

Proteome Discoverer Workshop

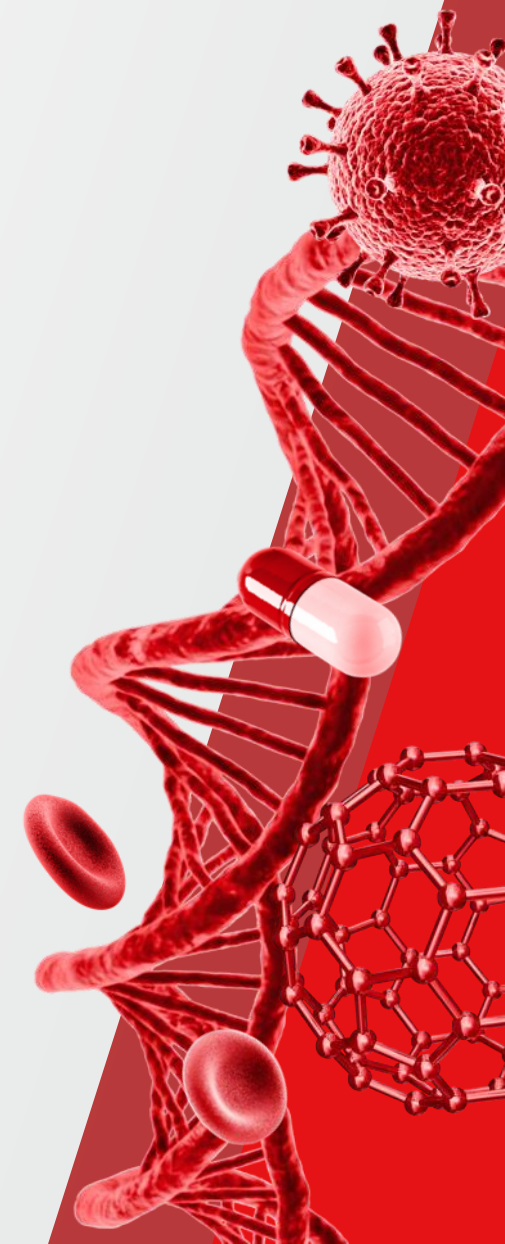
Hradec Králové

Ing. Michaela Scigelova, Ph.D.

LSMS Factory Support Groups

Bremen, Germany

 The world leader in serving science

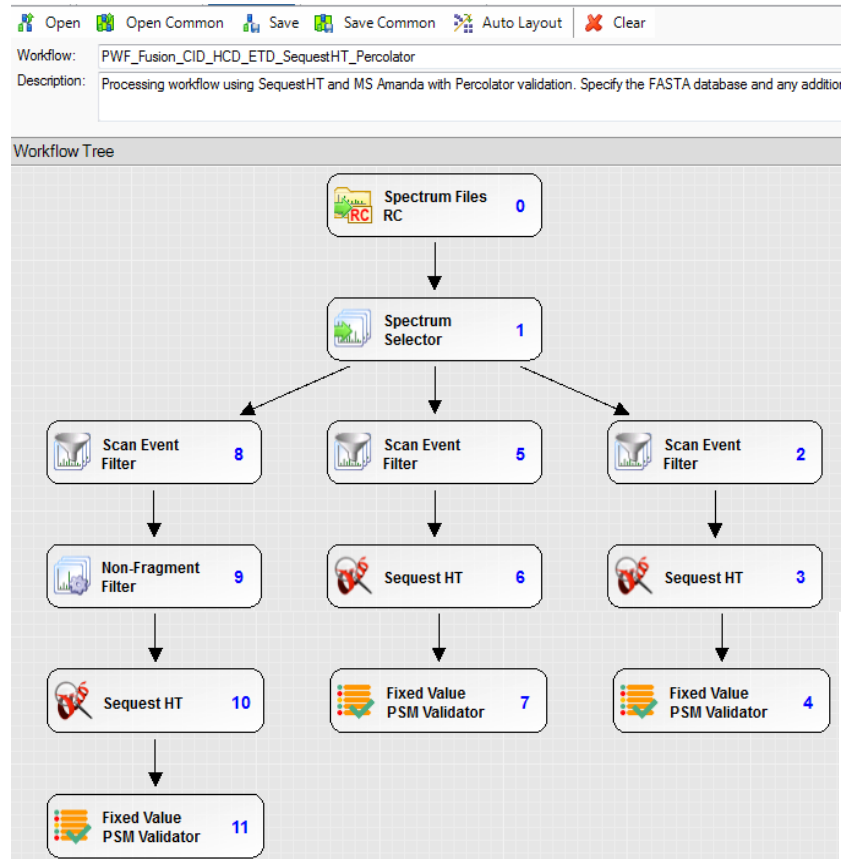


Agenda

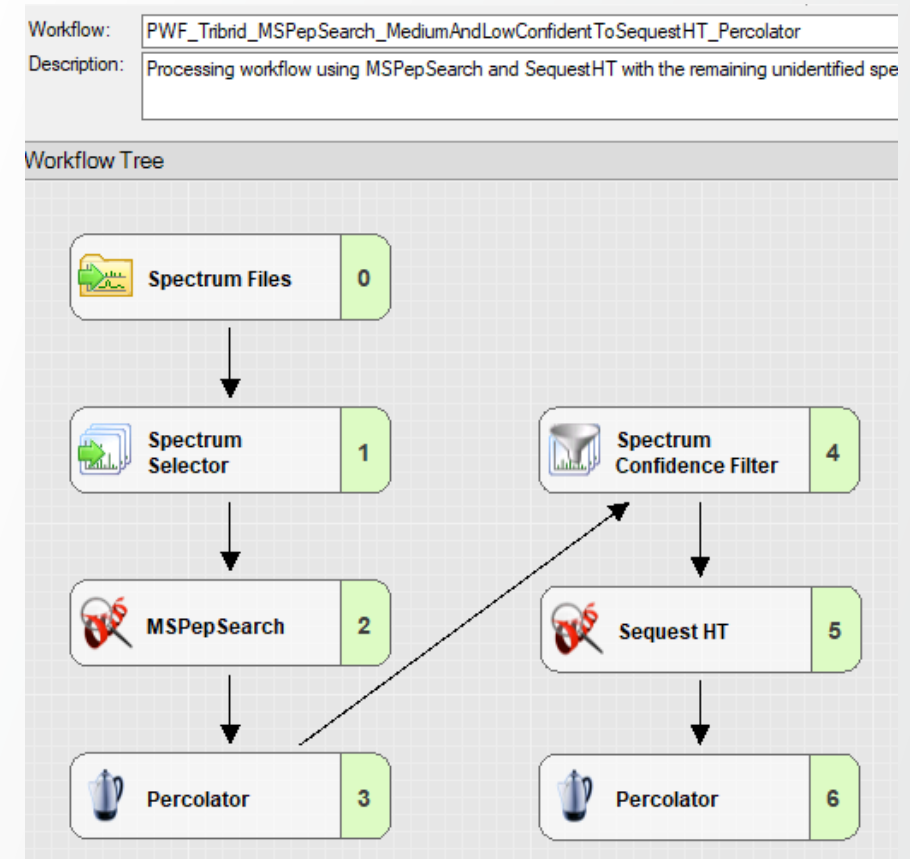
- What's new
 - Enhanced workflow capabilities
 - Spectral libraries
 - INFERYS
 - CHIMERYYS
- TMT quantification
 - TMTpro 18plex
 - PSM filters for quan
 - Normalisation and scaling
 - Multiplexed TMT study design

Enhanced Workflow Capabilities

- Branched

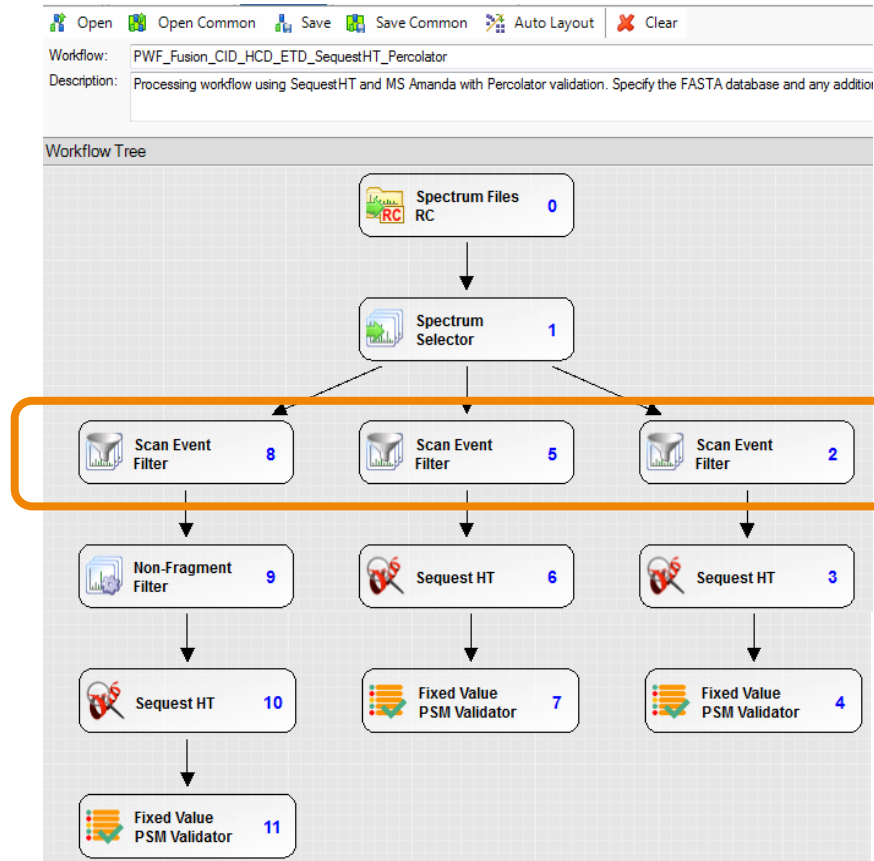


Iterative



Branched WF

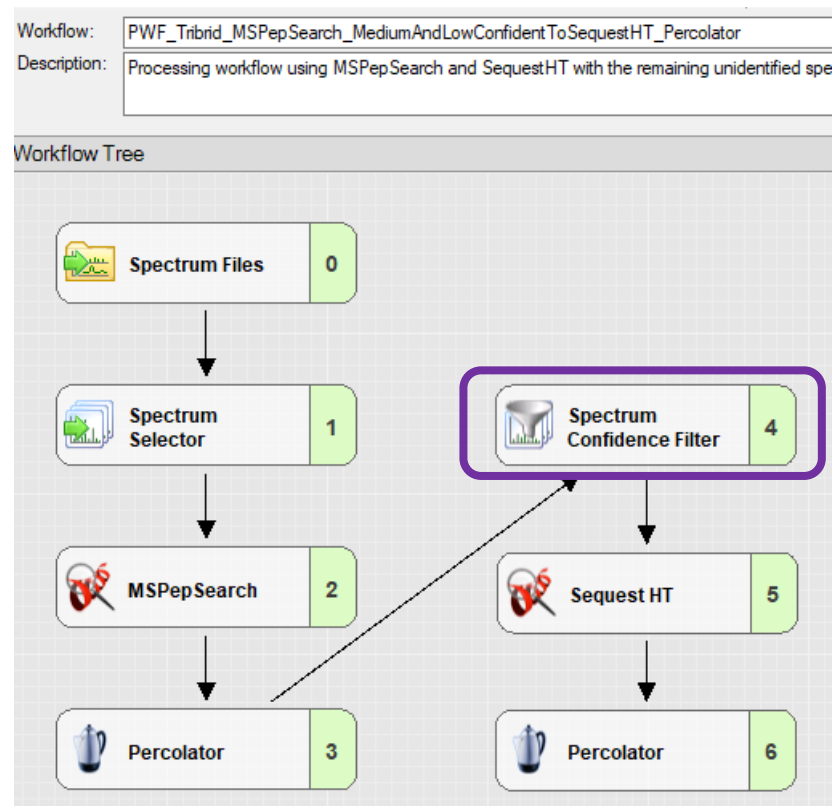
- Use Scan Event Filter node



Parameters of 'Scan Event Filter'	
Show Advanced Parameters	
Filter Settings	
Mass Analyzer	(Not specified)
MS Order	(Not specified)
Activation Type	(Not specified)
Min. Collision Energy	0
Max. Collision Energy	1000
Scan Type	(Not specified)
Polarity Mode	(Not specified)

Iterative WF

- Use **Spectrum Confidence Filter** node
 - Sends selected spectra for the next round of processing



Processing Workflow | Consensus Workflow

Parameters of 'Spectrum Confidence Filter'

