (1)

x	F1	TKOTT11_1ms3_1	TMT 11plex SG253828	Sample Type: [Control, Sample], Yeas
---	----	----------------	---------------------	--------------------------------------

Ready

Ratios based on study factor, not quan channel

Study Management in PD – Yeast Triple Knockout

- View results based on study factor rather than quan channel

Thermo Proteome Discoverer 2.2.0.388

File View Administration Tools Window Help

Start Page x Study: Triple Knockout

Proteins Protein Groups

Specialized Traces

Abundance Ratios

Abundance Ratio

Checked Protein FDI Master Gene Symbol Accession Description

1 High PET9 P18239 ADP/ATP carrier protein 2 [OS=Saccharomyces cerevisiae]

2 High ATP2 P00830 ATP synthase subunit beta, mitochondrial [OS=Saccharomyces cerevisiae]

3 High TDH1 P00360 glyceraldehyde-3-phosphate dehydrogenase 1 [OS=Saccharomyces cerevisiae]

4 High SHM2 P37291 Serine hydroxymethyltransferase, cytosolic [OS=Saccharomyces cerevisiae]

5 High ADE1 P27616 phosphoribosylaminoimidazole-succinocarboxamide synthetase [OS=Saccharomyces cerevisiae]

6 High GCV2 P49095 Glycine dehydrogenase (Decarboxylating), mitochondrial [OS=Saccharomyces cerevisiae]

7 High FAS2 P19097 Fatty acid synthase subunit alpha [OS=Saccharomyces cerevisiae]

8 High YBR085C-A O43137 Uncharacterized protein YBR085C-A [OS=Saccharomyces cerevisiae]

9 High MET6 P05694 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase [OS=Saccharomyces cerevisiae]

10 High RPS21B Q3E754 40S ribosomal protein S21-B [OS=Saccharomyces cerevisiae]

11 High MDH1 P17505 Malate dehydrogenase, mitochondrial [OS=Saccharomyces cerevisiae]

12 High ARG1 P22768 Argininosuccinate synthase [OS=Saccharomyces cerevisiae]

13 High SAM1 P10659 S-adenosylmethionine synthase 1 [OS=Saccharomyces cerevisiae]

14 High QCR6 P00127 Cytochrome b-c1 complex subunit 6 [OS=Saccharomyces cerevisiae]

15 High GLK1 P17709 Glucokinase-1 [OS=Saccharomyces cerevisiae S288C]

16 High HSP26 P15992 heat shock protein 26 [OS=Saccharomyces cerevisiae S288C]

17 High SAC6 P32599 fimbrin [OS=Saccharomyces cerevisiae S288C]

18 High ADE6 P38972 Phosphoribosylformylglycinamidine synthase [OS=Saccharomyces cerevisiae]

19 High ADE3 P07245 C-1-tetrahydrofolate synthase, cytoplasmic [OS=Saccharomyces cerevisiae]

Abundance Ratios

(Met6) / (Parental)

(His4) / (Parental)

(Ura2) / (Parental)

Abundance Ratio

1.646 1.228 1.145 1.4e-13 6.8e-5 2.2e-4 131.2 97.9 91.3 79.7 1.23 1.2

1.518 1.153 1.165 1.4e-13 3.0e-4 8.1e-5 125.6 95.4 96.3 82.7 0.21 1.5

2.310 0.958 1.116 1.4e-13 6.4e-1 1.7e-2 171.6 71.2 82.9 74.3 1.17 2.7

2.354 1.443 1.932 1.4e-13 2.5e-5 3.1e-13 139.9 85.8 114.9 59.4 0.19 1.0

1.786 1.318 1.639 1.4e-13 2.5e-5 3.1e-13 124.4 91.8 114.1 69.6 0.89 1.6

2.106 1.374 1.612 1.4e-13 4.0e-5 4.7e-6 138.3 90.2 105.8 65.7 1.25 0.5

1.224 1.094 1.204 1.4e-13 2.9e-5 3.1e-13 108.3 96.8 106.5 88.4 0.09 0.5

1.595 1.612 1.401 1.4e-13 6.1e-13 5.2e-6 113.8 115.0 99.9 71.3 1.08 1.4

0.073 1.318 1.251 1.4e-13 2.3e-2 3.6e-2 8.0 144.7 137.4 109.8 9.91 0.2

0.770 1.098 0.914 4.5e-7 3.4e-4 1.5e-4 81.4 116.1 96.6 105.8 0.68 0.5

1.590 1.194 1.071 1.9e-6 4.0e-4 1.0e-2 131.0 98.4 88.2 82.4 0.67 1.2

3.735 3.209 3.877 3.1e-6 1.6e-5 3.7e-6 126.4 108.6 131.2 33.8 2.66 2.5

1.531 1.070 1.181 3.6e-6 3.1e-2 2.5e-4 128.1 89.5 98.8 83.7 1.60 0.5

1.988 1.806 1.491 3.6e-6 1.9e-5 2.6e-5 126.5 114.9 94.9 63.6 1.91 1.4

1.456 1.048 0.964 3.7e-6 1.7e-1 1.4e-1 130.3 93.8 86.3 89.5 1.06 2.7

2.077 0.925 0.869 3.7e-6 2.5e-1 1.0e-2 170.5 76.0 71.4 82.1 1.26 3.7

1.226 1.099 1.284 4.0e-6 3.0e-4 2.2e-6 106.4 95.4 111.4 86.8 0.79 0.4

1.230 1.073 1.239 4.1e-6 1.5e-3 4.6e-6 108.3 94.5 109.1 88.1 0.20 1.7

1.445 1.194 1.488 4.3e-6 3.0e-4 4.6e-6 112.8 93.2 116.1 78.0 2.01 1.0

Show Associated Tables

Ready 1010/1275 Proteins; 1010 Protein Groups; 5855 Peptide Groups; 6338 PSMs; 16687 MS/MS Spectrum Info; 1/2 Input Files; 2 Specialized Traces

User customizable workflows for processing complex datasets

Thermo Proteome Discoverer 2.2.0.388

File View Administration Tools Window Help

Start Page x Study: Triple Knockout Example x

Open Containing Folder New Analysis Open Analysis Template

Analysis Results Workflows Grouping & Quantification

Open Open Common Save Save Common Auto Layout Clear

Workflow: PWF_OT_ETD_CID_SequestHT_Percolator

Description: Workflow for processing LTQ Orbitrap raw files with ETD and CID spectra, with Sequest HT. Add the modifications in the Sequest HT node.

Workflow Tree

List of nodes

- Data Input
 - Spectrum Files
 - Spectrum Files RC
- Spectrum Retrieval
 - ProSightPD Spectrum Selector
 - Spectrum Selector
- Feature Detection & Quantification
 - Minora Feature Detector
 - Reporter Ions Quantifier
- Spectrum Processing
 - Noise Peak Filter
 - Non-Fragment Filter
 - Parallel Xtract
 - ProSightPD Top Down High/High cRA...
 - ProSightPD Top Down Low/High cRA...
 - Spectrum Grouper
 - Spectrum Normalizer
 - Top N Peaks Filter
- Spectrum Filters
 - Scan Event Filter
 - Spectrum Confidence Filter
 - Spectrum Properties Filter
- Sequence Database Search
 - MS Amanda
 - Mascot
 - PMI-Byonic
 - PMI-Preview
 - ProSightPD Absolute Mass Search
 - ProSightPD BioMarker Search
 - ProSightPD Gene Restricted Absolute...
 - ProSightPD Gene Restricted BioMarke...
 - Sequest HT
- Spectral Library Search
 - MCScanSearch

Customizable workflow tree

```
graph TD; A[Spectrum Files RC 0] --> B[Spectrum Selector 1]; B --> C[Scan Event Filter 2]; B --> D[Scan Event Filter 10]; C --> E[Non-Fragment Filter 3]; D --> F[Sequest HT 11]; E --> G[Spectrum Properties Filter 4]; E --> H[Spectrum Properties Filter 7]; F --> I[Percolator 12]; G --> J[Sequest HT 5]; H --> K[Sequest HT 8]; J --> L[Percolator 6]; K --> M[Percolator 9]
```

Analysis

Consensus Step

Workflow: CWF_Comprehensive_Enhanced Annotation

Result File: Enter result file name.

Child Steps: (1)

Processing Step

Workflow: PWF_OT_ETD_CID_SequestHT_Percolator

Result File: Enter result file name.

Files for Analysis: (0)

Clear All

Drag and drop from Input Files here

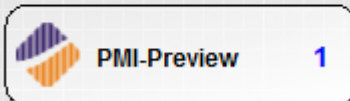
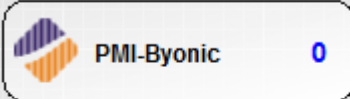
Search engine nodes available in the PD framework

- Proteome Discoverer framework includes several search engines
- The nodes can be used in series or in parallel to identify more peptides than each individual search engine
- More search engines are currently being developed by 3rd parties
- Byonic, ProSightPD, Mascot require purchase of the standalone search engines

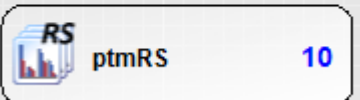
 Sequest HT 17	Default search engine	 PMI-Byonic 14	Glycosylation/PTM
 Mascot 16	Connects to a local Mascot server	 PMI-Preview 15	Quick preview search
 MS Amanda 13	IMP Vienna (Mechtler group)	 ProSightPD Absolute Mass Search 19	Top down protein identification
 MSPepSearch 18	Library search engine from NIST	 ProSightPD BioMarker Search 20	Truncated protein identification

Third party nodes (or nodes that encapsulate 3rd party algorithms)

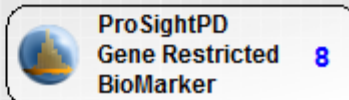
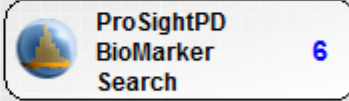
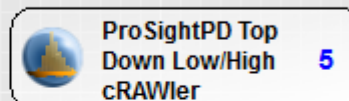
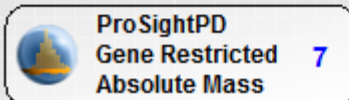
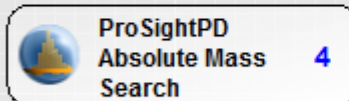
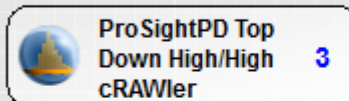
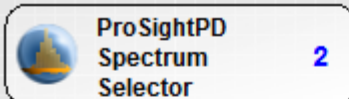
Protein Metrics



IMP Vienna



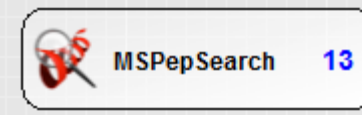
Proteinaceous/Northwestern



Nodes distributed only by external sites:

- pd-nodes.org (IMP Vienna)
- www.openms.de (PD 2.1)
- SuperQuant (PD 2.1)

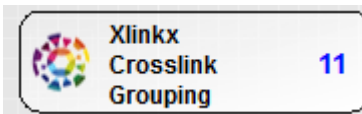
NIST



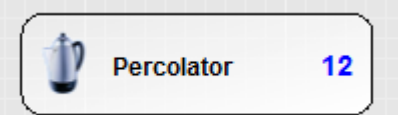
Matrix Science



University of Utrecht



UW/KTH



Proteome Discoverer Visualization – Hierarchical tables and row filters

Thermo Proteome Discoverer 2.2.0.388

File View Administration Tools Window Help

Start Page x Study: Triple Knockout Example x TKOTT11_1ms3_1 x

Display Filter

Load Save Clear Clear All Apply Cancel

ON Proteins
ON Protein Groups
ON Peptide Groups
ON PSMs

Proteins

AND Add group

Master is equal to Master Remove

KEGG Pathways contains Methionine Remove

Proteins Protein Groups Peptide Groups PSMs MS/MS Spectrum Info Input Files Specialized Traces