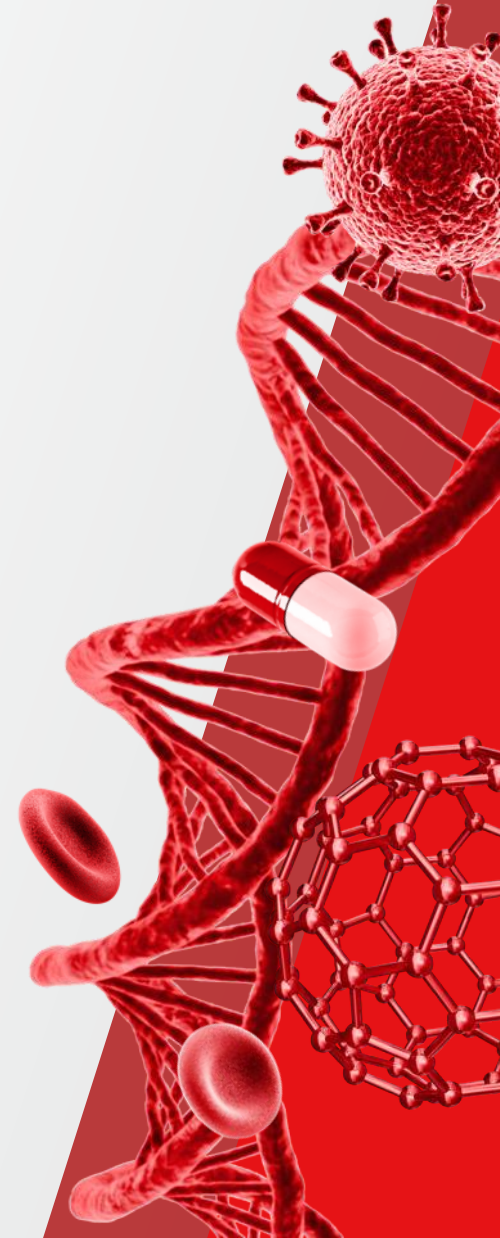
Show Advanced Parameters

▼ **Filter Settings**

Spectrum Confidence	Worse Than High
---------------------	-----------------

Spectral Libraries

 The world leader in serving science

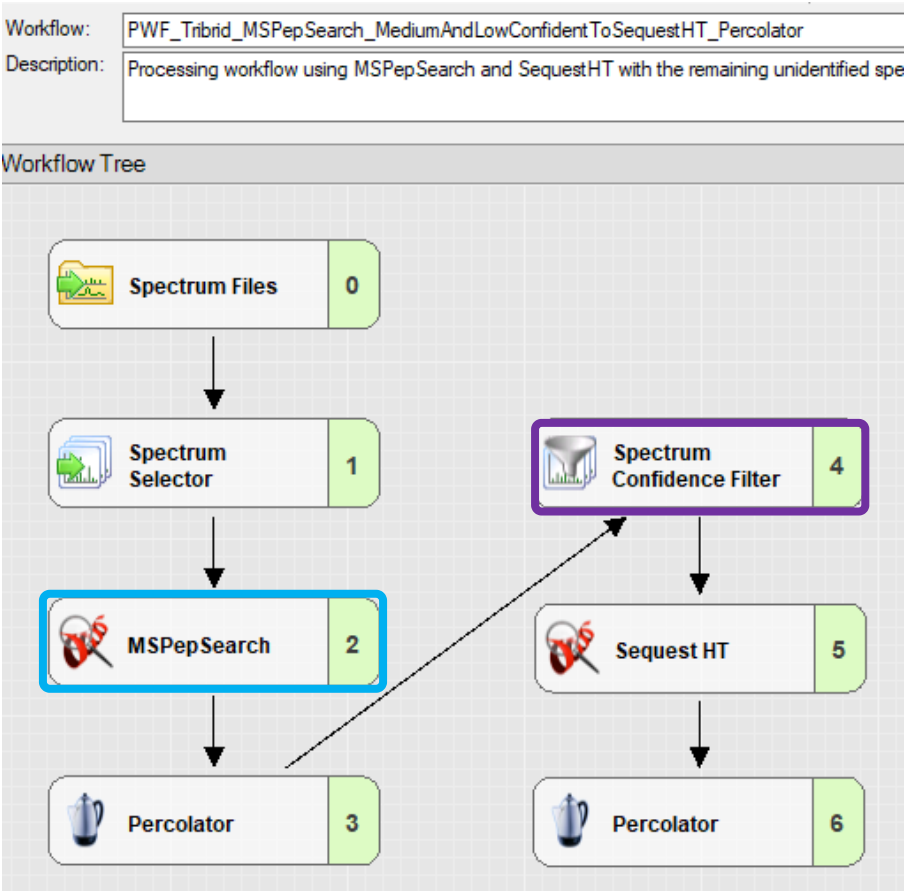


Spectral Libraries in PD

- Library search engine
 - Uses fragment abundances in the correlation of the experimental and library spectra
- Complements database search
 - Fast, multithreaded in PD 2.5, more sensitive
 - With confidence estimate based on decoy library search (FDR)
- Libraries can be directly downloaded in PD Administration
 - **NIST** – large number of spectra, various modifications
 - **ProteomeTools** – libraries of synthetic peptides, some modified peptides included, acquired at various CE settings

Spectral Library Search

- MSPepSearch deployed as part of an iterative WF



Parameters of 'MSPepSearch'

Show Advanced Parameters

- 1. **Input Data**
 - 1. Spectral Library ProteomeTools_HCD28_PD
 - 1. Protein Database Homo sapiens (SwissProt TaxID=9606).fasta
 - 2. Spectral Library
 - 2. Protein Database
 - 3. Spectral Library
 - 3. Protein Database
- 2. **Search Settings**
 - Precursor Mass Toleranc 10 ppm
 - Fragment Mass Toleranc 0.02 Da

Processing Workflow Consensus Workflow

Parameters of 'Spectrum Confidence Filter'

Show Advanced Parameters

- Filter Settings
 - Spectrum Confidence Worse Than High

Interpreting Library Search Results

- Empirically derived score thresholds (my own; not a general recommendation)
 - dot Score ≥ 500
 - rev-dot Score ≥ 800
- Set as PSM filters in **MSF Files** node

The screenshot displays the 'Consensus Workflow' tab in the software. The 'Parameters of MSF Files' section is expanded, showing various settings. The 'PSM Filters' section is highlighted, with the following parameters:

Parameter	Value
Maximum Delta Mass	0 ppm
1. Score	MSPepSearch: dot Score
1. Threshold	500
2. Score	MSPepSearch: rev-dot Score
2. Threshold	800
3. Score	
3. Threshold	0
4. Score	
4. Threshold	0

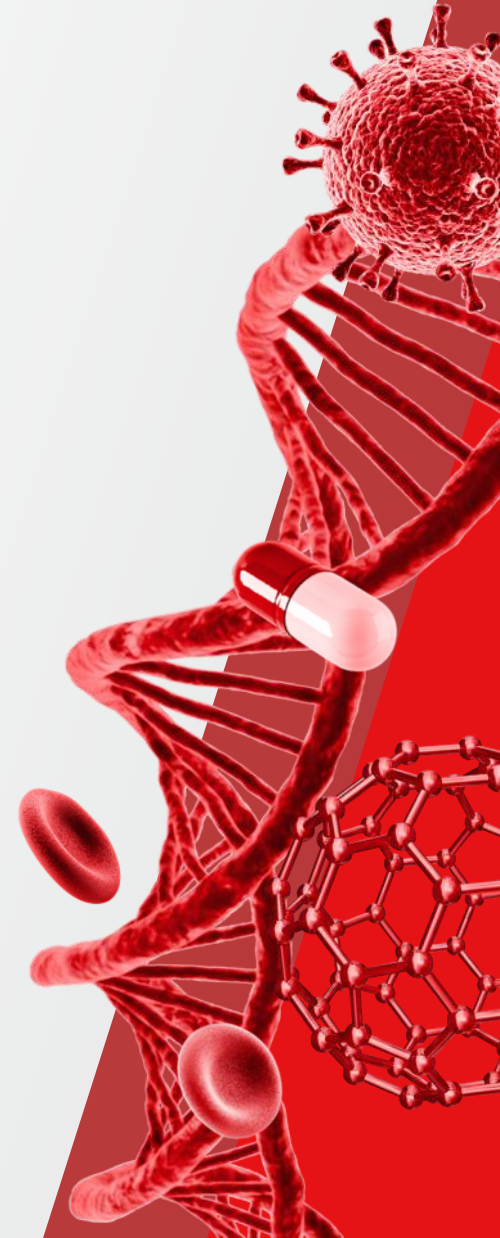
The 'Workflow Tree' on the right shows the workflow steps:

- MSF Files (0)
- PSM Grouper (1)

The 'Workflow' field is set to 'PD_25Dev_02_Freeze' and the 'Description' is 'Result filtered for high confidence amount per channel and scale'.

INFERYS - Predicted Spectral Libraries




 The world leader in serving science



Predicted spectral libraries

- Deep learning-based model for peptide MS/MS spectral prediction
- Predicts peptide spectrum (including relative intensities of fragments)
- Remarkable accuracy

Generating high-quality libraries for DIA-MS with empirically-corrected peptide predictions

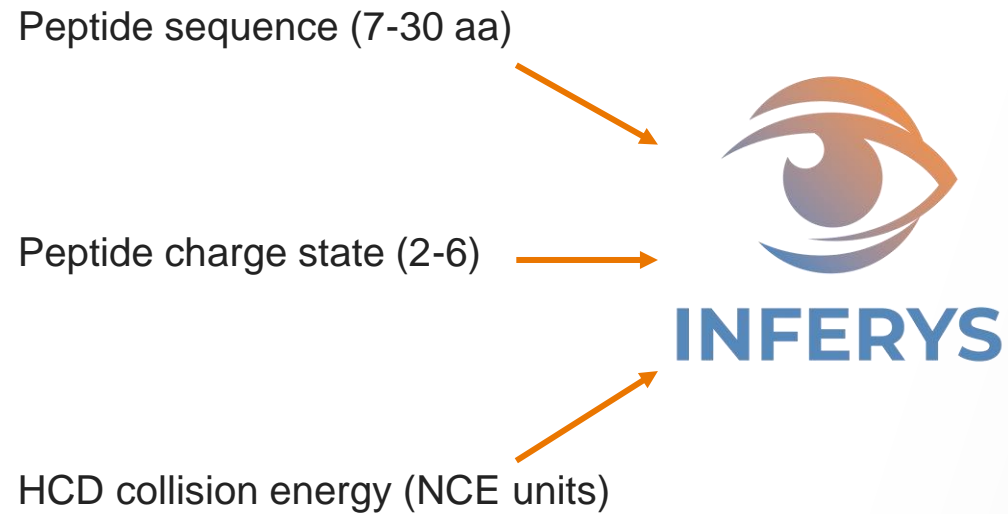
 Brian C. Searle,  Kristian E. Swearingen, Christopher A. Barnes,  Tobias Schmidt,  Siegfried Gessulat,  Bernhard Kuster,  Mathias Wilhelm

doi: <https://doi.org/10.1101/682245>

Now published in *Nature Communications* doi: [10.1038/s41467-020-15346-1](https://doi.org/10.1038/s41467-020-15346-1)

Predicted Spectral Libraries

- Deep learning-based model for peptide MS/MS spectra prediction



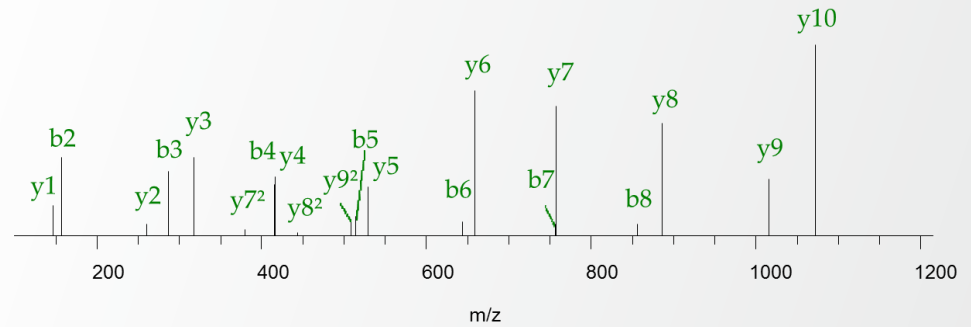
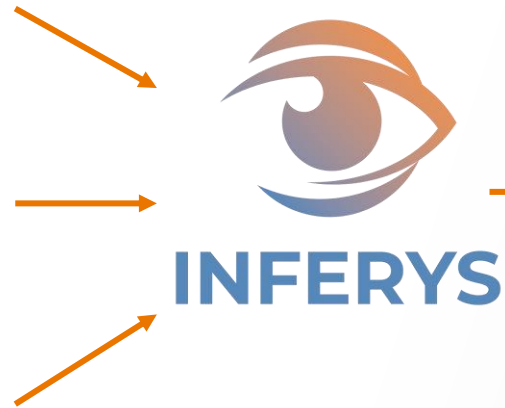
Predicted Spectral Libraries

- Deep learning-based model for peptide MS/MS spectra prediction

VGEEVEIVGIK

$z = 2+$

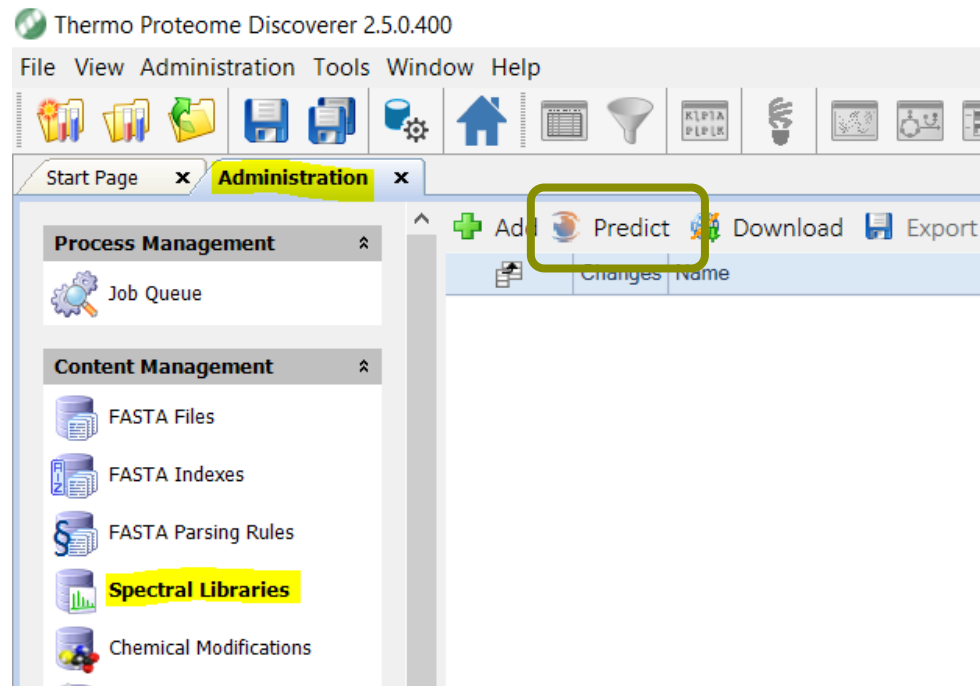
NCE = 26



Predicted MS/MS spectrum

Predicted Spectral Libraries

- Administration → Maintain Spectrum Libraries → Predict
- Create a predicted spectral library for all peptides in a FASTA file
- Use predicted libraries for searches with MSPepSearch node



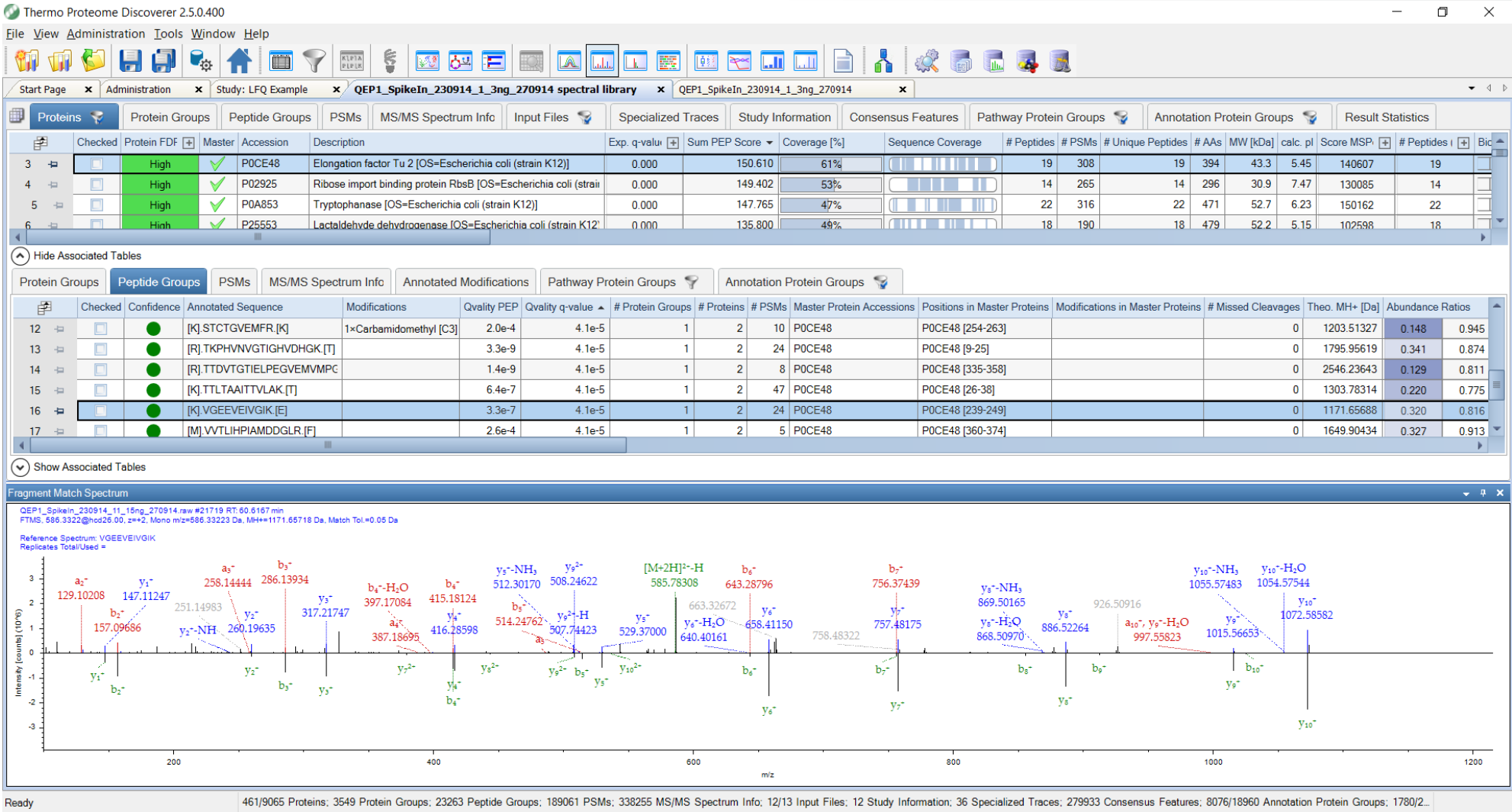
Predict a Spectrum Library

1. General	
FASTA File	
Library Name	
Organism	
Activation Type	HCD
Collision Energy	28
Comment	
2. Digest	
Enzyme	Trypsin
Min. Peptide Length	7
Max. Peptide Length	30
Min. Precursor Charge	2
Max. Precursor Charge	3
Maximum Missed Cleavage Sites	2
1. Static Modification	Carbamidomethyl (C)
1. Dynamic Modification	
Max. Equal Modifications Per Peptide	3

Search Result with Predicted Library



- Match between experimental and predicted spectrum



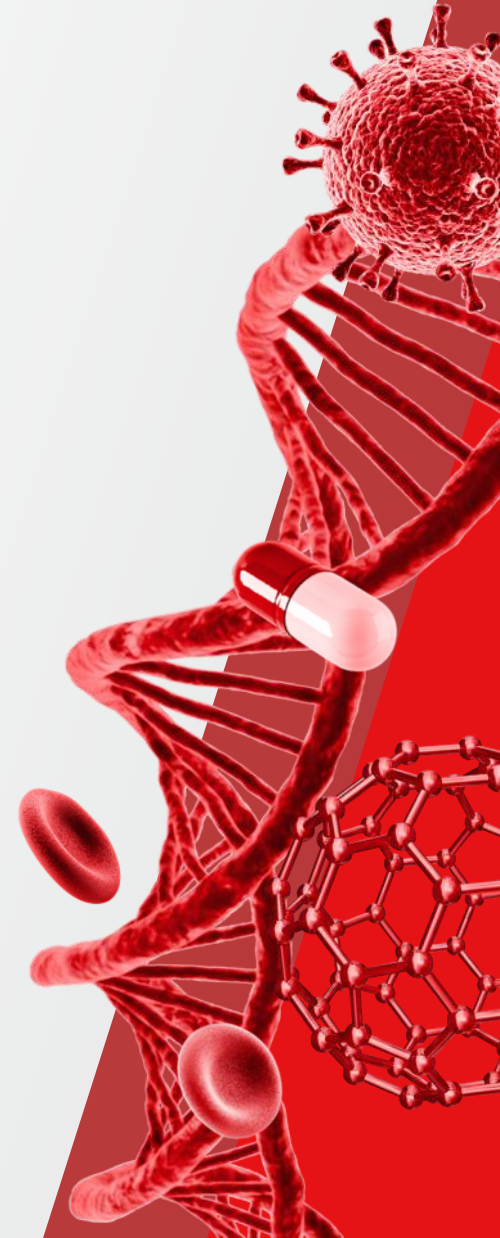
- Prediction of spectral libraries has a **limited practical use**
 - Recommended for small FASTA files
 - O.K. to characterize a recombinant protein or a protein complex
 - High consumption of disc space

Caveats

- High consumption of disk space
- *Example: working with human SwissProt-derived library*
 - Predicted library size 11 - 12 GB
 - Decoy DB of approx. the same size needed
 - When a library is being predicted, there are intermediate files (the prediction queries and the prediction result files) being saved on the system (100 GB free disc space needed)
 - The translation of the prediction format to the NIST binary format required for the searches also needs a lot of space
 - PD stores the annotated reference spectra in a separate library file for a quick display. This file is 10-20 GB
- It takes about 1 day

INFERYS - Rescoring

 The world leader in serving science



Prediction-based rescoring for improved confidence

- **INFERYS Rescoring** node
 - predicts MS/MS spectra for PSMs identified by Sequest HT
 - subsequently compares the predicted and experimental spectra
- Updated **Percolator** node
 - includes new features for improved peptide ID confidence

<https://doi.org/10.1002/rcm.9128>



SPECIAL ISSUE PAPER

INFERYS rescoring: Boosting peptide identifications and scoring confidence of database search results

Daniel P. Zolg, Siegfried Gessulat, Carmen Paschke, Michael Graber, Magnus Rathke-Kuhnert, Florian Seefried, Kai Fitzemeier, Frank Berg, Daniel Lopez-Ferrer, David Horn, Christoph Henrich, Andreas Huhmer, Bernard Delanghe ✉, Martin Frejno ✉ ... [See fewer authors](#) ^

First published: 20 May 2021 | <https://doi.org/10.1002/rcm.9128>

[Read the full text](#) >

PDF TOOLS SHARE

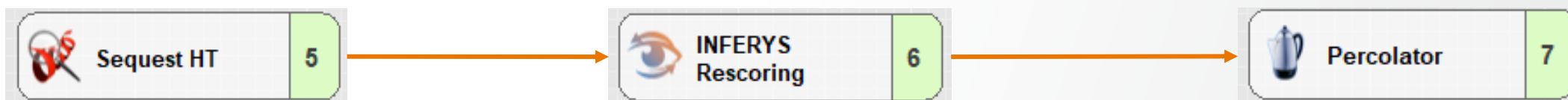
Abstract

Database search engines for bottom-up proteomics largely ignore peptide fragment ion intensities during the automated scoring of tandem mass spectra against protein databases. Recent advances in deep learning allow the accurate prediction of peptide fragment ion intensities. Using these predictions to calculate additional intensity-based scores helps to overcome this drawback.

Here, we describe a processing workflow termed INFERYS™ rescoring for the intensity-based rescoring of Sequest HT search engine results in Thermo Scientific™ Proteome Discoverer™ 2.5 software. The workflow is based on the deep learning platform INFERYS capable of predicting fragment ion intensities, which runs on personal computers without the need for graphics processing units. This workflow calculates intensity-based scores comparing peptide spectrum matches from Sequest HT and predicted spectra. Resulting scores are combined with classical search engine scores for input to the false discovery rate estimation tool Percolator.

INFERYS Rescoring Node

Additional features calculated as input into Percolator

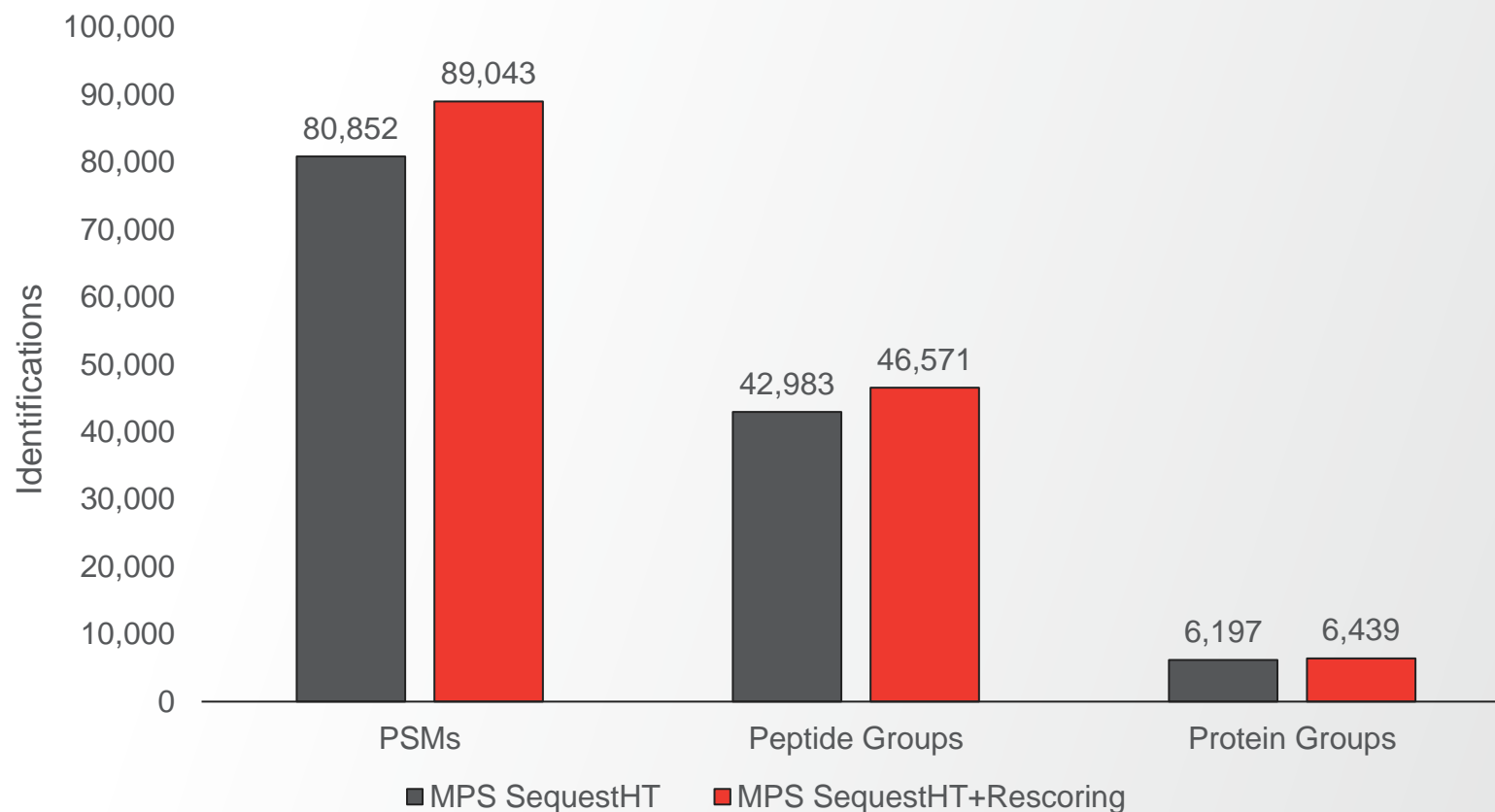


- Search against FASTA file
- Produce a list of PSMs from the target and decoy database searches
- Produce up to rank 10 hits per spectrum
- Determine optimal collision energy for input into INFERYS algorithm
- Calculate all predicted spectra for all target and decoy hits
- Calculate spectral angle and other figures-of-merit for correlation of predicted and measured spectra
- Use the standard 35 features
- If INFERYS Rescoring node is used, add 15 more features
- Calculate FDR thresholds using SVM score calculated from all input features
- For information about Percolator, read Matthew The *et al*, JASMS 2016

Results – Standard bottom-up proteomics data set

200 ng HeLa Thermo Scientific™ Orbitrap Exploris™ 480 MS with FAIMS CV50 CV70

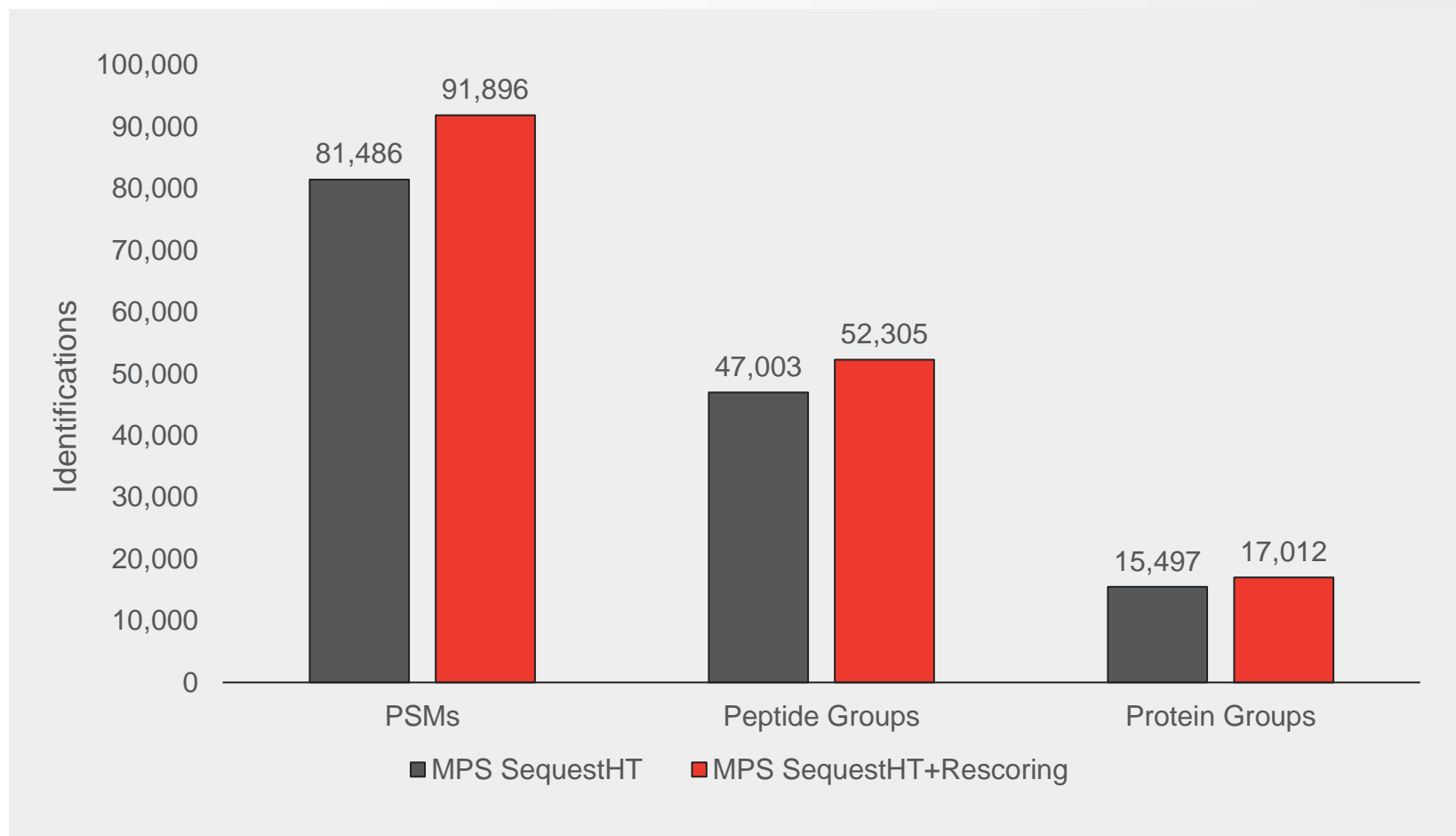
- Trypsin digest
- 10% increase in PSMs
- 8% more unique peptides
- 4% more proteins
- Rescues short peptides and peptides where classical scores fail to separate targets from decoys