Protein list

Checked	Protein FDR	Master	Gene Symbol	Accession	Description	Abundance Ratios	Abundance Ratio Adj.	P-Value	Abundances (Grouped)	Abundances (Grouped) CVs [%]	# Unique Peptides	Found in S										
<input type="checkbox"/>	High	✓	MET6	P05694	5-methyltetrahydropteroyltrimethylglutamate-homocysteine methyltransferase [O	0.073	1.318	1.251	1.4e-13	2.3e-2	3.6e-2	8.0	144.7	137.4	109.8	9.91	0.21	1.18	3.07	32	✓	
2	<input type="checkbox"/>	High	✓	MDH1	P17505	Malate dehydrogenase, mitochondrial [OS=Saccharomyces cerevisiae S28C]	1.590	1.194	1.071	1.9e-6	4.0e-4	1.0e-2	131.0	98.4	88.2	82.4	0.67	1.25	2.27	0.12	7	✓
3	<input type="checkbox"/>	High	✓	SAM1	P10659	S-adenosylmethionine synthase 1 [OS=Saccharomyces cerevisiae S288C]	1.531	1.070	1.181	3.6e-6	3.1e-2	2.5e-4	128.1	89.5	98.8	83.7	1.60	0.90	0.94	2.37	6	✓
4	<input type="checkbox"/>	High	✓	SAH1	P39954	Adenosylhomocysteinase [OS=Saccharomyces cerevisiae S288C]	1.352	0.964	1.044	8.1e-6	2.5e-1	1.3e-1	124.0	88.5	95.8	91.7	0.25	0.48	2.12	2.04	15	✓

Peptide groups for selected protein

Reference	Annotated Sequence	Modifications	Quan Info	Quan Usage	Abundance Ratios	Abundance Ratio Adj.	P-Value	Abundances (Grouped)	Abundances (Grouped) CVs [%]									
2	[K] FVVPDCLGK[T]	1×Carbamidomethyl [C7]; 1×TMT6plex [K10]; 1×TMT6plex [N-Term]		Used	0.010	1.413	1.425	1.6e-3	1.0e0	1.0e0	1.0	146.9	148.1	104.0	83.45	4.40	2.95	2.69
3	[K] GTISAEYEK[F]	1×TMT6plex [K10]; 1×TMT6plex [N-Term]		Used	0.053	1.273	1.249	2.3e-1	2.8e-1	5.9	142.5	139.8	111.9		3.24	9.29	2.51	

PSMs for selected peptide group

Confidence	Annotated Sequence	Modifications	# Protein Groups	Master Protein Accessions	Charge	Rank	m/z [Da]	ΔM [ppm]	Average Reporter S/N	Abundances									
100	[K] FVVPDCLGK[T]	N-Term(TMT6plex); C7(Carbamidomethyl)	1	P05694	2	1	847.46016	-0.35	341.6	2.0	3.1	10.7	476.9	481.7	454.5	490.9	492.3	477.0	323.4

Quan results for selected protein or peptide group

Quan Channel Values

Fragment Match Spectrum

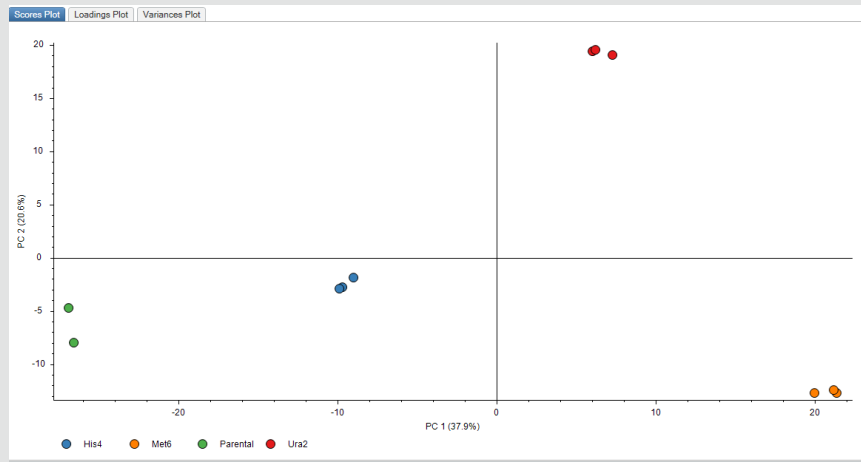
TKOTT11_1ms3_1.raw #26200 RT: 61.4491 min
ITMS, 847.4609@cid35.00, z=+2, Mono m/z=847.46016 Da, MH+=1693.91303 Da, Match Tol=0.6 Da

Annotated MS/MS spectrum

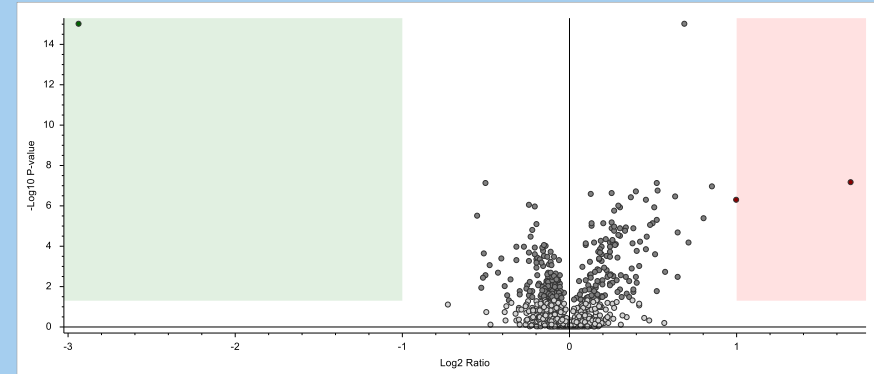
17/1275 Proteins; 1010 Protein Groups; 5855 Peptide Groups; 6338 PSMs; 16687 MS/MS Spectrum Info; 1/2 Input Files; 2 Specialized Traces

Proteome Discoverer Tools for statistical and biological interpretation

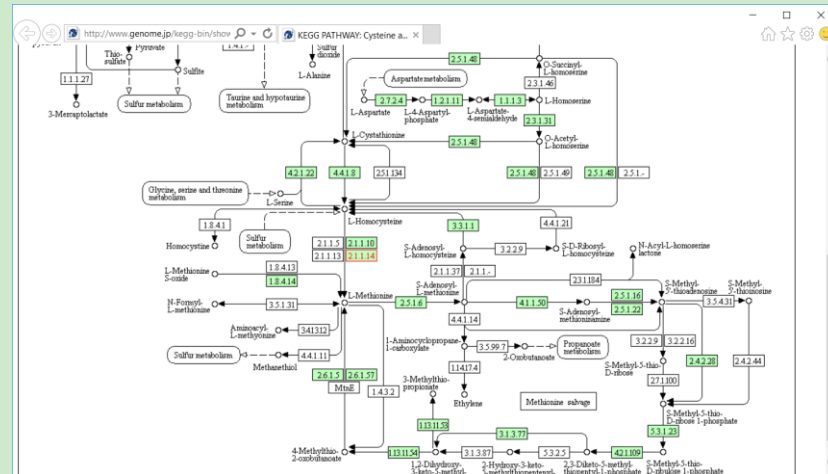
- PCA



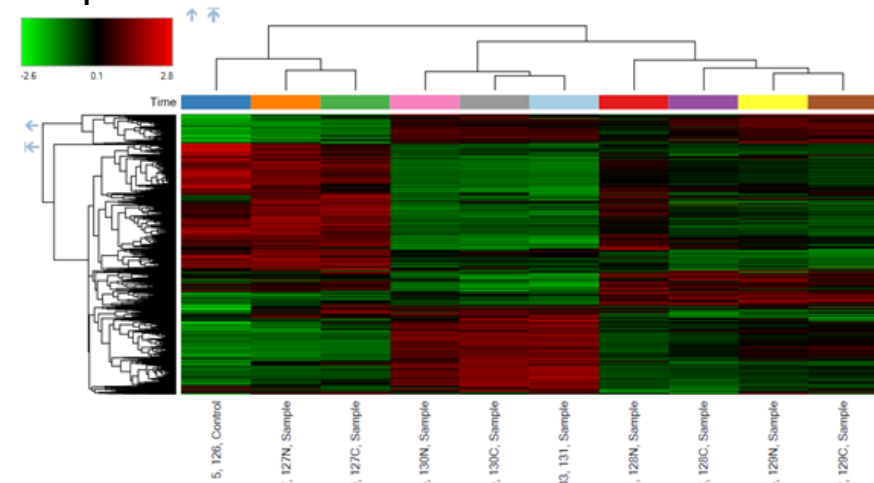
- Volcano plots



- Pathway maps



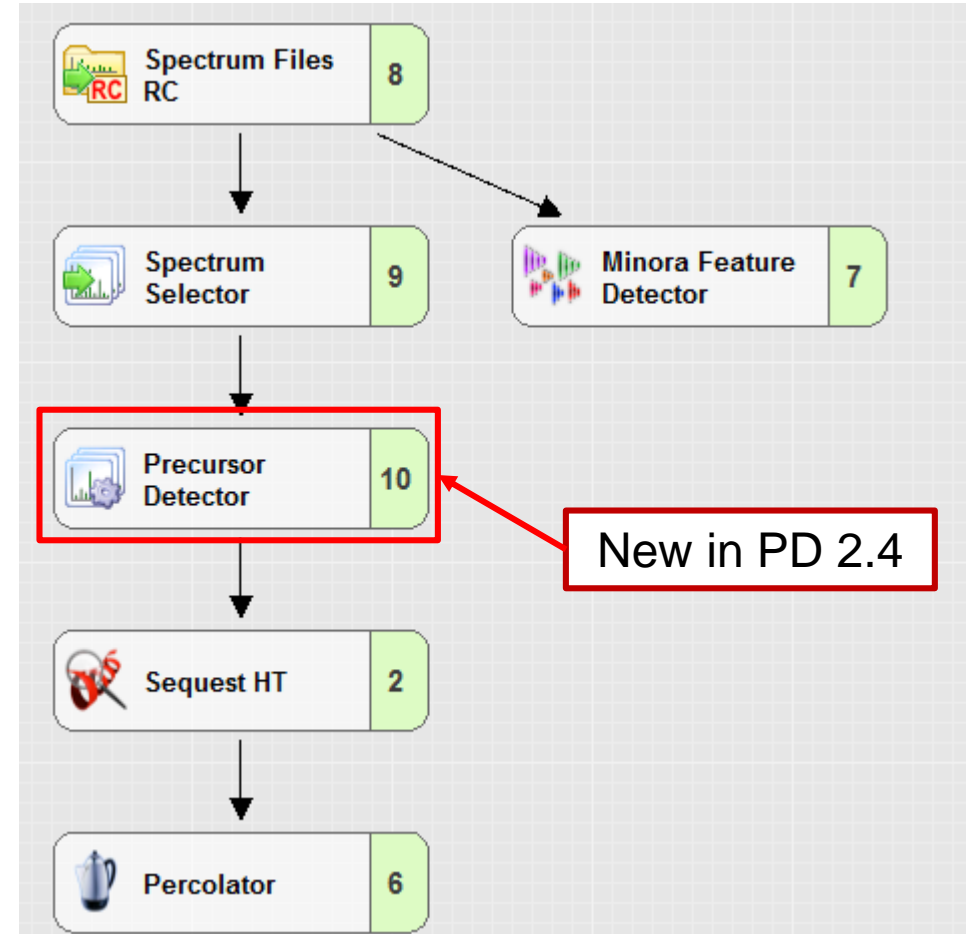
- Heat maps



- Mixed spectra/chimeric spectra support
 - Precursor Detector node
 - Works with Sequest HT and MSPepSearch search engines only
- Scripting node
 - Incorporate R, Python, or other scripts into Proteome Discoverer workflows
- New TMTpro 16plex method
- FAIMS LFQ – license no longer needed
- New Flexera licensing (like Thermo Scientific™ BioPharma Finder™ software)
- Updated Proteome Tools libraries coming soon
- Updated Familiarization Exercises