Results – Standard bottom-up proteomics data set

Metaproteomics – Human stool samples (*Rechenberger et al., Proteomes. 2019*)

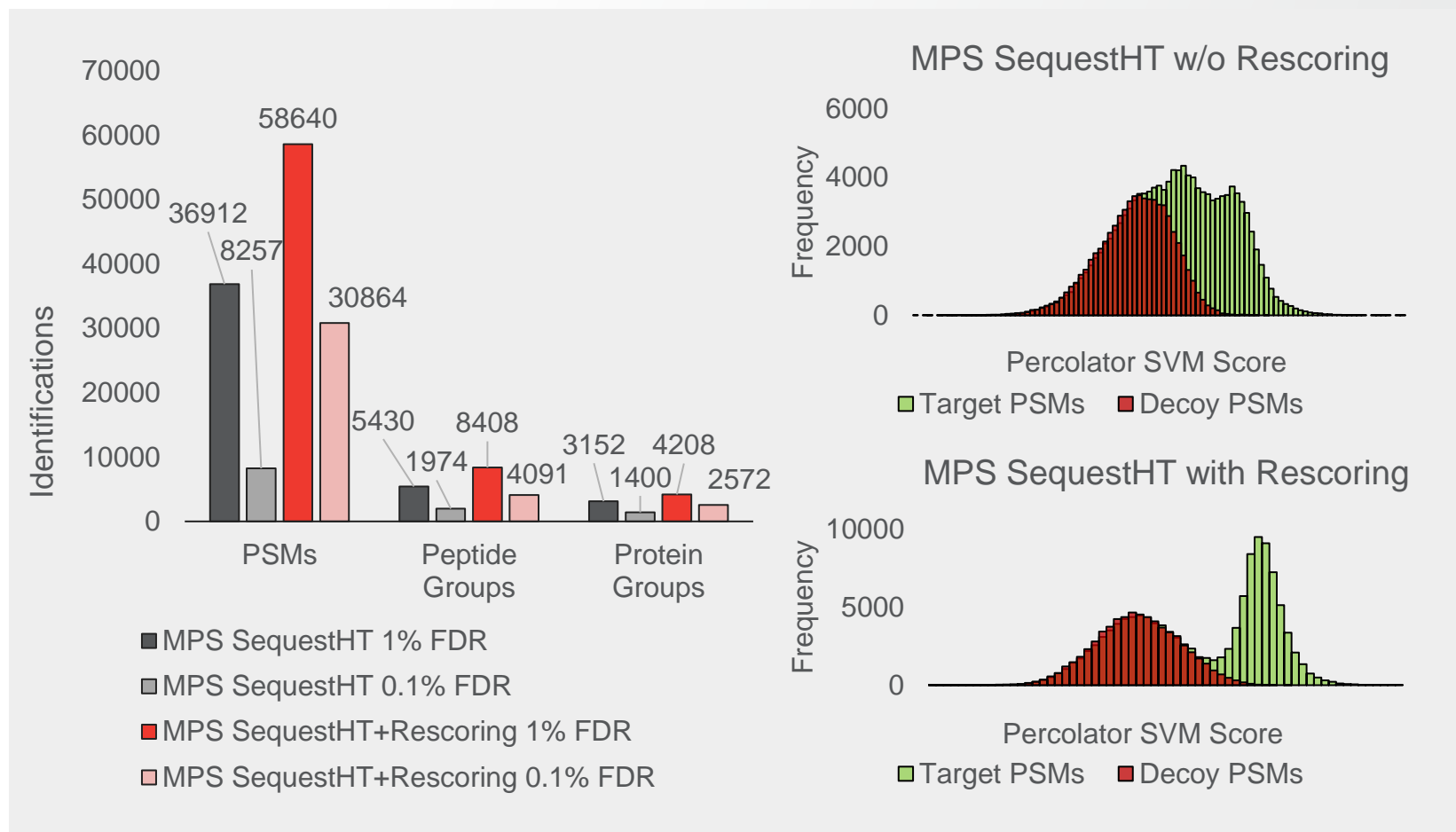
- Trypsin digest
- Complex sample with huge search space
- 13% increase in PSMs
- 11% more peptides
- 10% more proteins
- Rescues short peptides and peptides where classical scores fail to separate targets from decoys



Results – Immunopeptidomics data set

HLA Class I data set – Patient derived melanoma cell line (*Chong. et al, Nat. Com. 2020*)

- No enzyme search
- Huge search space and very similar peptide properties
- 59% increase in PSMs
- 55% more peptides
- 34% more proteins
- Reduces peptide loss at 0.1% compared to 1% FDR from 64% to 25%



Default Processing Workflows

- Q Exactive

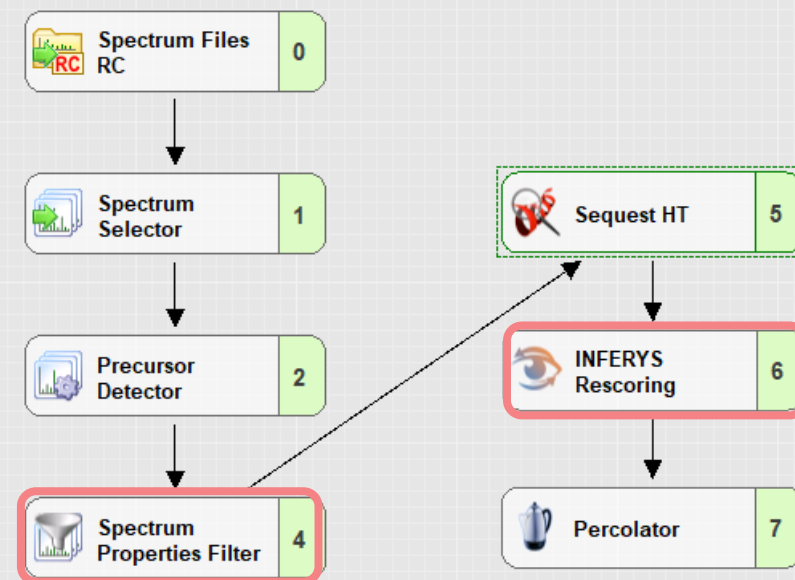
Name

- PWF_QE_Basic_SequestHT
- PWF_QE_Dimethylation_Quan_Sequest_HT_Percolator
- PWF_QE_INFERYS_Rescoring_SequestHT_Percolator
- PWF_QE_INFERYS_Rescoring_SequestHT_Percolator_2stage
- PWF_QE_Iterative_SequestHT_Percolator
- PWF_QE_Precursor_Quan_and_LFQ_MPS_SequestHT_Percolator
- PWF_QE_Precursor_Quan_and_LFQ_SequestHT_Percolator
- PWF_QE_Reporter_Based_Quan_SequestHT_Percolator
- PWF_QE_SequestHT_MSAmanda_Percolator
- PWF_QE_SequestHT_MSAmanda_Percolator_phosphoRS
- PWF_QE_SequestHT_MSAmanda_Percolator_ptmRS

Workflow: PWF_QE_INFERYS_Rescoring_SequestHT_Percolator

Description:

Workflow Tree



Default Processing Workflows

- **INFERYS** node
 - Define collision energy in “advanced parameters”
- Must include **Spectrum Properties Filter** node

Parameters of 'INFERYS Rescoring'

Hide **Advanced Parameters**

General	
Automatic Mode	True
Collision Energy	28

Parameters of 'Spectrum Properties Filter'

Show Advanced Parameters

1. Spectrum Properties	
Lower RT Limit	0
Upper RT Limit	0
Lowest Charge State	2
Highest Charge State	6
Min. Precursor Mass	0 Da
Max. Precursor Mass	0 Da
2. Spectrum Properties Filter	
First Scan	0
Last Scan	0
Ignore Specified Scans	
3. Thresholds	
Total Intensity Threshold	0
Minimum Peak Count	1
4. Filter Spectra by Peak Properties	
Filter Mode	Pass Through Matching Spectra
Peak Masses	
Neutral Loss Masses	
Match Tolerance	0.02 Da

Summary – INFERYS Rescoring

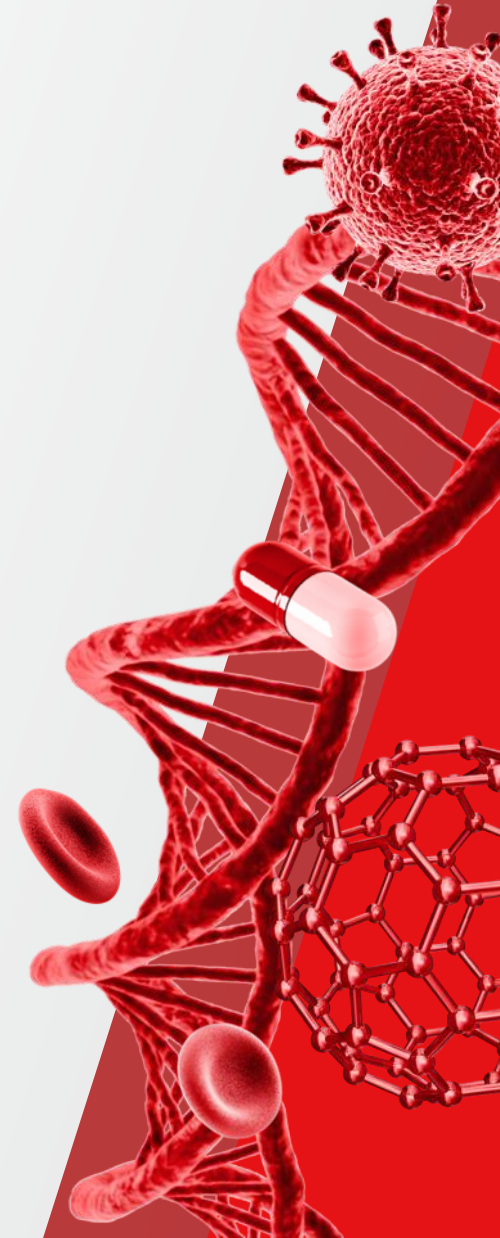
- **Significant increase in IDs**
- Useful for extremely large search space in particular
 - no-enzyme searches
 - metaproteomics

Caveats

- Input from .raw file
- High res HCD spectra
- Charge state 2-6
- Only with Sequest HT (does not work with Mascot)
- Not trained for TMT-labeled peptides

CHIMERYYS

 The world leader in serving science



CHIMERYS

Deconvolutes chimeric spectra based on predicted fragment ion intensities.

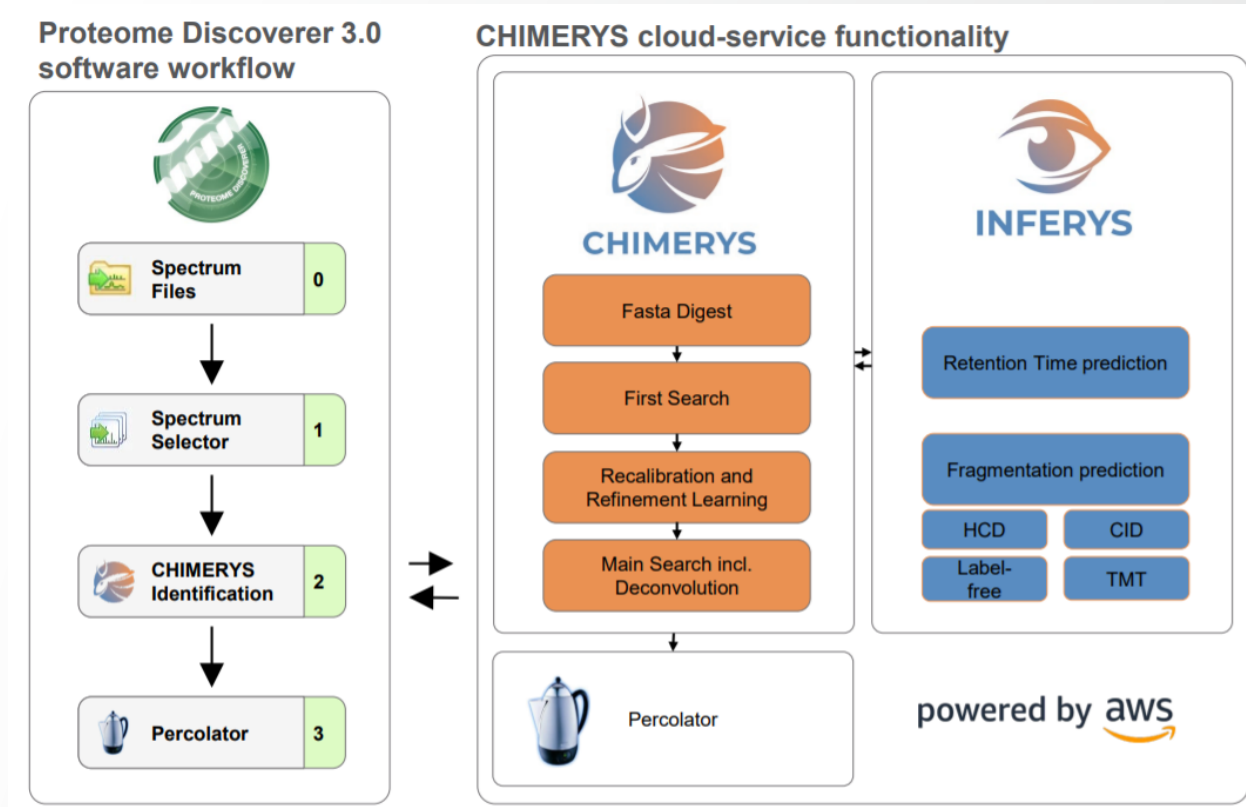
A cloud-native search algorithm that uses accurate predictions of peptide fragment ion intensities and retention times provided by the deep learning framework INFERYS 2.0.

- Artificial intelligence-powered search algorithm
- Implemented in **PD 3.0** (release February 2022)
- 2x peptide IDs compared to Sequest HT alone
- Protein coverage increase by 2.5-fold
- More than 80% identification rate



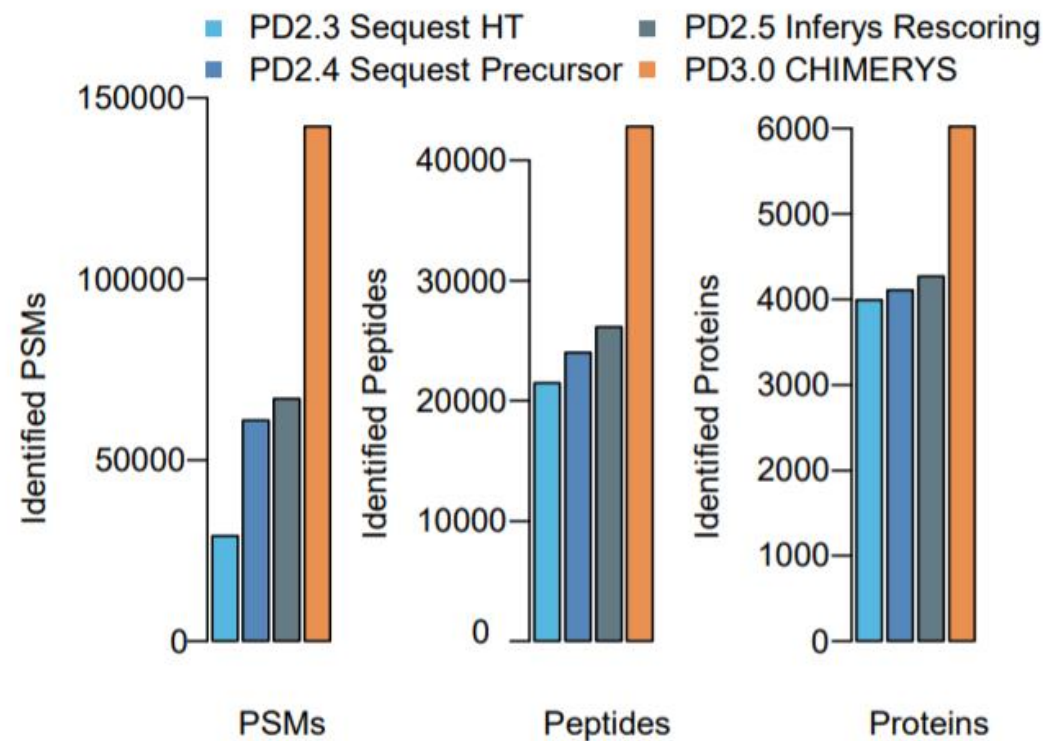
Deconvolutes chimeric spectra based on predicted fragment ion intensities

- All candidates in the isolation window of a given tandem mass spectrum are considered simultaneously
- Compete for measured fragment ion intensity in one concerted step
- CHIMERYS aims to explain as much measured intensity with as few candidate peptides as possible, resulting in the **deconvolution of chimeric spectra**
- Peptide spectrum match (PSM)-level false discovery rate (FDR)-control is performed using Percolator

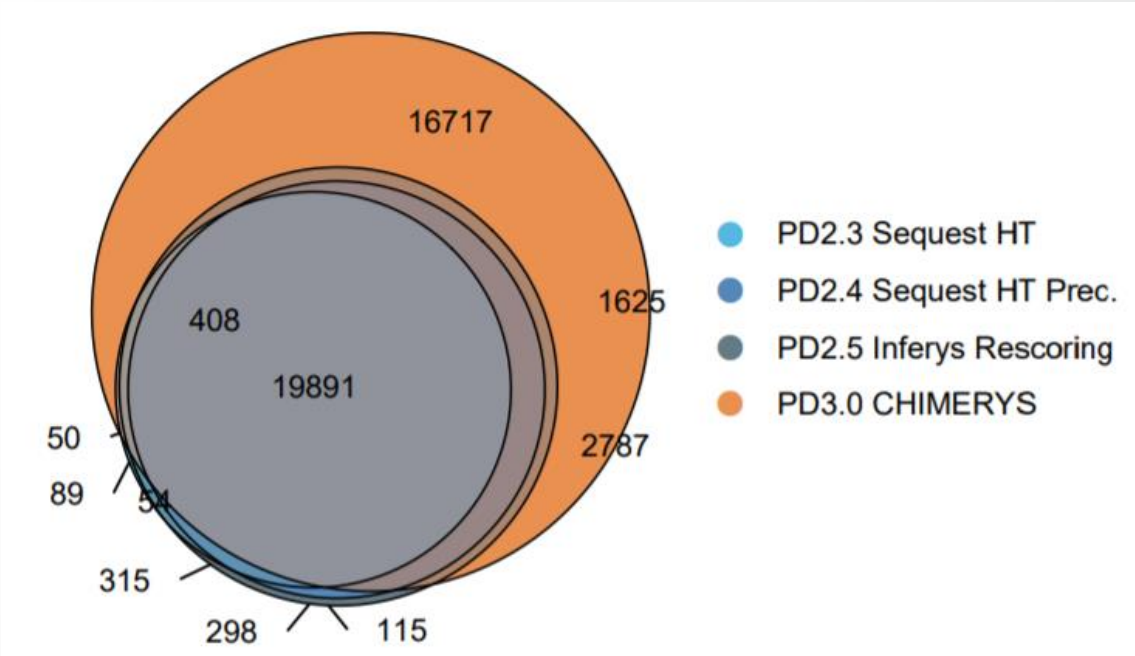


Peptide and Protein IDs

HeLa cell lysate, 1-hour gradient, Thermo Scientific™ Orbitrap Exploris™ 480



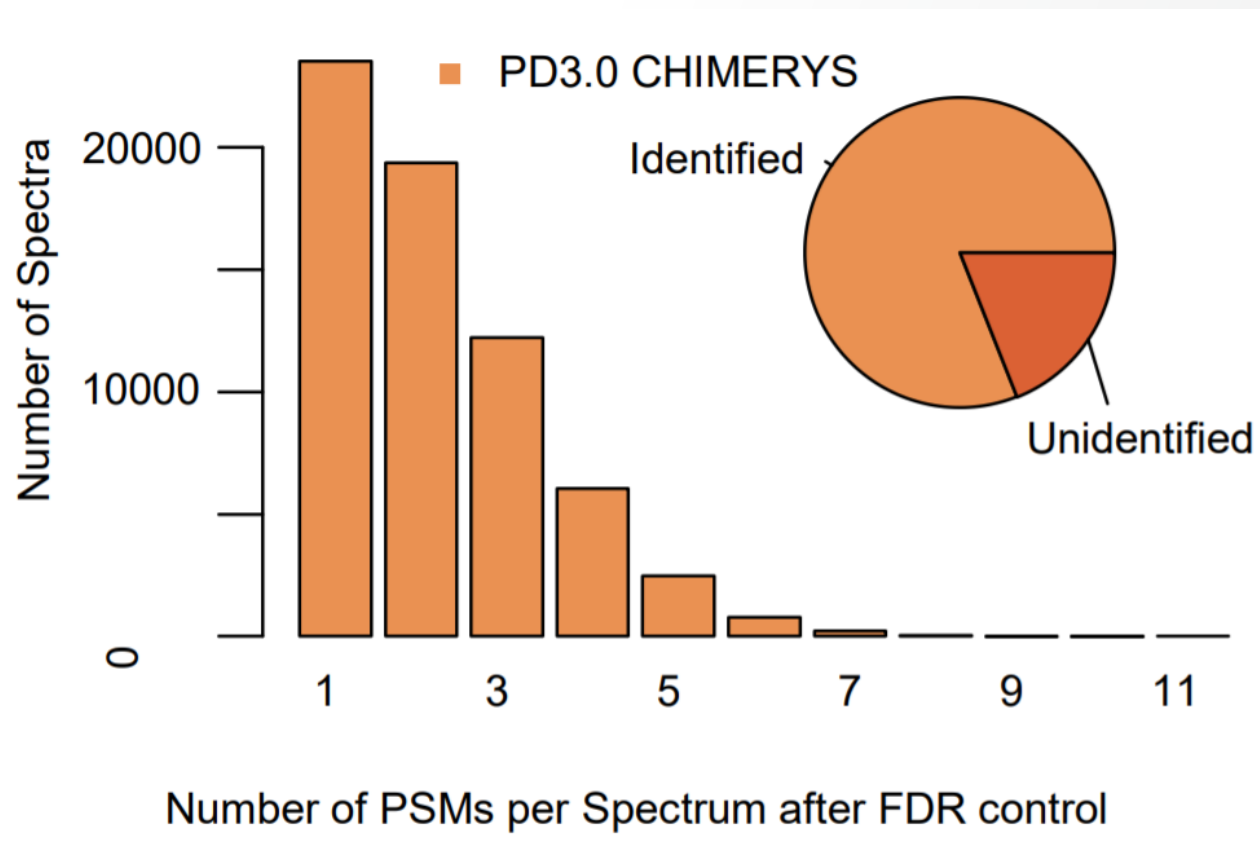
Overlap of peptide identifications



Extent of Chimeric Spectra Problem

Number of PSMs per spectrum and identification rate achieved by CHIMERY5

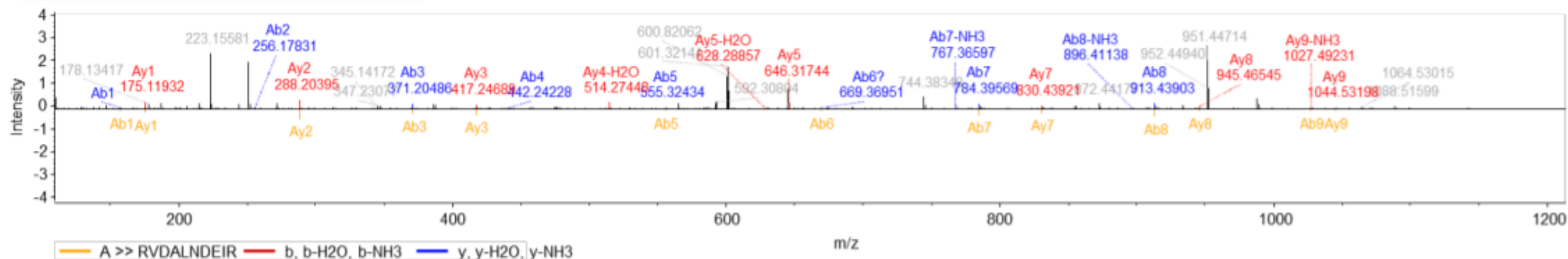
- More than 80% identification rate



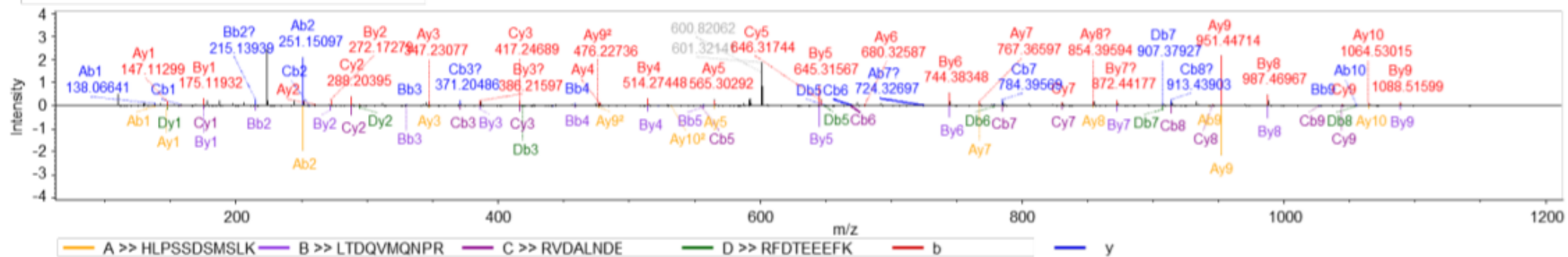
Proteome Discoverer Spectrum Viewer

Visualise the proportional contributions of the individual peptides in a mirror plot