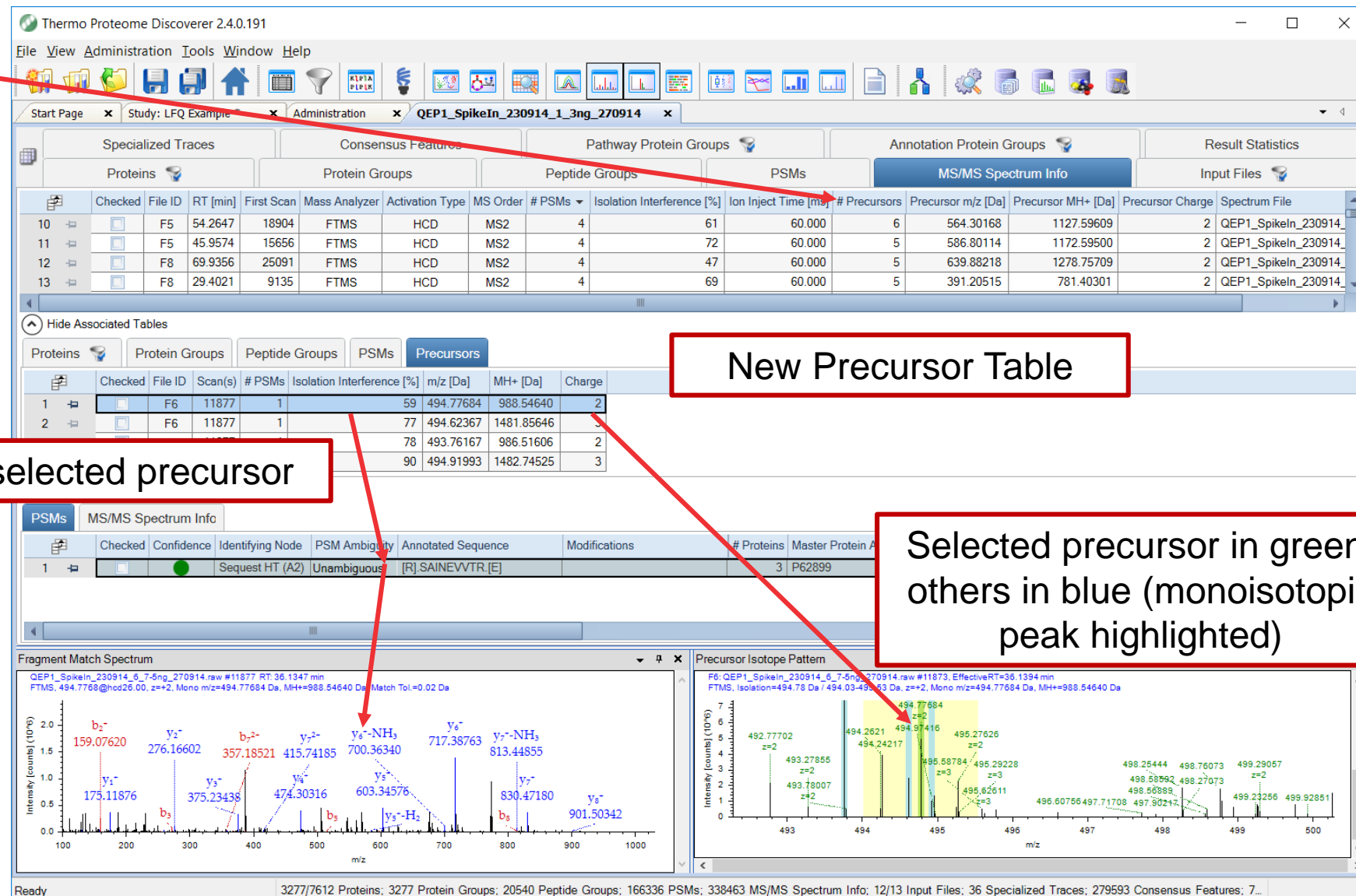
Proteome Discoverer 2.4 – Precursor Detector

- Spectrum Selector by default chooses isotopic cluster selected by instrument
- **Precursor Detector** node
 - Detects other isotopic clusters with peaks within the precursor isolation window
 - Duplicates MS/MS spectrum for each new precursor mass
 - Only input parameter: Input S/N
- Search engine can identify multiple peptides in the same MS/MS spectrum.
- Works **only** with Sequest HT and MSPepSearch
- Can increase proteins IDs up to 10-20%, peptide group IDs up to 15-30%



Precursor Detector Node produces new Precursor Table

New # Precursors Column in MS/MS Spectrum Info table



Link to PSM for selected precursor

New Precursor Table

Selected precursor in green, others in blue (monoisotopic peak highlighted)

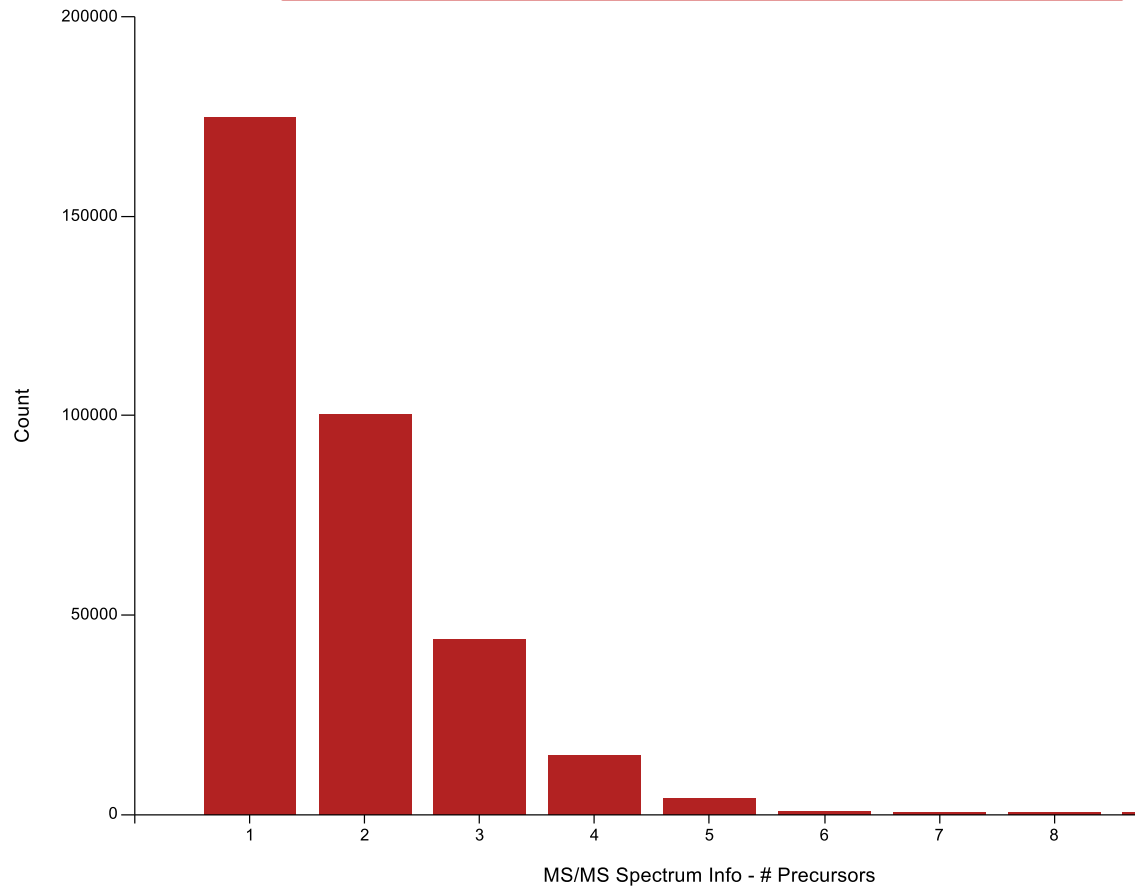
MS/MS spectrum with annotations for selected PSM

18

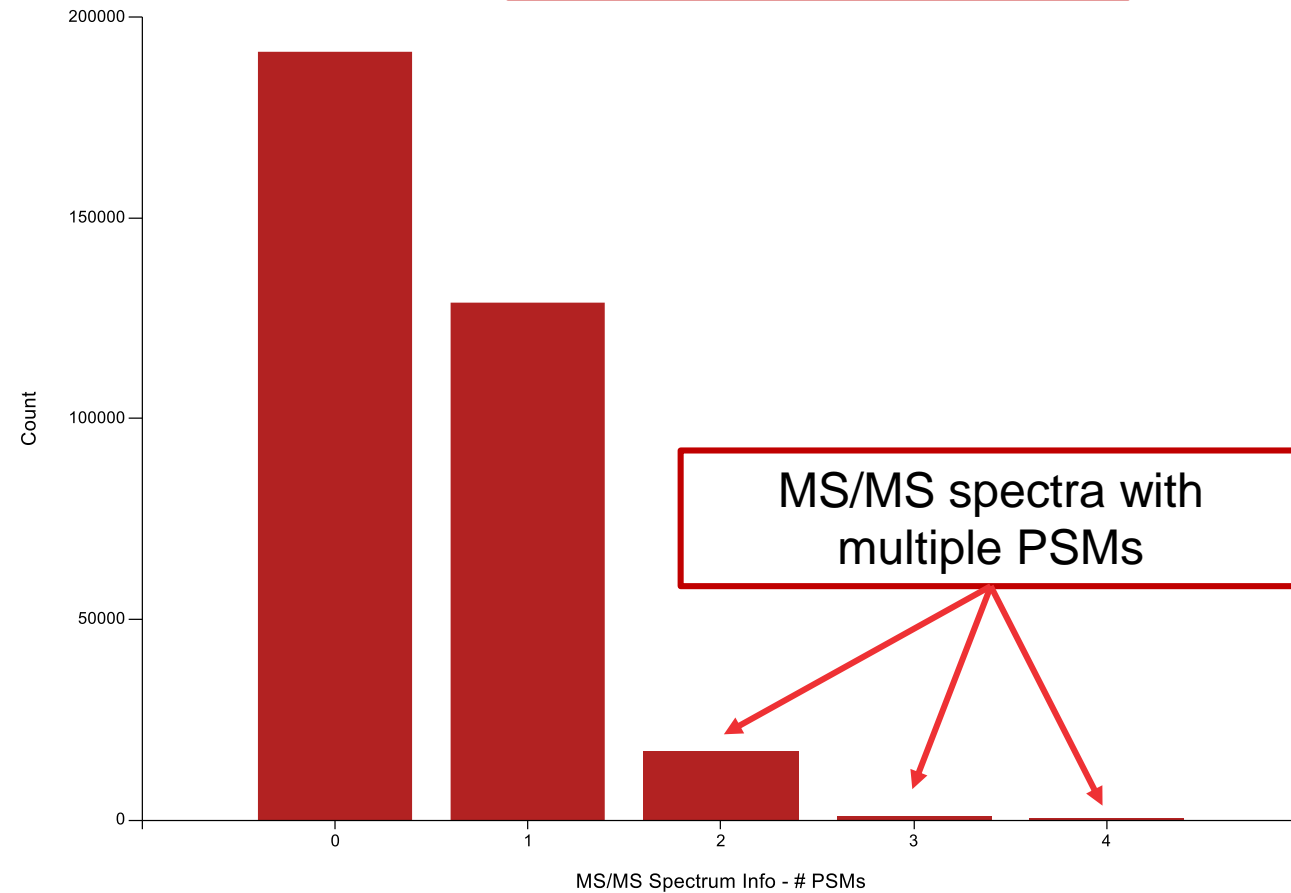
ThermoFisher
SCIENTIFIC

Precursors and **Sequest** HT PSMs per MS/MS spectrum using Precursor Detector node