PD 2.5 – Sequest HT + INFERYS Rescoring: 1 PSM

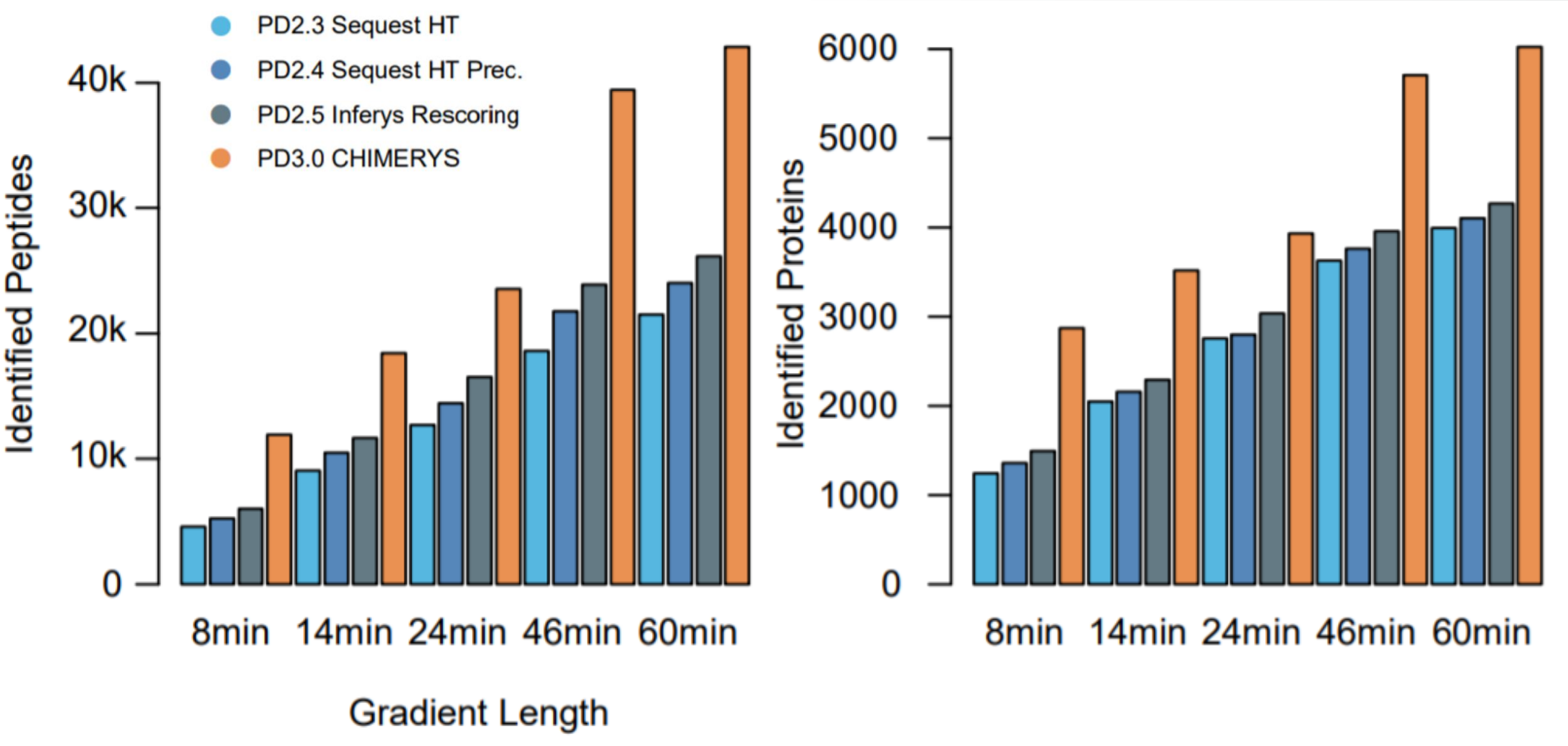


PD 3.0 – CHIMERYYS: 4 PSMs



Increased Throughput

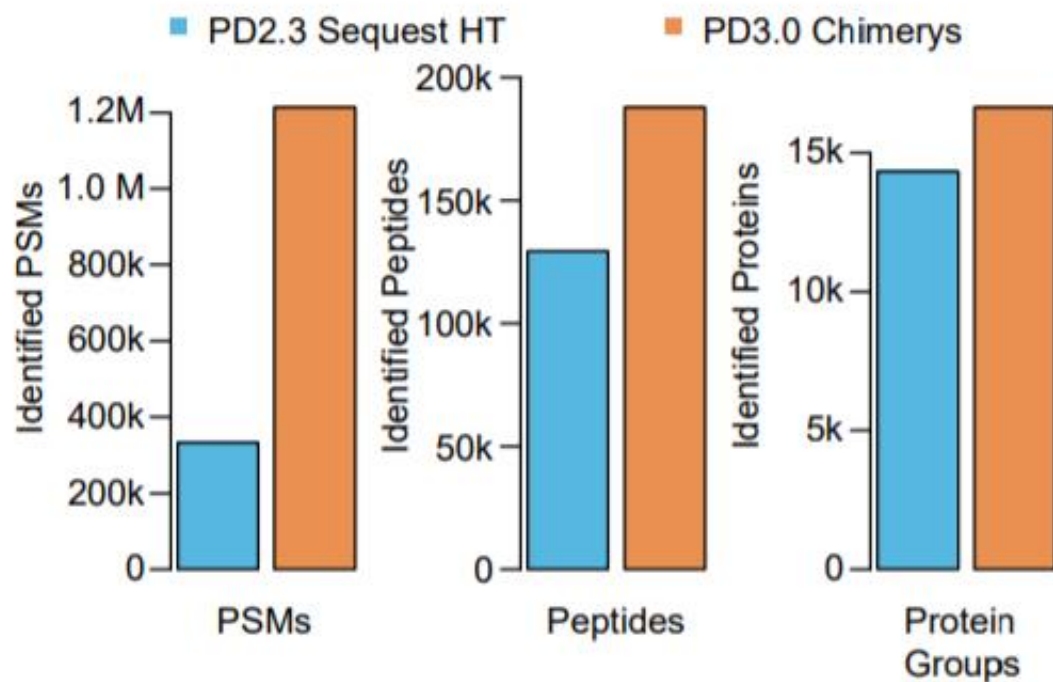
Same number of peptides and protein groups in 1/3 of the measurement time



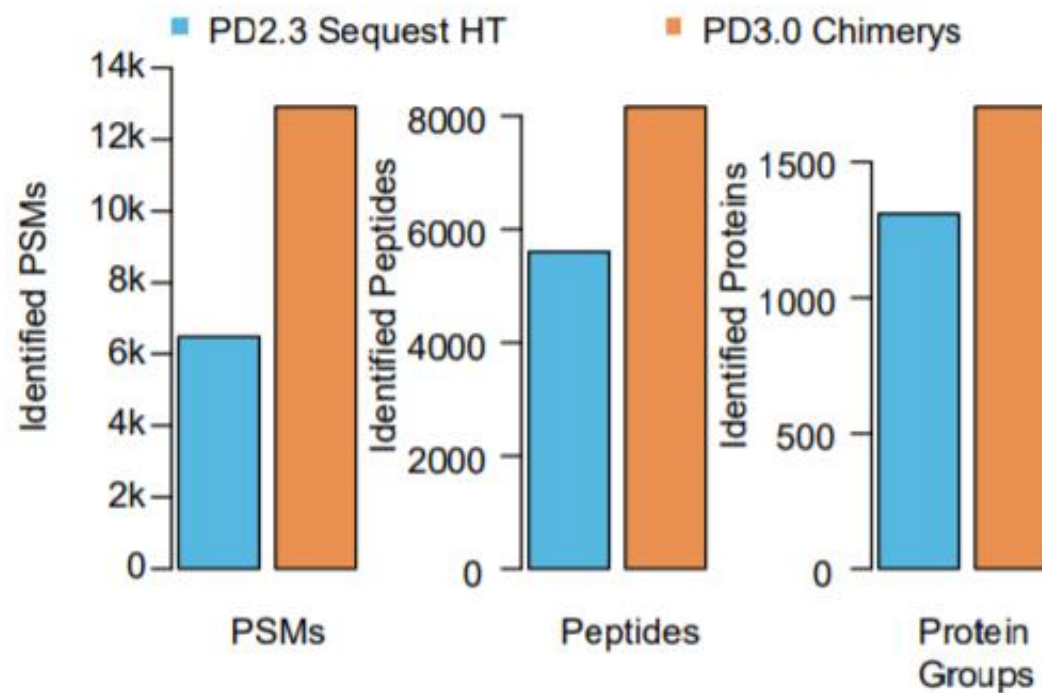
Applicability

Organisms from all kingdoms of life as well as less complex samples like body fluids

Fractionated *Arabidopsis thaliana* proteome
PRIDE Project PXD019483



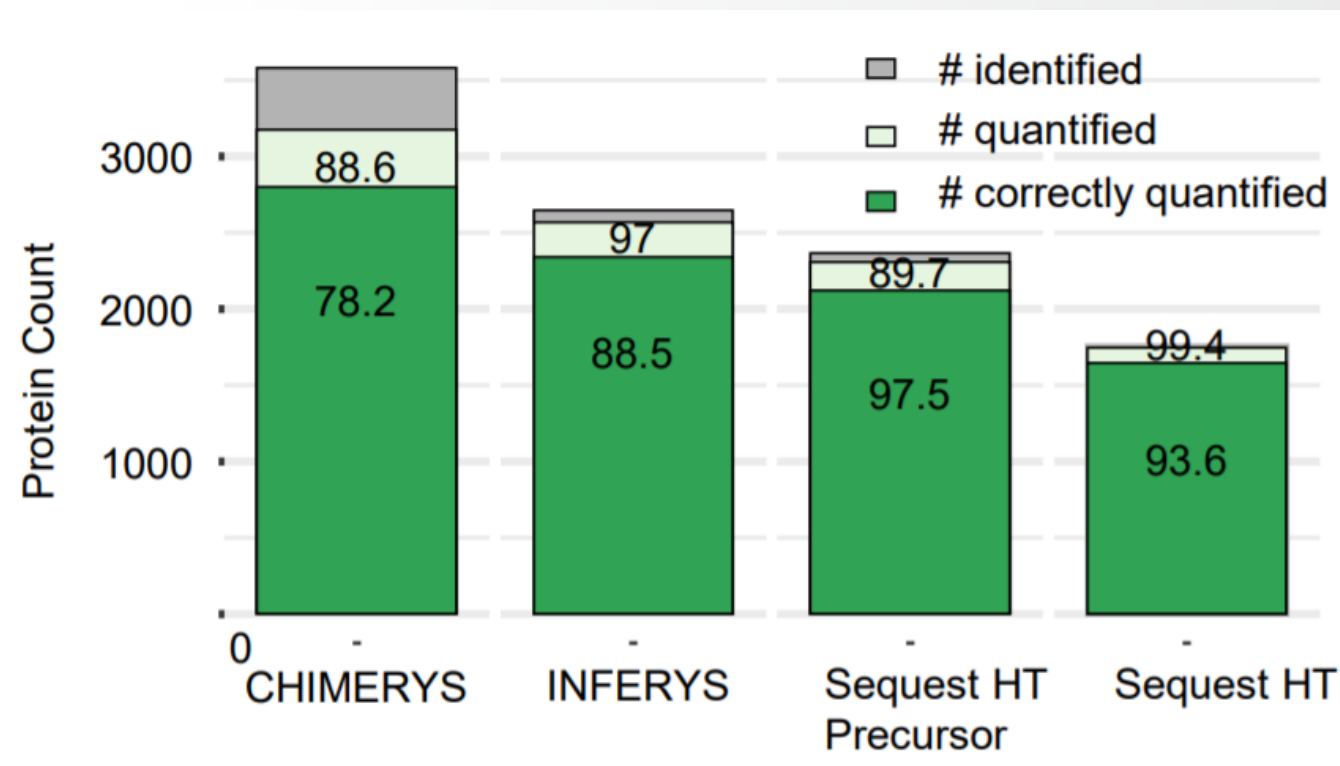
A single 30 min urine proteome file
PRIDE Project PXD015087



LFQ with CHIMERYS

A two-organism dilution series

- Distribution of quantitative yeast protein ratios from dilution experiment
(correct: $0.5 * \text{expected} < r < 2 * \text{expected}$)
- CHIMERYS produces more, especially lower abundant quantified peptides and proteins
- 75% more correctly quantified proteins compared to Sequest HT



CHIMERYYS - Summary

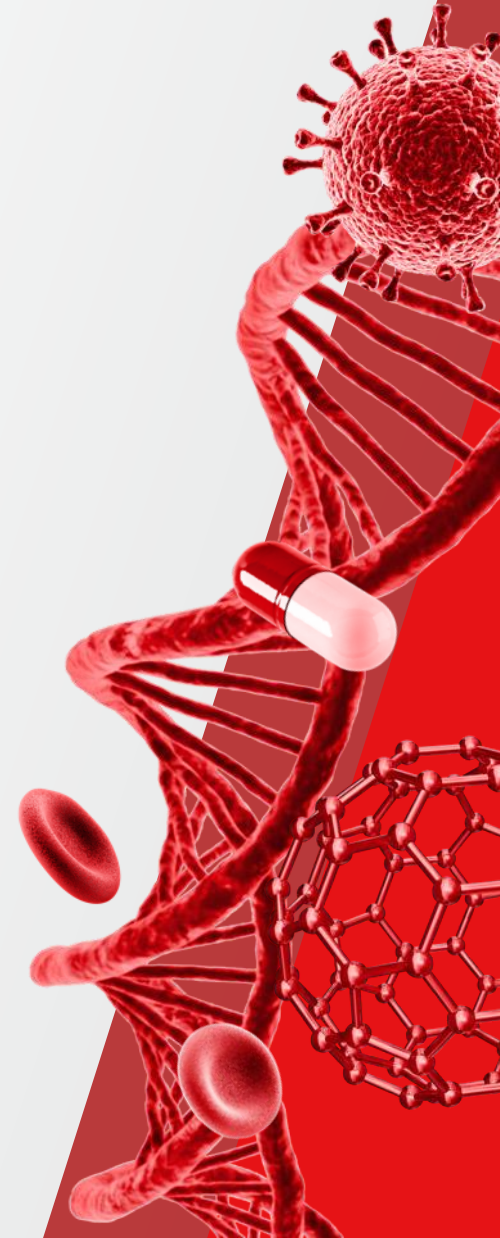
Deconvolutes chimeric spectra based on predicted fragment ion intensities.

- Cloud-native search algorithm that uses AI-based predictions to deconvolute chimeric spectra
- Fully integrated into Proteome Discoverer 3.0 software
- CHIMERYYS results in drastically increased numbers of PSM, peptide and protein group identifications, higher sequence coverage and more confident quantification
- CHIMERYYS excels at analyzing complex samples, enabling more efficient measurements, advanced acquisition settings and shorter gradients
- Enhanced proteomic throughput, productivity and efficiency



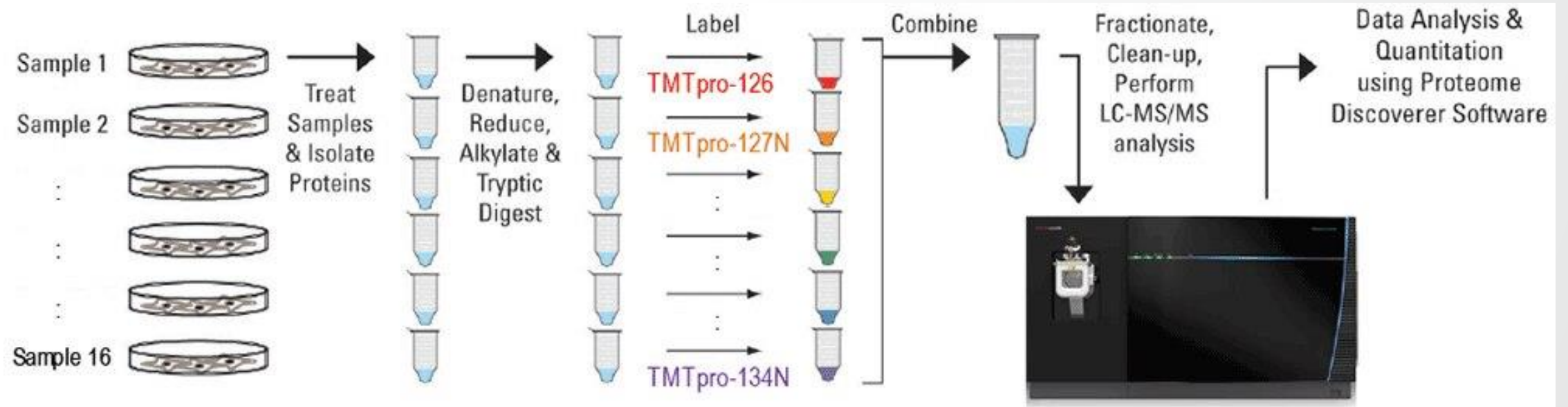
TMTpro 18plex

 The world leader in serving science



General Considerations

- Sample complexity
- Number of samples
- Multiplexed experiments
- Instrument acquisition method
- Data processing