of Precursors Per Spectrum
produced by Precursor Detector



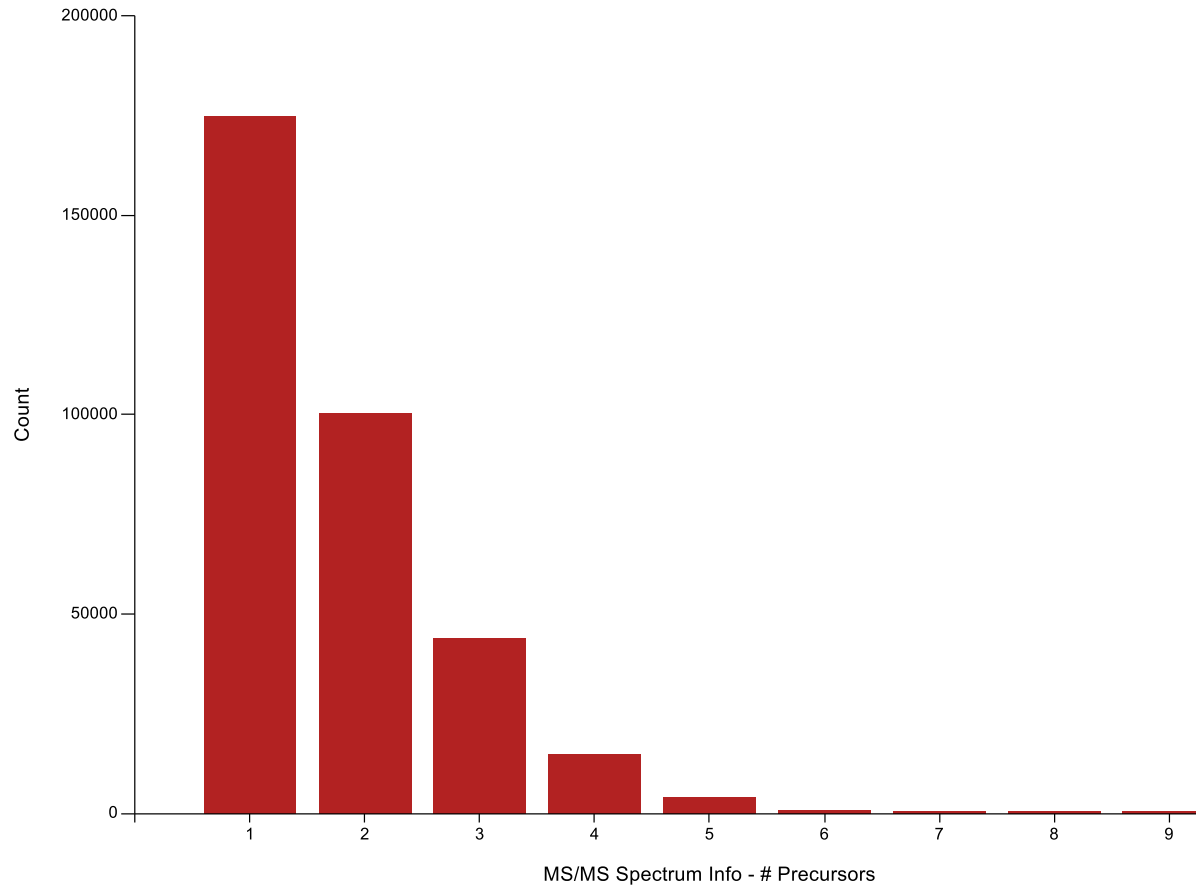
IDs per MS/MS spectrum
for Sequest HT search



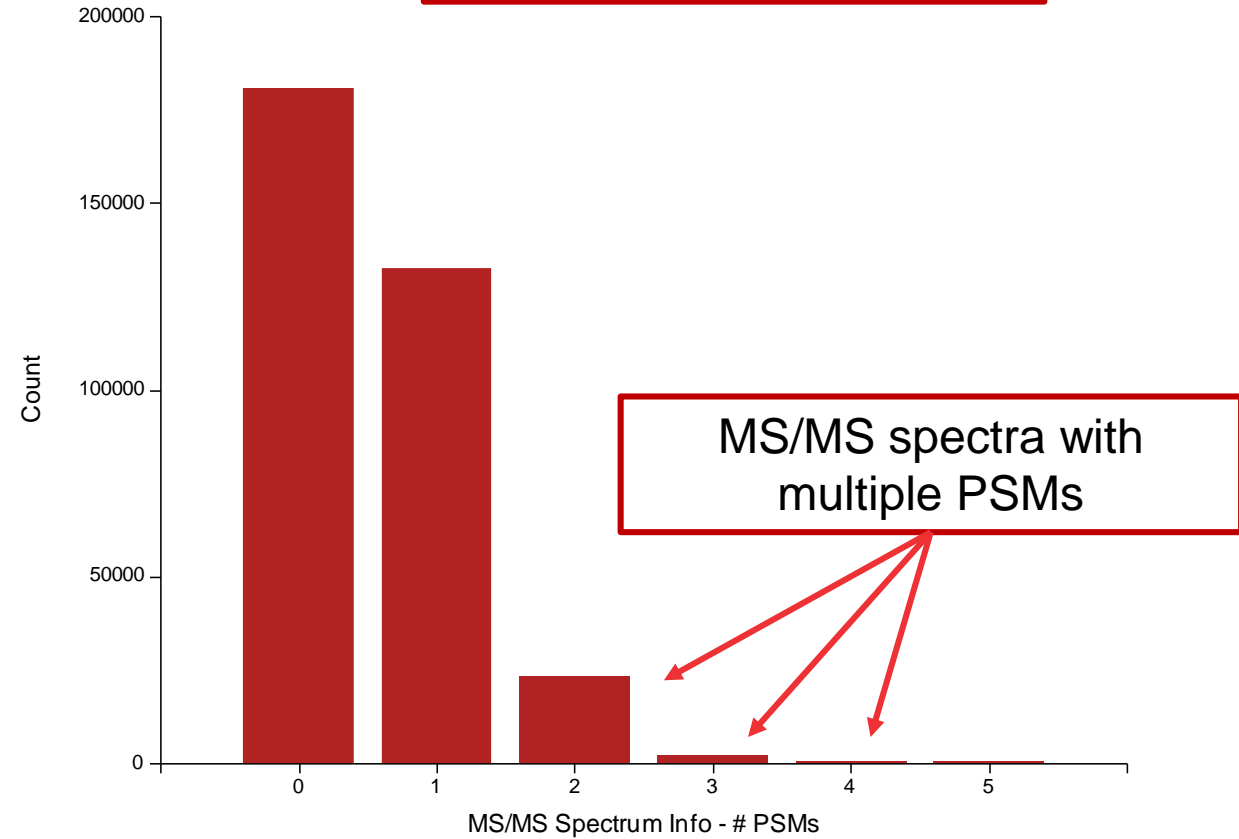
Data from Exercise 3 in Familiarization Guide

Precursors and **MS**Pep**Search** PSMs per MS/MS spectrum using Precursor Detector node

of Precursors Per Spectrum
produced by Precursor Detector



IDs per MS/MS
MSPep**Search** search
(NIST OT HCD library)



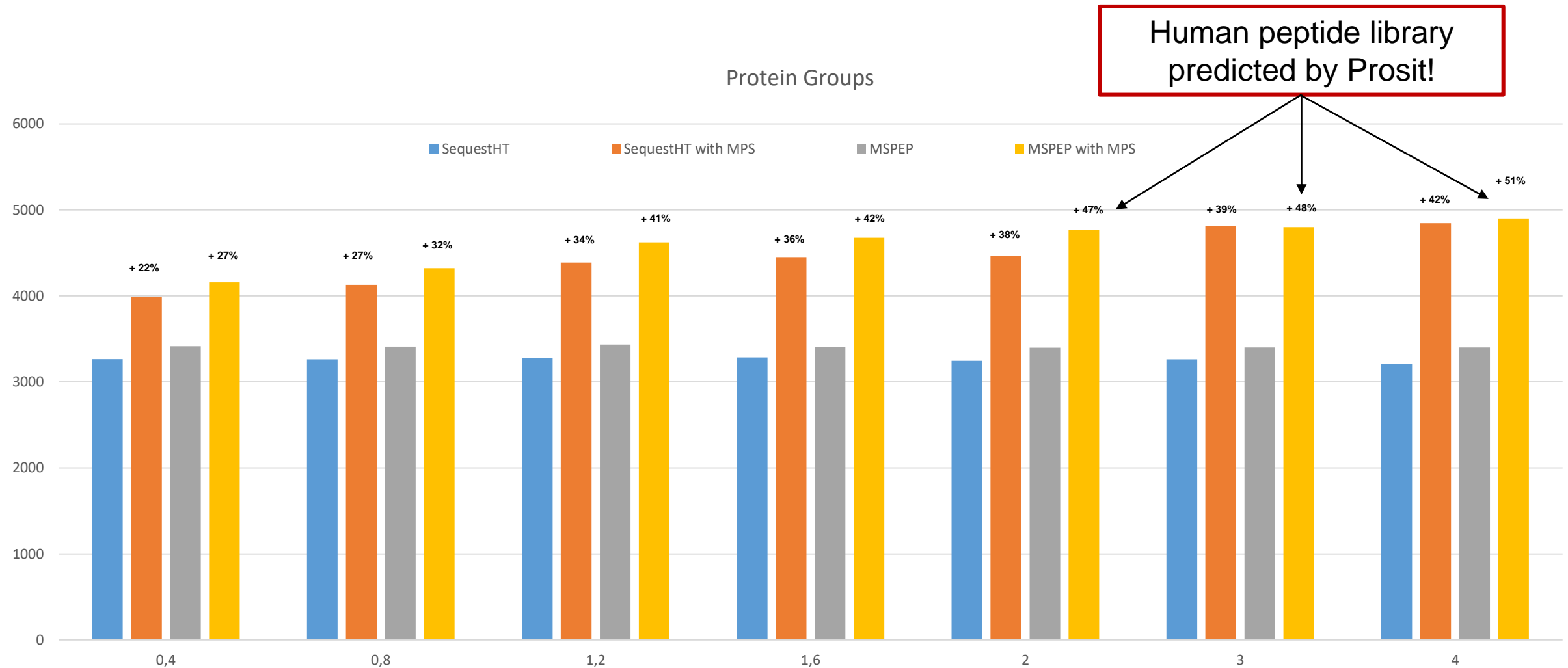
Data from Exercise 3 in Familiarization Guide

LFQ Example data Familiarization Exercise 3 – With and Without Precursor Detector

	Sequest HT + Percolator			MSPepSearch (NIST OT HCD library) + Percolator		
	PD 2.4 with Precursor PD 2.3 Detector Improvement			PD 2.4 with Precursor PD 2.3 Detector Improvement		
PSMs	129705	166336	+28%	135710	185457	+37%
Peptide Groups	17288	20540	+18%	17771	22500	+26%
Quantified Peptides	16250	19122	+18%	16659	20827	+25%
Proteins	2931	3277	+12%	2797	3233	+16%
Quantified Proteins	2757	3062	+11%	2607	3043	+17%

- Without Precursor Detector, Sequest HT and MSPepSearch perform similarly
- With Precursor Detector, MSPepSearch identifies **~10% more unique peptides** than Sequest HT.
- MSPepSearch requires a comprehensive spectral library, which is currently only available for the unlabeled human proteome (NIST OT HCD)

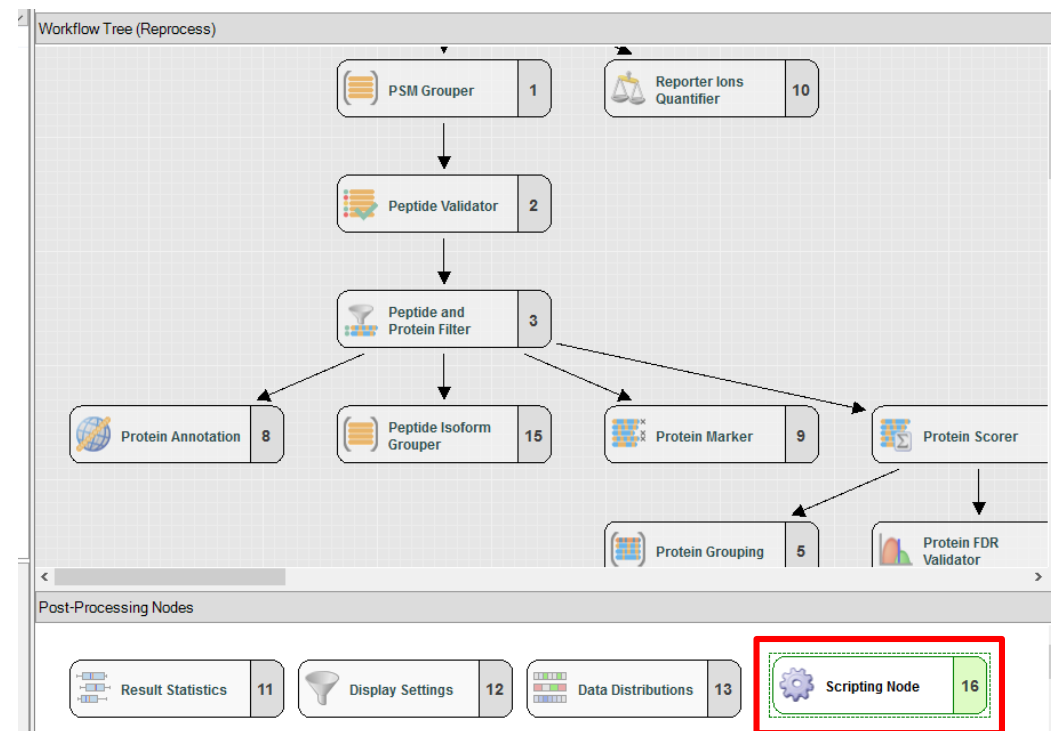
Figure from Bernard Delanghe's ASMS 2019 poster (MP 414)



Use of Precursor Detector (MPS) leads to increased IDs with wider isolation windows!

Proteome Discoverer 2.4 – Scripting Node

- **New Scripting Node** in Proteome Discoverer 2.4 and Thermo Scientific™ Compound Discoverer™ 3.1
- Enable users to incorporate R and Python scripts (or any other executable) to perform custom data analysis
- Installed as a Post-Processing node
- Can also register custom scripting nodes as a Processing, Consensus, or Post-Processing node
- Can access any information from any visible table
 - No access to MS/MS spectra or study information



New Scripting Node
available by default as a
Post Processing node

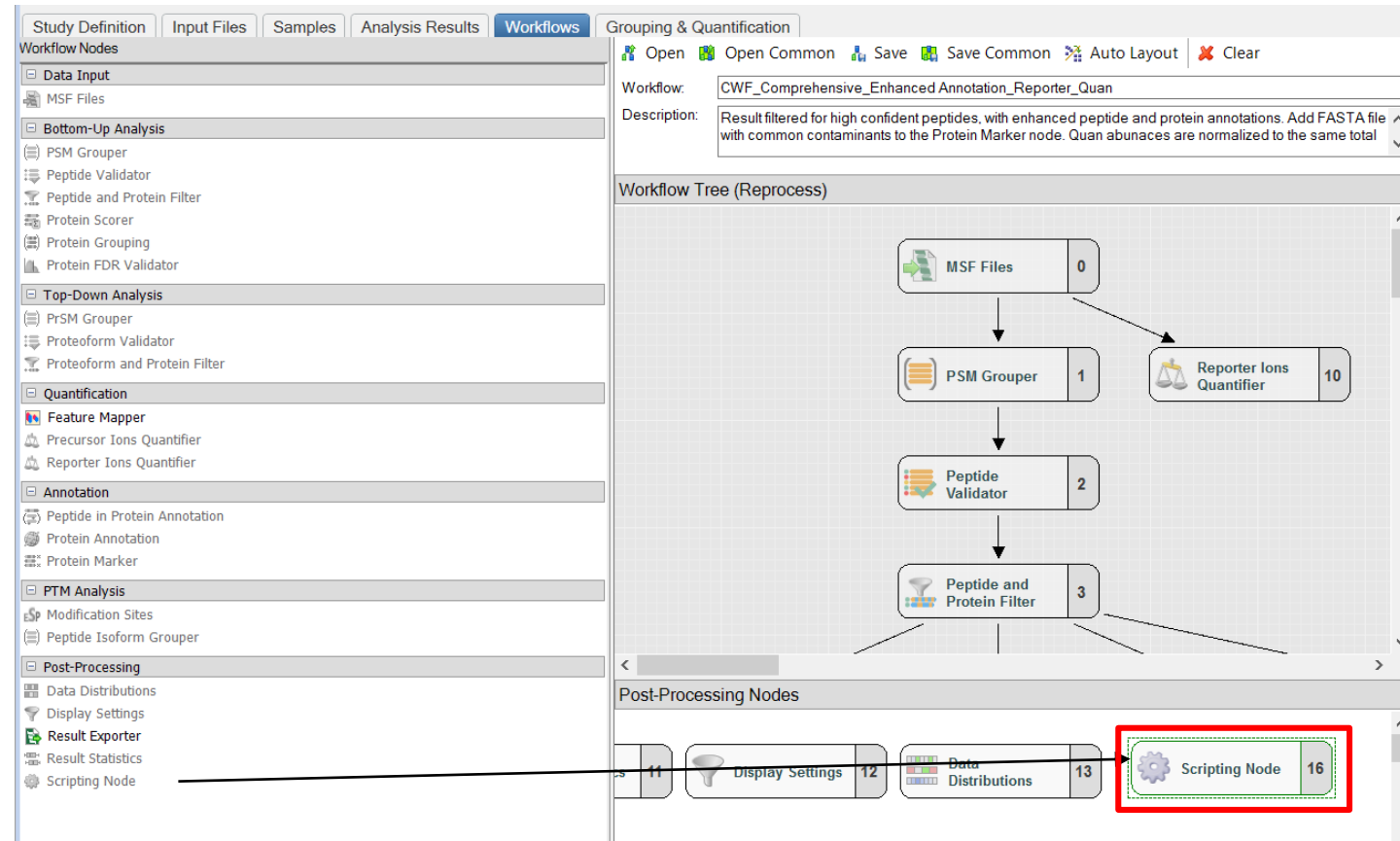
What about customers who know R, Python, Java, etc?

- In past releases, it is possible to create nodes in C# that call R, Python scripts or other executables
- These required learning C# .NET and using Visual Studio in past releases.
- **New Scripting Node** in Proteome Discoverer 2.4 and Thermo Scientific™ Compound Discoverer™ 3.1
- Enable users to incorporate R and Python scripts to perform their own type of data analysis
- Scripts have access to data from any of the result tables

- Caveats:
 - Scripting node for PD 2.4 does not have access to:
 - The raw mass spectra
 - The study information

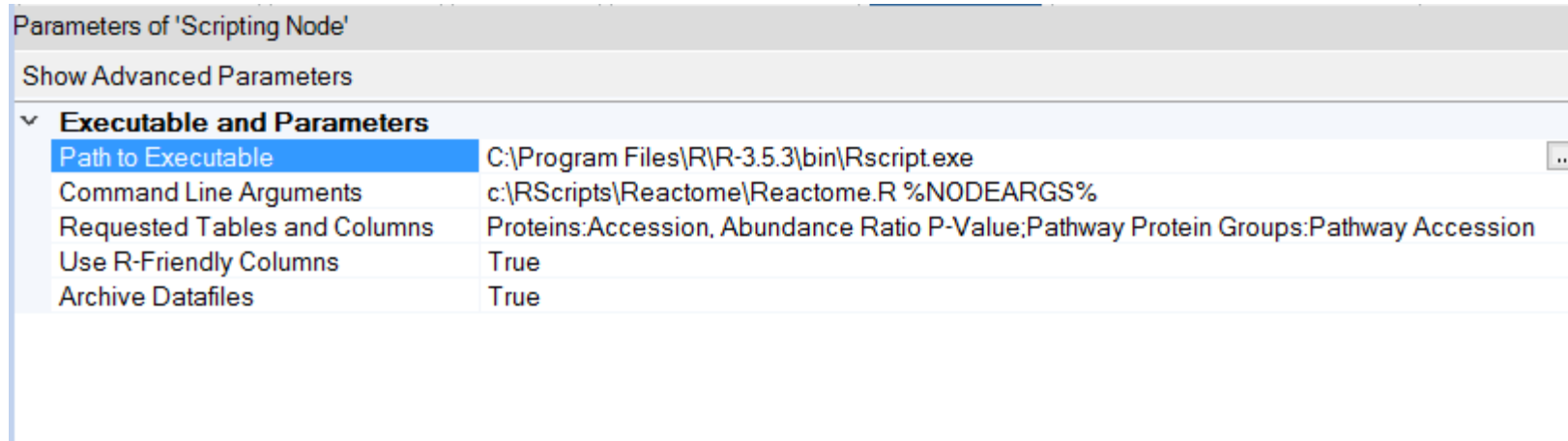
Scripting node – primary use will be as a post-processing node

- Scripting node in Proteome Discoverer 2.4 is shown as a Post-Processing node:



- It is possible to register a scripting node as a Processing or Consensus workflow node

Scripting Node parameters



Parameters of 'Scripting Node'

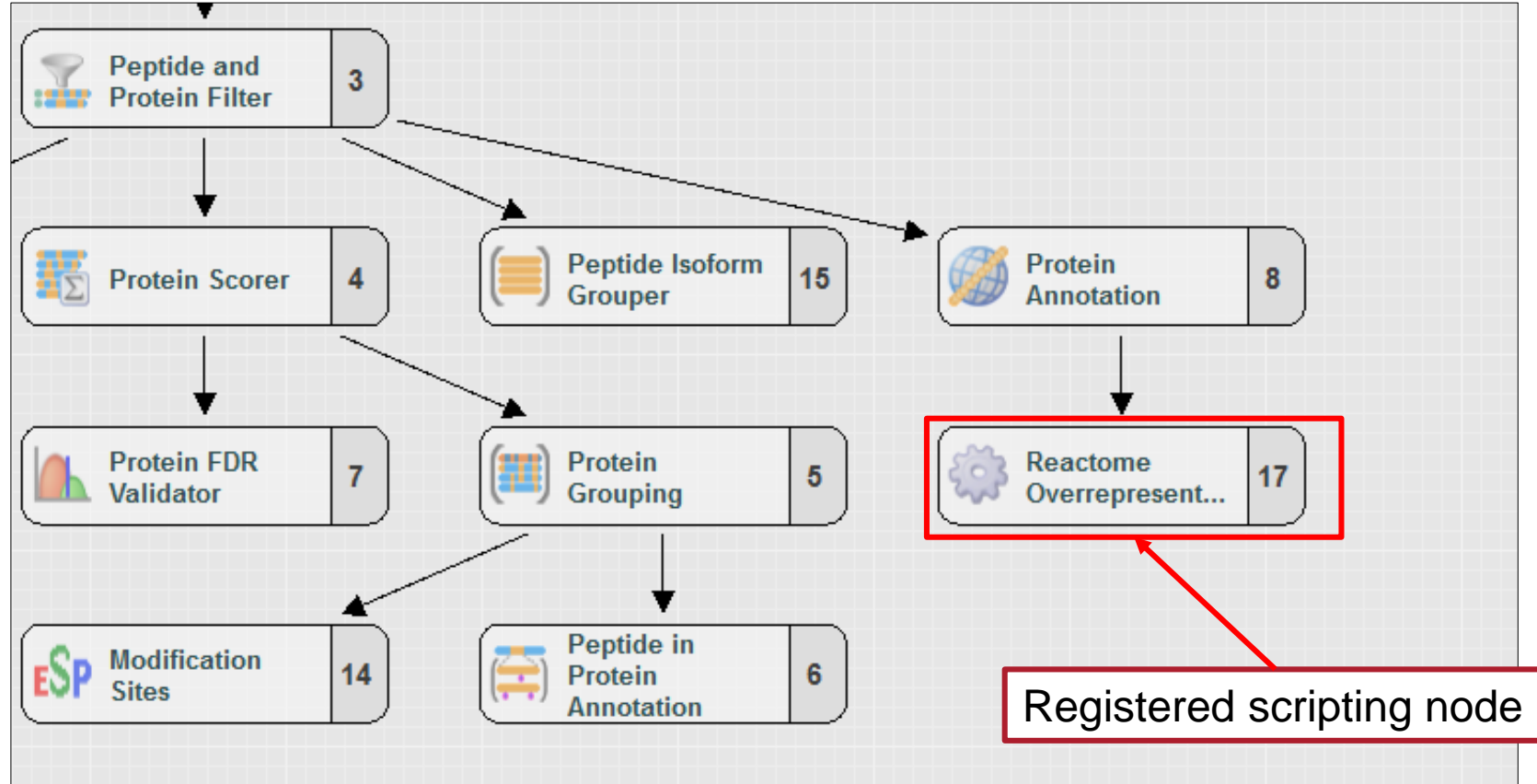
Show Advanced Parameters

▼ **Executable and Parameters**

Path to Executable	C:\Program Files\R\R-3.5.3\bin\Rscript.exe
Command Line Arguments	c:\RScripts\Reactome\Reactome.R %NODEARGS%
Requested Tables and Columns	Proteins:Accession, Abundance Ratio P-Value;Pathway Protein Groups:Pathway Accession
Use R-Friendly Columns	True
Archive Datafiles	True