Considerations – Number of samples

- **TMTpro 18plex**
 - Concurrent MS analysis of up to 18 samples derived from cells, tissues, or biological fluids
 - Robust as increased multiplex capability results in fewer missing quantitative values among samples and higher confidence among replicates
 - Efficient amine-reactive, NHS ester-activated reagents ensure efficient labeling of all peptides regardless of protein sequence or proteolytic enzyme specificity
 - Optimized for use with high resolution Thermo Scientific MS/MS platforms

Journal of
proteome
research

pubs.acs.org/jpr

TMTpro-18plex: The Expanded and Complete Set of TMTpro Reagents for Sample Multiplexing

Jiaming Li, Zhenying Cai, Ryan D. Bomgarden, Ian Pike, Karsten Kuhn, John C. Rogers, Thomas M. Roberts, Steven P. Gygi,* and Joao A. Paulo*



Cite This: <https://doi.org/10.1021/acs.jproteome.1c00168>

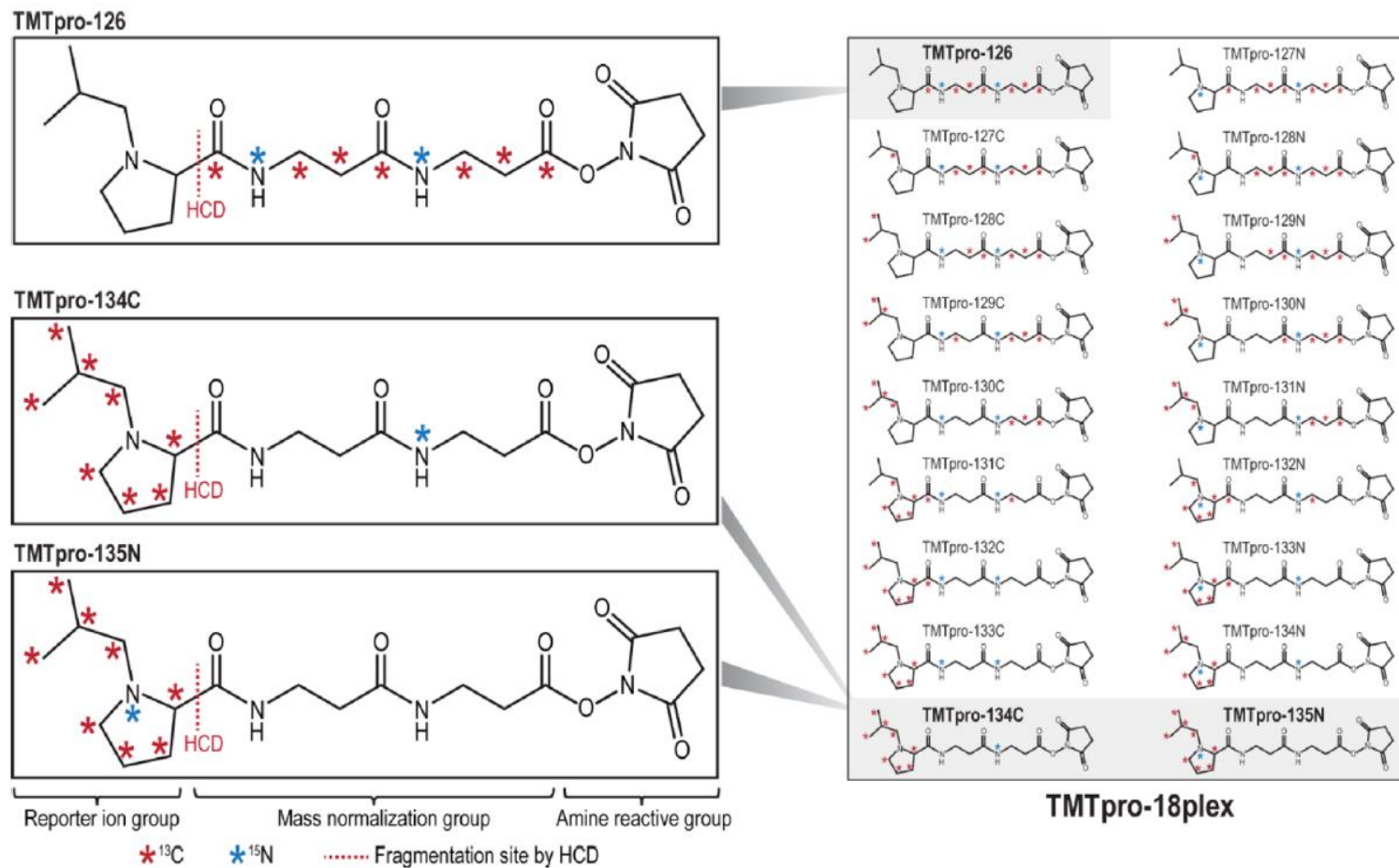


Read Online

<https://doi.org/10.1021/acs.jproteome.1c00168>

TMTpro 18plex

- PD 2.5 quan method available
- <https://www.thermofisher.com/order/catalog/product/A52045>



Considerations – Acquisition Method Parameters

- **TMTpro** labeled peptides fragment easier
 - Decrease collision energy
 - MS2: 28-32 NCE
 - MS3: 45-55 NCE
- **TMTpro** benefits from a better IT scan
 - Use “rapid” scan rather than “turbo”
- **TMTpro** requires stronger signal
 - 16 (18) vs 11-way split of signal
 - Increase ion fill
 - Load more sample
 - Use wider columns 100-120µm
 - Use longer columns (> 35cm)

Properties	Fusion MS2 120 min TMT 11 plex	Fusion MS2 120 min TMT pro 16 plex
Resolution Full MS	120000	120000
AGC target Full MS	4e5	4e5
MS max IT, ms	50	50
Scan range, m/z	400-1400	400-1400
Top Speed, s	3	3
MS2 max IT, ms	120	120
MS2 Isolation window, Th	0.7(2-3)-0.5 (4+)	0.7
MS2 NCE, %	38-40	35
MS2 Intensity threshold	5e4	5e4
Dynamic exclusion, s	60, single charge	60, single charge
MS2 Resolution	50000	50000
MS2 AGC target	1e5	1.2e5
MS3 AGC target		
SPS Isolation window, Th		
SPS NCE, %		
SPS max IT, ms		
SPS settings: # notches, mass range, Tag Exclusion	m/z 110	m/z 110

Considerations – Acquisition Strategy

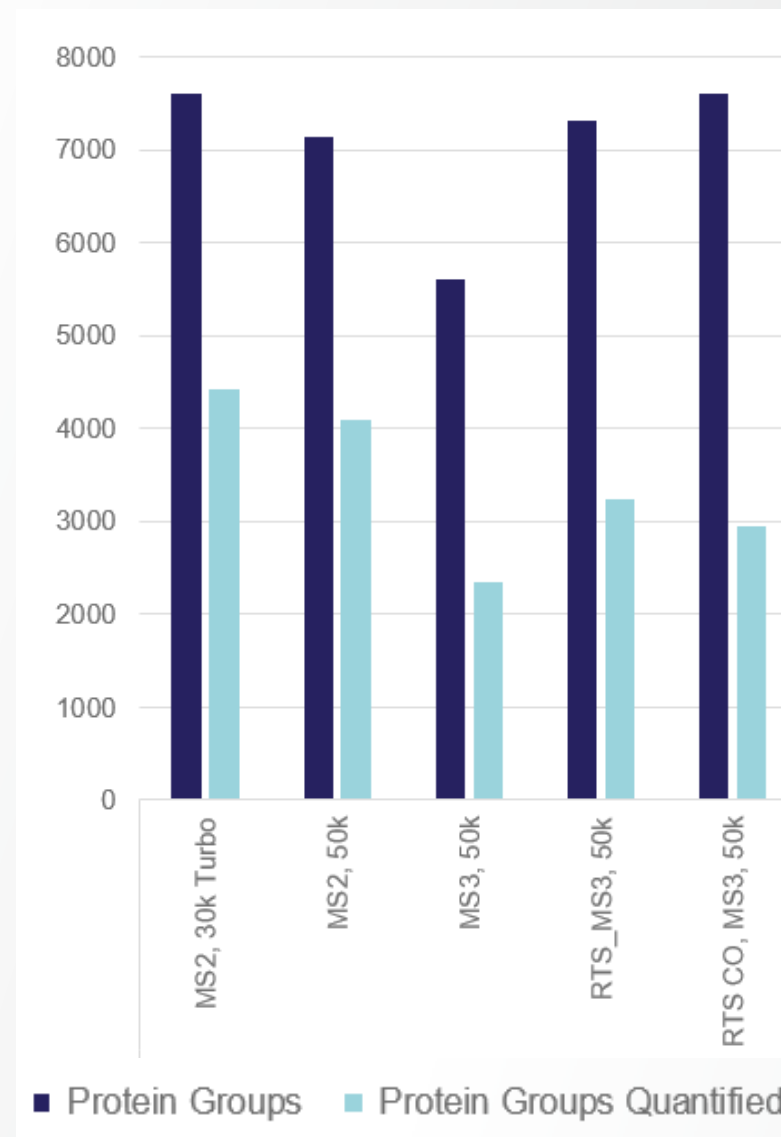
- FAIMS
 - Can decrease co-isolation interference (ratio compression)
- MS2 or MS3 strategy
 - Is ratio accuracy that important?
- Real Time Search
 - Available on Eclipse
- Close Out function
 - Specify the maximum number of peptides per protein to be quantified
- Turbo/Rapid scan
 - Higher number of IDs in MS2 run
 - “Rapid” recommended for TMTpro

Considerations – Acquisition Strategy

Complex mixture, 50 cm column, 90 min gradient, 6 fractions

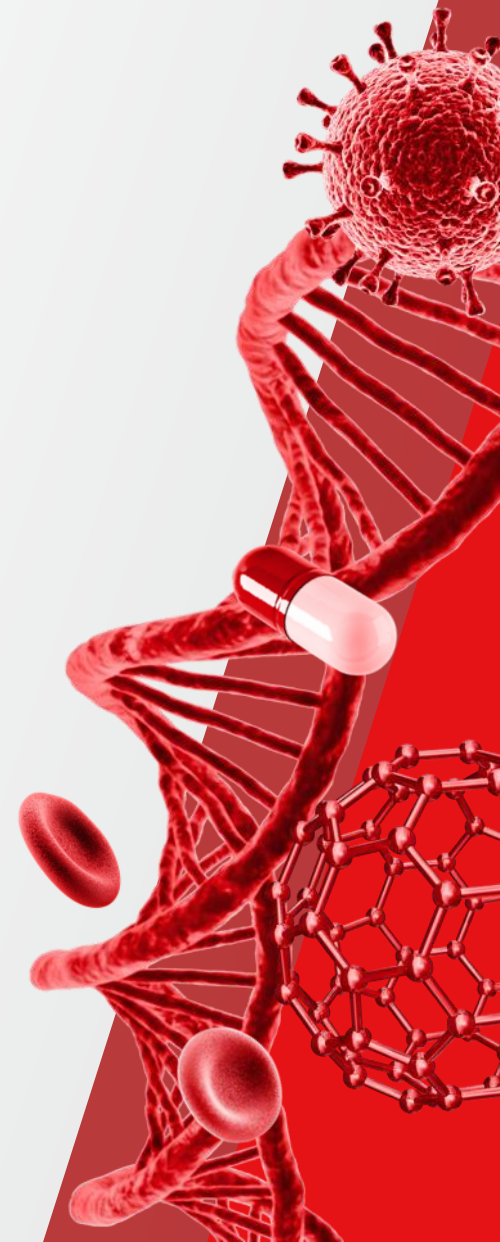
- Which acquisition strategy is most appropriate?
- MS2 30k Turbo scan
- MS2 50k
- MS3 50k
- RTS, MS3 50k
- RTC CO, MS3 50k
- Caveat
 - Protein ID being the only figure of merit
 - Quan accuracy and precision not considered

Data courtesy of Jenny Ho (TFS Hemel Hempstead)



Reporter Ions Quantifier node PSM Filters for quantification

 The world leader in serving science



Reporter Ions Quantifier node

Discussion of parameter settings

- **Co-isolation threshold** – tries to minimize ratio compression due to chimeric spectra
- **Average reporter S/N threshold** – avoids issues due to ion statistics
- **SPS mass matches threshold** – considers number of relevant ions selected from MS2
- **Minimum channel occupancy** - only peptides with at least this percentage of non-zero channel values would be considered for quantification

Parameters of 'Reporter Ions Quantifier'	
Show Advanced Parameters	
1. General Quantification Settings	
Peptides to Use	Unique + Razor
Consider Protein Groups for Peptide Uniqueness	True
Use Shared Quan Results	True
Reject Quan Results with Missing Channels	False
2. Reporter Quantification	
Reporter Abundance Based On	Automatic
Apply Quan Value Corrections	False
Co-Isolation Threshold	50
Average Reporter S/N Threshold	10
SPS Mass Matches [%] Threshold	65
Minimum Channel Occupancy [%] Threshold	0
3. Normalization and Scaling	
Normalization Mode	Total Peptide Amount
Proteins For Normalization	
Scaling Mode	On All Average
4. Exclude Peptides from Protein Quantification	
For Normalization	Use All Peptides
For Protein Roll-Up	Use All Peptides
For Pairwise Ratios	Exclude Modified
1. Considered Peptide Modification	None
2. Considered Peptide Modification	None
3. Considered Peptide Modification	None
N-Terminal Considered Peptide Modification	None
5. Quan Rollup and Hypothesis Testing	
Protein Ratio Calculation	Protein Abundance Based
Maximum Allowed Fold Change	100
Imputation Mode	None
Hypothesis Test	ANOVA (Individual Proteins)