- **Path to Executable** – usually the executable for the scripting language (e.g. Rscript.exe) but any executable will work
- **Command Line Arguments** - Name of the script + what will be the link to the input data for the script
- **Requested Tables and Columns** –PD result tables and columns to be sent to the script
- **Use R-Friendly Columns** – removes special characters (e.g. #, /, (), []) from column headers
- **Archive data files** – saves a copy of files used by and produced by scripting node

Registering the scripting node for use in Processing or Consensus Workflows



Registered scripting nodes are easier to transfer to other Proteome Discoverer installations

Scripting node in action – messages written by script appear in the Run Queue

Thermo Proteome Discoverer 2.4.0.173

File View Administration Tools Window Help

Start Page x Study: TMT Phosphopeptide Example x Administration x

Process Management

Job Queue

Content Management

- FASTA Files
- FASTA Indexes
- FASTA Parsing Rules
- Spectral Libraries
- Chemical Modifications
- Cleavage Reagents
- Annotation Aspects
- Quantification Methods

License Management

Licenses

Configuration

- Processing Settings
 - Annotation Server
 - Display Settings
 - IMP-ptmRS
 - Mascot
 - Minora Feature Detector
 - MSF Files
 - MSPeSearch
 - Sequest
 - Spectrum Files RC
 - Spectrum Libraries
- Server Settings
 - Temporary Files
 - Parallel Job Execution
 - Discoverer Daemon
 - FASTA Indexes

Pause Resume Abort Remove Refresh Open Results Open Study Display Verbose Messages

Job Queue:

Execution State	Details	Progress	Type	Name	Submitted at	Study	Data Source	Description
	Time	Processing Node	Level	Message				
	2:53 PM	(16): Scripting Node	Info	Storing data for table Pathway Protein Groups				
	2:53 PM	(16): Scripting Node	Info	The executable finished successfully in 35.2 s.				
	2:53 PM	(16): Scripting Node	Warni...	6.72% of input gene IDs are fail to map...				
	2:53 PM	(16): Scripting Node	Warni...	2: In bitr(geneList, fromType = "UNIPROT", toType = "ENTREZID", OrgDb = "org.Hs.eg.db") :				
	2:53 PM	(16): Scripting Node	Warni...	6.03% of input gene IDs are fail to map...				
	2:53 PM	(16): Scripting Node	Warni...	1: In bitr(geneList, fromType = "UNIPROT", toType = "ENTREZID", OrgDb = "org.Hs.eg.db") :				
	2:53 PM	(16): Scripting Node	Warni...	Warning messages:				
	2:53 PM	(16): Scripting Node	Info	[1] "Activation of DNA fragmentation factor"				
	2:53 PM	(16): Scripting Node	Info	2160 Levels: hsa00010 hsa00020 hsa00030 hsa00040 hsa00051 hsa00052 ... WP98				
	2:53 PM	(16): Scripting Node	Info	[1] R-HSA-211227				
	2:53 PM	(16): Scripting Node	Info	[1] "Found overrepresented pathway in PD results"				
	2:53 PM	(16): Scripting Node	Info	[1] "Apoptosis induced DNA fragmentation"				
	2:53 PM	(16): Scripting Node	Info	2160 Levels: hsa00010 hsa00020 hsa00030 hsa00040				
	2:53 PM	(16): Scripting Node	Info	[1] R-HSA-140342				
	2:53 PM	(16): Scripting Node	Info	[1] "Found overrepresented pathway in PD results"				
	2:53 PM	(16): Scripting Node	Info	[1] "Formation of Senescence-Associated Heterochromatin Foci (SAHF)"				
	2:53 PM	(16): Scripting Node	Info	2160 Levels: hsa00010 hsa00020 hsa00030 hsa00040 hsa00051 hsa00052 ... WP98				
	2:53 PM	(16): Scripting Node	Info	[1] R-HSA-2559584				
	2:53 PM	(16): Scripting Node	Info	[1] "Found overrepresented pathway in PD results"				
	2:53 PM	(16): Scripting Node	Info	[1] "mTORC1-mediated signalling"				
	2:53 PM	(16): Scripting Node	Info	2160 Levels: hsa00010 hsa00020 hsa00030 hsa00040 hsa00051 hsa00052 ... WP98				
	2:53 PM	(16): Scripting Node	Info	[1] R-HSA-166208				
	2:53 PM	(16): Scripting Node	Info	[1] "Found overrepresented pathway in PD results"				
	2:53 PM	(16): Scripting Node	Info	[1] "Deadenylation-dependent mRNA decay"				
	2:53 PM	(16): Scripting Node	Info	2160 Levels: hsa00010 hsa00020 hsa00030 hsa00040 hsa00051 hsa00052 ... WP98				
	2:53 PM	(16): Scripting Node	Info	[1] R-HSA-429914				
	2:53 PM	(16): Scripting Node	Info	[1] "Found overrepresented pathway in PD results"				
	2:53 PM	(16): Scripting Node	Info	[1] "Apoptotic execution phase"				
	2:53 PM	(16): Scripting Node	Info	2160 Levels: hsa00010 hsa00020 hsa00030 hsa00040 hsa00051 hsa00052 ... WP98				
	2:53 PM	(16): Scripting Node	Info	[1] R-HSA-75153				
	2:53 PM	(16): Scripting Node	Info	[1] "Found overrepresented pathway in PD results"				

Ready

Output from R script shown in the Run Queue

Scripting node can add new columns and tables to the result

Checked	Group ID	Pathway Accession	Pathway Level	Pathway Description	Pathway Source	# Master Proteins	# Proteins	-10LogPValue Insulin Con	-10LogPValue IGF-1 Control
<input checked="" type="checkbox"/>	1045	R-HSA-1266695	Leaf	Interleukin-7 signaling	Reactome	7	11	163.43770	
<input checked="" type="checkbox"/>	35	R-HSA-427389	Leaf	ERCC6 (CSB) and EHMT2 (G9a) positively regulate	Reactome	12	36	138.30881	
<input checked="" type="checkbox"/>	930	R-HSA-3214842	Leaf	HDMS demethylate histones	Reactome	12	14	124.73350	
<input checked="" type="checkbox"/>	369	R-HSA-3214815	Leaf	HDACs deacetylate histones	Reactome	22	14	123.03750	
<input checked="" type="checkbox"/>	379	R-HSA-5625886	Leaf	Activated PKN1 stimulates transcription of AR (and	Reactome	8	26	121.05480	
<input checked="" type="checkbox"/>	381	R-HSA-2299718	Leaf	Condensation of Prophase Chromosomes	Reactome	9	30	114.76200	
<input checked="" type="checkbox"/>	384	R-HSA-5250924	Leaf	B-WICH complex positively regulates RNA expres	Reactome	16	35	113.49040	
<input checked="" type="checkbox"/>	385	R-HSA-73728	Leaf	RNA Polymerase I Promoter binding	Reactome	5	23	111.98370	
<input checked="" type="checkbox"/>	372	R-HSA-5334118	Leaf	DNA methylation	Reactome	6	25	110.14020	

2 new columns:
-10logP value for overrepresentation for
proteins from insulin or IGF-1 stimulation

Checked	Group ID	Pathway Accession	Pathway Level	Pathway Description	Pathway Source	# Master Proteins	# Proteins	-10LogPValue Insulin Con	-10LogPValue IGF-1 Control
<input checked="" type="checkbox"/>	1101	R-HSA-1912406	Leaf	mTORC1 transcription and translation	Reactome	1	1	180	
<input checked="" type="checkbox"/>	374	R-HSA-977225	Leaf	Amyloid fiber formation	Reactome	1	1	80	
<input checked="" type="checkbox"/>	271	R-HSA-5617472	Leaf	Activation of anterior HOX genes in hindbrain develo	Reactome	1	1	266	
<input checked="" type="checkbox"/>	375	R-HSA-8939236	Leaf	RUNX1 regulates transcription of genes involved in c	Reactome	1	1	120	
<input checked="" type="checkbox"/>	376	R-HSA-3214847	Leaf	HATs acetylate histones	Reactome	1	1	133	
<input checked="" type="checkbox"/>	1648	R-HSA-211227	Leaf	Activation of DNA fragmentation factor	Reactome	1	1	66	
<input checked="" type="checkbox"/>	204	R-HSA-5674400	Leaf	Constitutive Signaling by AKT1 E17K in Cancer	Reactome	1	1	93	
<input checked="" type="checkbox"/>	1646	R-HSA-2559584	Leaf	Formation of Senescence-Associated Heterochroma	Reactome	1	1	169	
<input checked="" type="checkbox"/>	564	R-HSA-72163	Leaf	mRNA Splicing - Major Pathway	Reactome	1	1	4202	
<input checked="" type="checkbox"/>	1473	R-HSA-166208	Leaf	mTORC1-mediated signaling	Reactome	1	1	6	
<input checked="" type="checkbox"/>	187	R-HSA-198323	Leaf	AKT phosphorylates targets in the cytosol	Reactome	1	1	49	
<input checked="" type="checkbox"/>	281	R-HSA-72187	Leaf	mRNA 3'-end processing	Reactome	1	1	75	
<input checked="" type="checkbox"/>	212	R-HSA-389357	Leaf	CD28 dependent PI3K/Akt signaling	Reactome	1	1	158	
<input checked="" type="checkbox"/>	778	R-HSA-106688	Leaf	Release of Growth Transcript in the Transition B	Reactome	1	1		

New table with links to existing tables

ProteinStartsWith	NumberOfTimes	ProteinStartsWithLetter
1	180	A
2	80	B
3	266	C
4	120	D
5	133	E
6	66	F
7	93	G
8	169	H
9	4202	I
10	6	J
11	49	K
12	75	L
13	158	M

ProteinStartsWith
36 items shown (0 filtered out)

Hide Associated Tables

Checked	Protein F	Master	Accession	Description	Exp. q-val	Sum PEP Score	Coverage [%]	Sequence Coverage	# Peptides	# Isoforms	# PSMs	# Unique Peptides	# AAs	MW [kDa]	calc. pI
<input checked="" type="checkbox"/>	High	✓	Q9H4G0-1	Band 4.1-like protein 1 [OS=Homo sapiens]	0.000	85.699	17%		19	32	34	19	881	98.4	5.62
<input checked="" type="checkbox"/>	High	✓	Q9NYF8-1	Bcl-2-associated transcription factor 1 [OS=Homo sapiens]	0.000	83.189	18%		18	21	22	7	920	106.1	9.98
<input checked="" type="checkbox"/>	High	✓	Q9BRD0	BUD13 homolog [OS=Homo sapiens]	0.000	56.150	22%		17	21	24	17	619	70.5	9.86
<input checked="" type="checkbox"/>	High	✓	Q8NFC6	Biorientation of chromosomes in cell division prote	0.000	46.794	5%		13	13	14	13	3051	330.3	5.08
<input checked="" type="checkbox"/>	High	✓	Q9UHR4	Brain-specific angiogenesis inhibitor 1-associated	0.000	41.677	16%		4	10	11	4	511	56.8	8.68
<input checked="" type="checkbox"/>	High	✓	P35612-1	Beta-adducin [OS=Homo sapiens]	0.000	40.442	6%		5	11	11	5	726	80.8	5.92
<input checked="" type="checkbox"/>	High	✓	O43491-1	band 4.1-like protein 2 [OS=Homo sapiens]	0.000	32.169	7%		7	10	10	7	1005	112.5	5.44
<input checked="" type="checkbox"/>	High	✓	Q13425-1	Beta-2-syntrophin [OS=Homo sapiens]	0.000	24.952	13%		4	8	8	4	540	57.9	8.82
<input checked="" type="checkbox"/>	High	✓	Q9UIF9-1	Bromodomain adjacent to zinc finger domain prote	0.000	22.719	4%		5	9	9	5	1905	211.1	6.64
<input checked="" type="checkbox"/>	High	✓	Q95817	BAG family class B transcription factor 3 [OS=Homo sapiens]	0.000	21.031	12%		5	8	8	5	575	61.6	6.05

Show Associated Tables

Ready 2930/7180 Proteins; 2930 Protein Groups; 8992 Peptide Groups; 9468 Peptide Isoforms; 8656 Modification Sites; 10160 PSMs; 52022 MS/MS Spectrum Info; 41294 Quan Spectra; 1/2 Inpu...

New Scripting Node section in the Help

- New scripting node section in the Help explains how to create nodes from scratch
- PD media includes “User Scripts” folder with example scripts in R, Python, and Java with tutorials on how to run

Java	7/29/2019 4:37 AM	File folder	
Python	7/29/2019 4:37 AM	File folder	
R	7/29/2019 4:38 AM	File folder	
Study	7/29/2019 4:48 AM	File folder	
.project	7/16/2019 1:16 AM	PROJECT File	1 KB
README.txt	7/16/2019 1:23 AM	Text Document	1 KB
Scripting Node Custom Script Examples.docx	7/28/2019 10:57 PM	Microsoft Word Document	80 KB
selftest.bat	7/16/2019 1:16 AM	Windows Batch File	1 KB

The screenshot shows the Proteome Discoverer application window. The left sidebar contains a 'Contents' pane with a tree view of the help topics. The main pane displays the 'Scripting Node Overview' page. The page title is 'Scripting Node Overview'. The text describes the Scripting Node as a customizable node for running executable files. It explains the data exchange mechanism between the application and the executable, which uses JSON files for structure and tab-separated text files for data. It also provides instructions on how to create a script for the Scripting Node, mentioning supported programming languages like Python, R, C#, and C++. A list of bullet points includes 'Use as a Post Processing Node (Default)' and 'Register a standalone Processing or Consensus node'. The page concludes with information on how to register a custom workflow node and a note that the Scripting Node cannot access MS/MS spectra or study information.

Scripting Node Overview

The Scripting Node is a customizable node that allows you to run an executable file as the final node in a workflow and to incorporate the result data into the application's result file.

The data exchange between the Proteome Discoverer application and the executable relies on a few mechanisms that do not need .NET code because the executable does not use the application API to read and write data. Instead, the user declares the columns for the tables from the result file to be exported into one or more tab-separated text files along with a JSON file describing the structure and location of those files. The user-defined script or executable should subsequently read these files, perform the user-defined type of analysis, and then optionally write out tab-separated text files containing information with another JSON file describing the structure of new tables and/or columns to be imported into the Proteome Discoverer results.

To create a script for the Scripting Node, programmers can use any programming language that supports running code from the command line, for example, python, R, C#, C++, and so on.

- Use as a Post Processing Node (Default)

For an example, see [this figure](#).

- Register a standalone Processing or Consensus node that is preconfigured to an executable and can be given to another Proteome Discoverer user.

For information about how to register a custom workflow node, see [Installing or Updating a Scripting Workflow Node](#).

The Scripting Node cannot node access MS/MS spectra or study information. In general, the Scripting Node can only act upon viewable columns from result file.

To integrate a search engine, a statistic algorithm, or a workflow that requires deeper access to the application code, you must program in C#. Contact pd.support@thermofisher.com for more information.

Implementing a Generic Scripting Node to a Standard Proteomics Workflow Processing Software

Frank Berg¹, Carmen Paschke¹, Kai Fritzemeier¹, Pedro Navarro¹, Torsten Ueckert¹, David Horn², Bernard Delanghe¹, ¹Thermo Fisher Scientific (Bremen) GmbH, Bremen, Germany; ²Thermo Fisher, San Jose, CA

ABSTRACT

Purpose: Implement an easy-to-use mechanism to enrich workflows with results of non-C# user algorithms in Thermo Scientific™ Proteome Discoverer™ framework.

Methods: Creating a family of preconfigured nodes as well as general mechanisms that integrate the calculation results of arbitrary external executables or scripts into Thermo Scientific™ Proteome Discoverer™ 2.4 software result files.

Results: We show by means of a custom R script that employs the widely used limma package [1] the integration of its results into Proteome Discoverer 2.4 software and use the additional statistical results of quantification data to compare them to the built-in Proteome Discoverer 2.4 software statistics algorithms. For this we use the rich set of plots and table presentations in Proteome Discoverer 2.4 software as well as R Studio.

INTRODUCTION

Proteome Discoverer software offers flexible analysis of proteomics mass spectrometry measurement data. Analyses are done by customizable workflows of configurable nodes that perform workflow subtasks, e.g., peptide identification, statistical validation or consolidation of protein findings. As of now, custom nodes may be implemented by third parties using a .NET programming language (typically C#) against the richly featured Proteome Discoverer API, thus extending the set of factory-provided analysis features. However, for rapid prototyping in context of, e.g., academic teaching or research contexts with compact and fast changing algorithmic ideas written in popular scripting languages like R or python this poses a certain cannon-on-sparrow situation.

Here we present a node family for PD that allows integrating arbitrary executables or scripts into an analysis workflow by using pre-implemented scripting nodes that adhere to a predefined data exchange protocol for external executables, thus providing an easy and fast method to extend workflows with user algorithms.

MATERIALS AND METHODS

The software was implemented within the Proteome Discoverer 2.4 framework using C#

Results

In principle our implementation offers to the user two ways of using a scripting integration in Proteome Discoverer:

- Predefined post processing scripting nodes for both the consensus and the processing workflow that only need a few parameters and an external script to be ready to go.
- Registration and creation of a custom standalone-node that follows the same principles as described above but additionally involves a registration process in PD. With this it appears as an "ordinary" workflow node that can also be given away to other users in a standalone fashion.

We now describe the principal mechanisms of doing the data exchange between Proteome Discoverer and an external process as defined by our implementation in the post processing node. Further below, the mentioned registration process is outlined.

The post processing nodes involve the following basic parameters (Figure 1):



- **Path to Executable:** Location of the executable or script. If only a filename is given the system PATH environment variable is used to find it.
- **Command Line Arguments:** Any additional argument the executable needs.
- **Requested Tables and Columns:** A string that encodes all information about the data tables from the current PD result file that should be provided to the executable. The data is exported as a CSV text export.

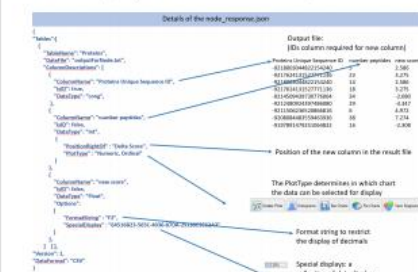
Prior to executing the script the node provides the requested data in CSV tables and additionally stores a json file named "node_args.json" (Figure 2) containing meta information about the data.



Figure 2. Example of exported Protein data from the PD result file as it is available to the scripting node.

- The following information is stored in file "node_args.json":
 - The path and name of every requested data table file
 - Type information about any column that is contained in the data table file
 - ID information, i.e., information about which columns are ID columns that are needed to insert new data into an existing table and to connect tables.

While the path information is vital to find the exported data, the type description is useful to parse the data values when importing them into a custom executable or script context.



After the external executable has calculated its results it may return data to the Proteome Discoverer result file. This is done by writing a very similar json file as "node_args.json" named "node_response.json" (Figure 3). This way the executable may perform the following actions:

- Add columns to existing tables (but not change existing columns/data)
- Add new tables with arbitrary data
- Add connections between tables. These connections may only be made between tables that are not related to each other yet.

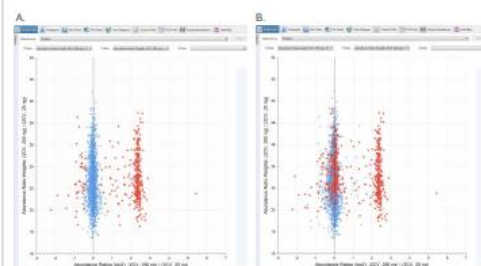
Case study: Integrating an R script for statistical analysis

To demonstrate the functionality of the scripting node, we implemented an R script using the limma package to calculate protein ratios and corresponding statistical values for a known mixture of human (HeLa) (1:1) and yeast (250ng; 25 ng) proteomes. Data reading and writing of json files (by using the RJSONIO package) are shown in the code snippets in (Figure 4 and 5). After running the scripting node, the data is available in Proteome Discoverer 2.4 and can be plotted using the plotting tools.



Figure 4. An example on how to import a table in R by using "node_args.json" properties. Here the table "Proteins" is imported, catching the index and corresponding abundances (by abundancesType) columns from the node_args.json file properties.

To compare the two different calculations we display the max. Abundance between the samples (in log scale) versus log2 of the sample ratio. The proteins with a significant q-value (< 0.01) are highlighted in red. The plot of the Proteome Discoverer 2.4 calculated values is shown on the left in Figure 6, values calculated by the R-script using limma package are plotted on the right.



Custom Deployable Scripting Nodes

As mentioned above the user can deploy a custom version of her scripting node by registering a node in a standalone fashion. Necessary steps are outlined in Figure 8. A special definition file named "node.json" (Figure 7) needs to be provided that contains the following information:

- Name, icon and target workflow (consensus or processing)
- Connection points that define where in the workflow the node can be placed
- Parameters to the node. Here standard parameters can be used that are known from ordinary PD nodes.
- All parameters needed for the scripting mentioned above.

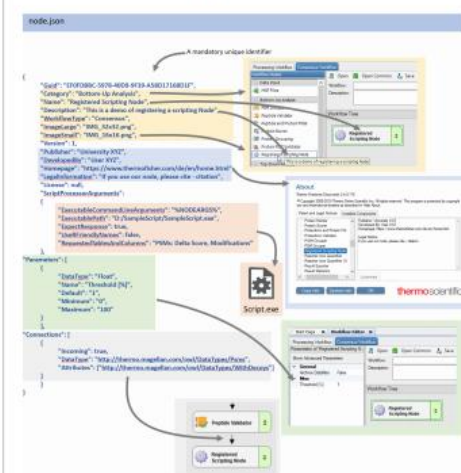


Figure 7. Example "node.json" file that defines the registration of a deployable scripting node

In the parameters section any types of parameters are available that can be used through the .NET API of Proteome Discoverer. The connections section describes a data contract that specifies the allowed connections of the scripting node to other nodes in the workflow.

Future Work

The current implementation of the scripting node mechanism involves some limitations that will be addressed in future versions of Proteome Discoverer.

1. An access to the study information that corresponds to the current analysis is not yet available.
2. Spectrum filters can not be implemented by scripting nodes.

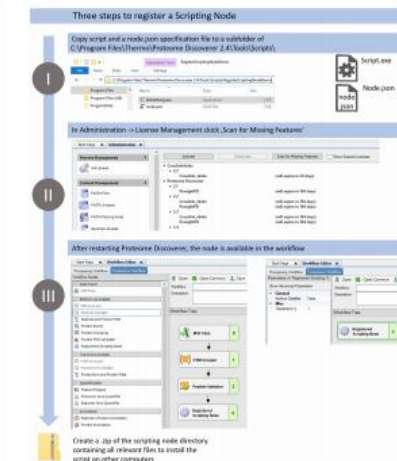


Figure 8. Three-step registration and deployment process for standalone scripting nodes in Proteome Discoverer.

CONCLUSIONS

We present a family of nodes that allow for rapid prototyping of proteomics algorithms in Proteome Discoverer™ 2.4. With these nodes the user can pass data to external executables or scripts and then import calculation results back into Proteome Discoverer. Moreover, the user can choose to define and register a deployable version of his scripting node for further distribution and sharing with collaborators. We demonstrate the usability by connecting the results of an R script that uses limma [1] to do statistics on a quantification workflow to Proteome discoverer™. For this we performed an analysis of proteomics data inspired by [2] and [3].

REFERENCES

- [1] Ritchie ME, Phipson B, Wu D, et al., limma powers differential expression analyses for RNA-seq and microarray studies. Nucleic Acids Res. 2015;43:e47.
- [2] Kai Kammers, R guide: Analysis of Cardiovascular Proteomics Data, http://www.biostat.hhs.edu/~kammers/software/CVproteomics/R_guide.html
- [3] Kai Kammers, D. Brian Foster, Ingo Ruzsinski, Analysis of Proteomic Data, In: Manual of Cardiovascular Proteomics Pages 275-292, Springer International Publishing Switzerland 2016

TRADEMARKS/LICENSES

© 2019 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.

TMTpro 16 plex support

- New TMTpro 16plex modification
- New table for TMTpro 16 plex correction factors
- New default workflows for TMTpro 16plex

Quantification Method Editor: newMethod1

Quan Channels

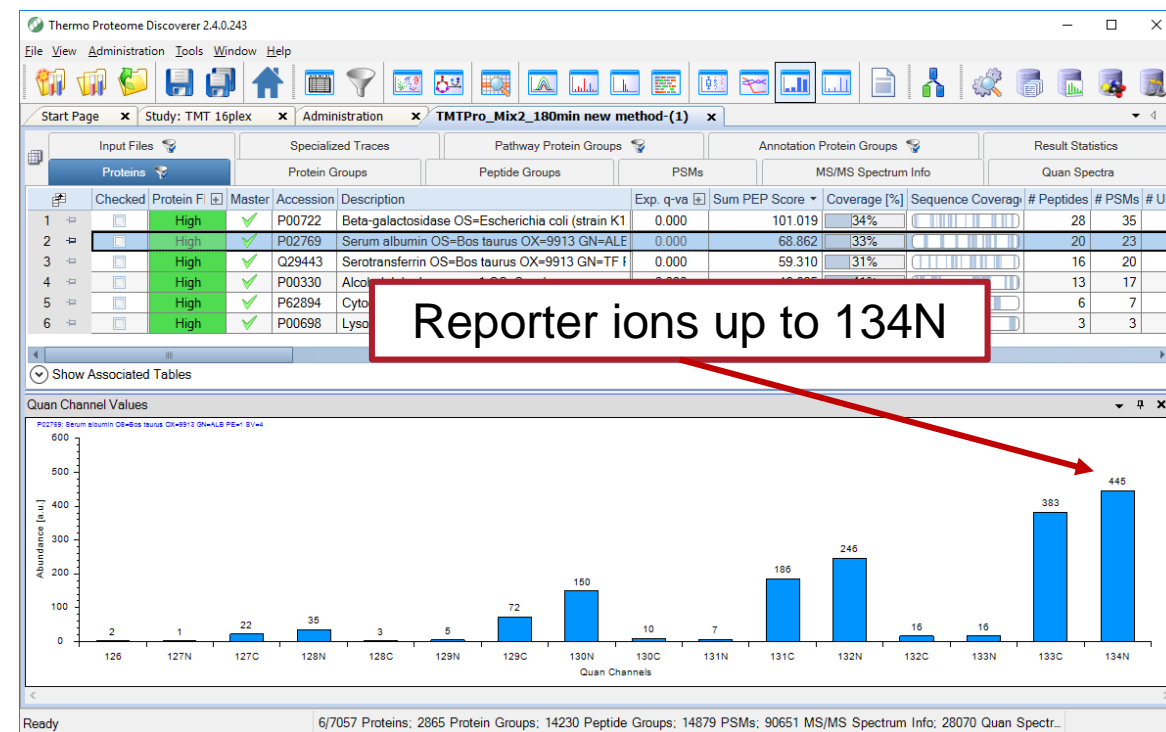
Residue Modification: TMTpro16plex / +304.207 Da K

N-Terminal Modification: TMTpro16plex / +304.207 Da

Mass Tag	Reporter Ion Mass	-2x13C	-13C-15N	-13C	-15N	Main	+15N	+13C	+15N+13C	+2x13C	Active
126	126.127726	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
127N	127.124761	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
127C	127.131081	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
128N	128.128116	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
128C	128.134436	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
129N	129.131471	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
129C	129.13779	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
130N	130.134825	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
130C	130.141145	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
131N	131.13818	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
131C	131.144499	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
132N	132.141535	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
132C	132.147855	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
133N	133.14489	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
133C	133.15121	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>
134N	134.148245	0	0	0	0	100	0	0	0	0	<input checked="" type="checkbox"/>

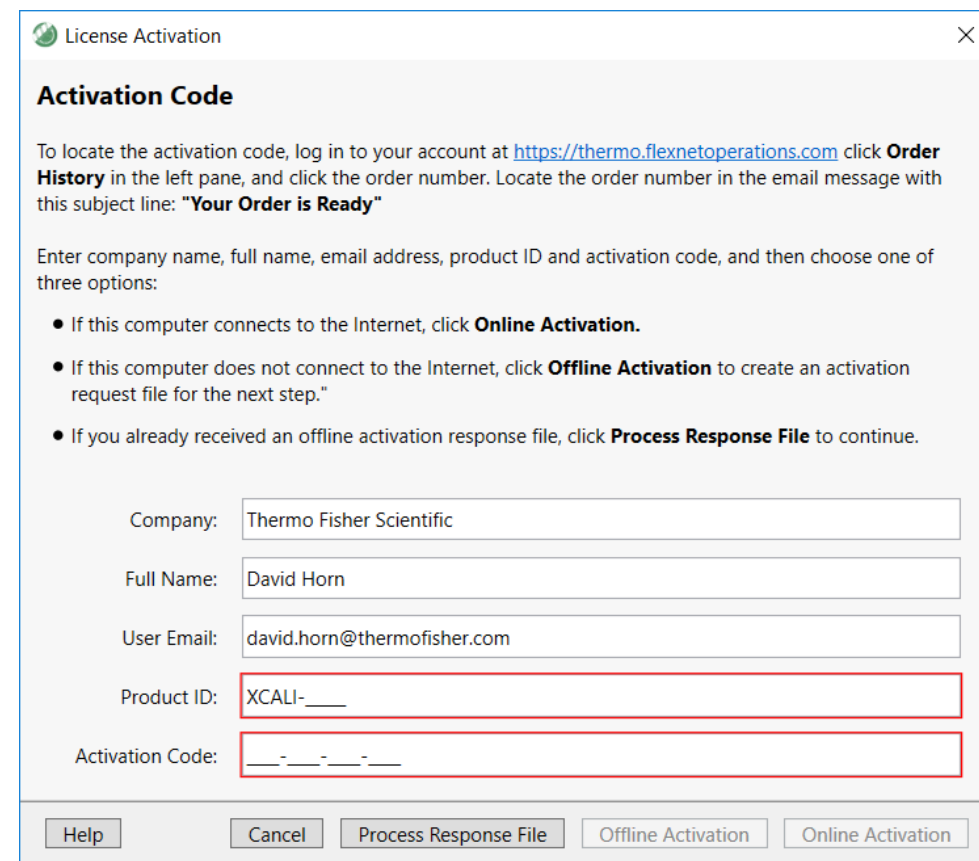
New correction factor table for TMTpro 16plex

OK Cancel Help



Flexera licensing

- Same licensing scheme as BioPharma Finder
- Activation now requires a part number and license key
- Upgrades from previous versions with active maintenance are still free (see next slide for instructions on how to upgrade)
- License keys are sent by e-mail (no more lost licenses!)
- **No more maintenance license for PD 2.4**
- **Protein Annotation service is now tied to the Base license with no expiration**



The image shows a 'License Activation' dialog box with a title bar and a close button. The main content area is titled 'Activation Code' and contains instructions on how to find the activation code. Below the instructions, there are three bullet points providing options for activation based on internet connectivity. At the bottom, there are five input fields: 'Company', 'Full Name', 'User Email', 'Product ID', and 'Activation Code'. The 'Product ID' and 'Activation Code' fields are highlighted with red borders. At the very bottom, there are five buttons: 'Help', 'Cancel', 'Process Response File', 'Offline Activation', and 'Online Activation'.

License Activation

Activation Code

To locate the activation code, log in to your account at <https://thermo.flexnetoperations.com> click **Order History** in the left pane, and click the order number. Locate the order number in the email message with this subject line: **"Your Order is Ready"**

Enter company name, full name, email address, product ID and activation code, and then choose one of three options:

- If this computer connects to the Internet, click **Online Activation**.
- If this computer does not connect to the Internet, click **Offline Activation** to create an activation request file for the next step."
- If you already received an offline activation response file, click **Process Response File** to continue.

Company:

Full Name:

User Email:

Product ID:

Activation Code:

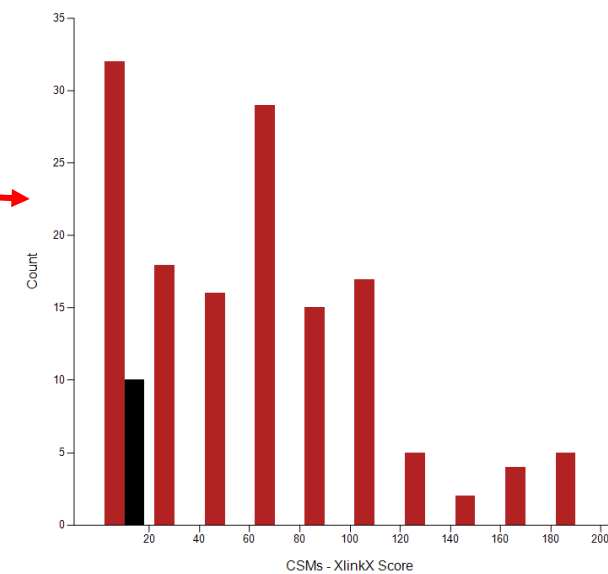
Procedure for customers with active maintenance to upgrade to 2.4

1. Install PD 2.4 on the same system as the previous version of Proteome Discoverer. PD 2.4 now automatically installs with a 60-day demo license without the need for a license key.
2. Send an e-mail to ThermoMSLicensing@thermofisher.com with the following information:
 1. Name
 2. E-mail address
 3. Institution
 4. Current active maintenance license key from previous PD installation (accessible from Administration->Manage Licenses, selecting the Discoverer Annotation license)
3. ThermoMSLicensing will send an e-mail to the e-mail address above with the license key
4. Open the Administration->Manage Licenses dialog and click Activate.
 1. For the Product ID, type XCALI-98057.
 2. For the Activation code, use the code provided by e-mail via ThermoMSLicensing.
5. If connected to the network, click Online Activation. If offline, click Offline Activation and follow the instructions to sending the license file to ThermoMSLicensing.

- MS Amanda 2.0 (IMP Vienna)
- Byonic and Preview nodes (Protein Metrics, Inc.)
- ProSightPD 3.0 (Proteinaceous, Inc.) – includes 60-day demo license
- XlinkX for PD 2.4 – includes 60-day demo license

- The Byonic and Preview nodes require the associated standalone software to be installed and licensed on the same PC. The Proteome Discoverer installation media includes the latest installers for both and a 30-day demo license key.
- The ProSightPD 3.0 nodes requires the standalone Thermo Scientific™ ProSightPC™ 4.1 software to be installed. A demo version can be downloaded from www.proteinaceous.net.

- Now have a “target-decoy” view for XlinkX score to help confirm FDR calculation
- Added compensation voltage as a column in CSM table for FAIMS data.
- Upgrades from previous versions of XlinkX are free. The license is automatically transferred from previous versions of PD.



Target CSMs (n = 143) Decoy CSMs (n = 10)