Reporter Ions Quantifier - Filters

- Co-isolation threshold

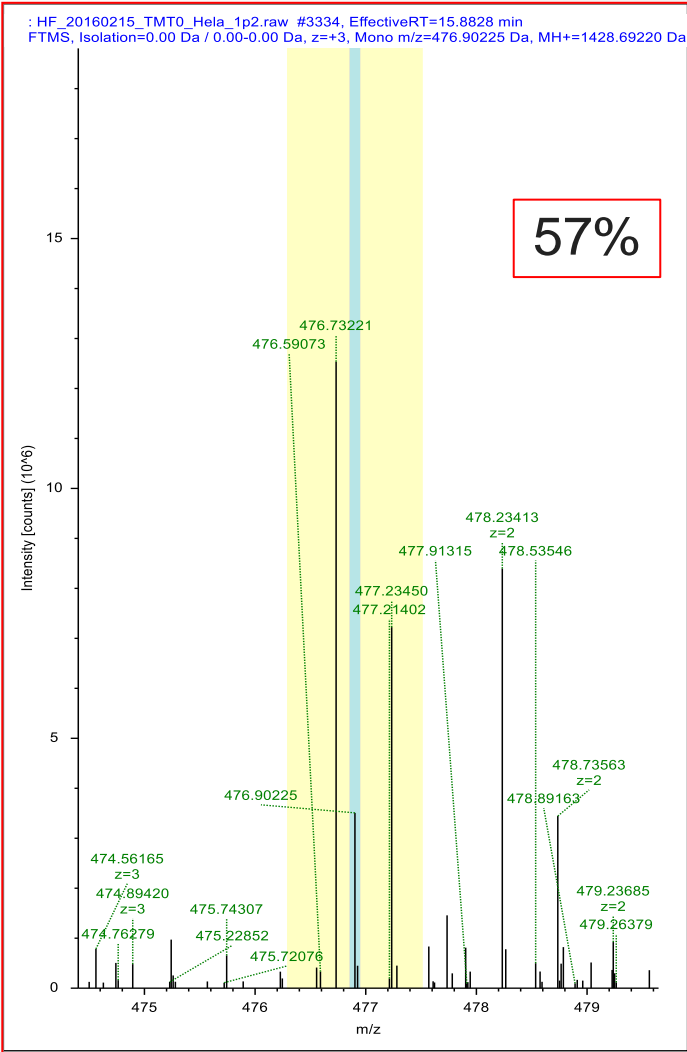
$$\%_{isolation_interference} = 100 \times \left[1 - \left(\frac{precursor_intensity_in_isolation_window}{total_intensity_in_isolation_window} \right) \right]$$

- If the contribution of unwanted peaks in the MS1 isolation window exceeds the threshold, this filter will exclude MS2 or a dependent MS3 spectrum from contributing to quantification

Reporter Ions Quantifier - Filters

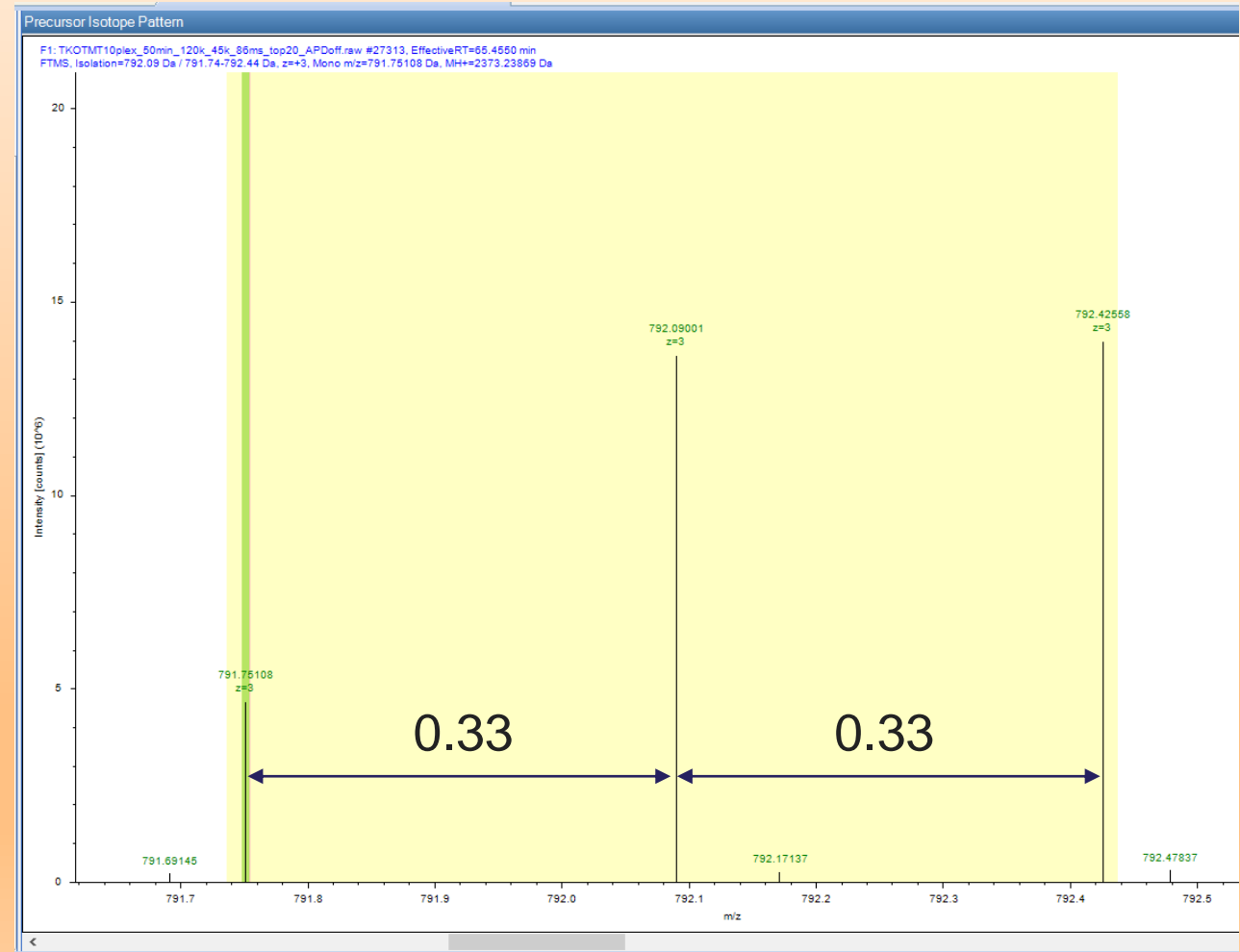
- Co-isolation threshold
- This PSM would be “*excluded by method*”

Parameters of 'Reporter Ions Quantifier'	
Show Advanced Parameters	
▼ 1. General Quantification Settings	
Peptides to Use	Unique
Consider Protein Groups for Peptide Uniqueness	True
Use Shared Quan Results	True
Reject Quan Results with Missing Channels	False
▼ 2. Reporter Quantification	
Reporter Abundance Based On	Automatic
Apply Quan Value Corrections	True
Co-Isolation Threshold	50
Average Reporter S/N Threshold	10
SPS Mass Matches [%] Threshold	65
▼ 3. Normalization and Scaling	



Q: Is there a bug in co-isolation threshold calculation?

- Isotopes of the selected parent ion **sometimes** considered “contaminants”
 - Example 3+ peptide dYGWTQTSLDDYPk
 - isolation interference reported = 86%
-
- Use isolation width <0.65 Da



Reporter Ions Quantifier - Filters

- **Average Reporter S/N Threshold**
- “Magic” number = 10
- Sum of S/N for all reporter ions in the spectrum = “10” x plex of the method

Parameters of 'Reporter Ions Quantifier'	
Show Advanced Parameters	
▼ 1. General Quantification Settings	
Peptides to Use	Unique
Consider Protein Groups for Peptide Uniqueness	True
Use Shared Quan Results	True
Reject Quan Results with Missing Channels	False
▼ 2. Reporter Quantification	
Reporter Abundance Based On	Automatic
Apply Quan Value Corrections	True
Co-Isolation Threshold	50
Average Reporter S/N Threshold	10
SPS Mass Matches [%] Threshold	65
▼ 3. Normalization and Scaling	

Example TMT 6plex

Sum of reporters S/N > 10 x 6

Example TMT 11plex

Sum of reporters S/N > 10 x 11

Reporter Ions Quantifier - Filters

- **SPS Mass Matches [%] Threshold**

- Only for **MS3**-based quan
- Min. percentage of MS2 fragments selected for MS3 fragmentation attributable to the precursor ion

Parameters of 'Reporter Ions Quantifier'	
Show Advanced Parameters	
▼ 1. General Quantification Settings	
Peptides to Use	Unique
Consider Protein Groups for Peptide Uniqueness	True
Use Shared Quan Results	True
Reject Quan Results with Missing Channels	False
▼ 2. Reporter Quantification	
Reporter Abundance Based On	Automatic
Apply Quan Value Corrections	True
Co-Isolation Threshold	50
Average Reporter S/N Threshold	10
SPS Mass Matches [%] Threshold	65
▼ 3. Normalization and Scaling	

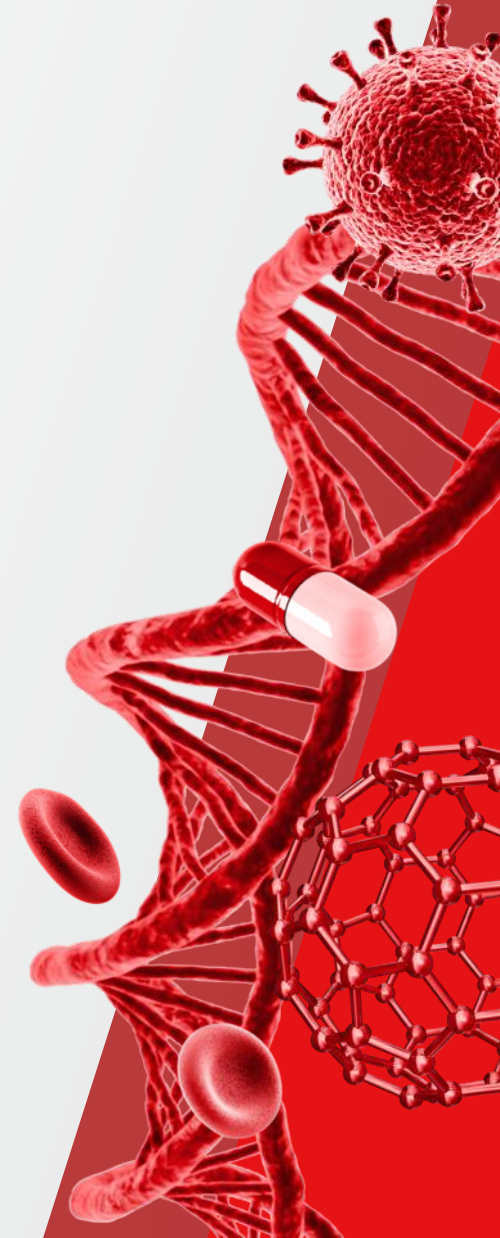
Q: Is there a bug in SPS Mass Matches Code?

TMT-labeled phosphopeptides

- For SPS MS3 data → SPS Mass Matches [%] Threshold = 0
- SPS Mass Match [%] Threshold considers MS1 precursor mass to check whether MS2 fragments belong to the peptide of interest
- The precursor carries the phosphate but most MS2 fragments are likely to contain dephosphorylated residues (dehydroalanine instead of phosphoserine)
- PD concludes that such fragments are not from the peptide of interest → Phosphopeptide MS3 spectra are unlikely to pass the filter...

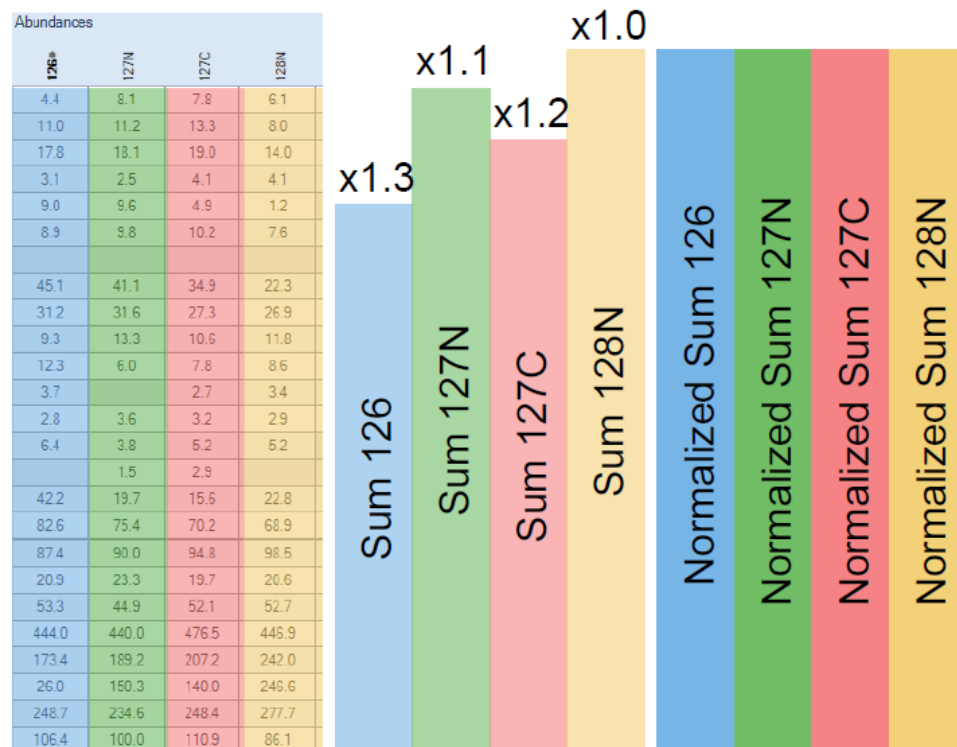
Parameters of 'Reporter Ions Quantifier'	
Show Advanced Parameters	
▼ 1. General Quantification Settings	
Peptides to Use	Unique + Razor
Consider Protein Groups for Peptide Uniqueness	True
Use Shared Quan Results	True
Reject Quan Results with Missing Channels	False
▼ 2. Reporter Quantification	
Reporter Abundance Based On	Automatic
Apply Quan Value Corrections	False
Co-Isolation Threshold	50
Average Reporter S/N Threshold	10
SPS Mass Matches [%] Threshold	0
Minimum Channel Occupancy [%] Threshold	0
▼ 3. Normalization and Scaling	

Reporter Ions Quantifier node Normalization and Scaling



Normalization

- Total peptide amount
 - Sum reporter ions per channel
 - Calculate normalization factor



Normalization Mode

Specifies how the normalization shall be performed to correct for experimental bias.

None: No normalization is applied.

Total Peptide Amount: This calculates the total sum of the abundance values for each channel over all peptides identified within a file. It then takes the channel with the highest total abundance as a reference and corrects all abundance values in all other channels by a constant factor per channel, so that at the end the total abundance is the same for all channels.

Specific Protein Amount: In this mode, the normalization only looks at the summed abundances of proteins in the specified FASTA file. This is useful if you know there are housekeeping proteins, or a bait protein from pull-downs, or similar proteins for which you can assume their abundance remaining constant across your treatments. Besides that the normalization is done similar to the normalization to total peptide amount.

Normalization

- Applied normalization values
- Result Summaries → Quantification → Derived Values

The screenshot shows the 'Result Summaries' window with the 'Quantification' tab selected. The 'Derived Quantification Values' section is visible, showing normalization settings and a list of applied normalization values. The 'Applied Normalization Values' list is highlighted with an orange box.

Result Summaries

Copy All Copy Section Copy Subsection

Samples & Files Analysis Settings Validation **Quantification** Configuration

Derived Quantification Values

Consensus Step
Processing Step A
Derived Values
Reported Ratios

Normalization/ Scaling:

Normalization Mode = Total Peptide Amount
Protein For Normalization = No FASTA File Selected
Scaling Mode = On All Average

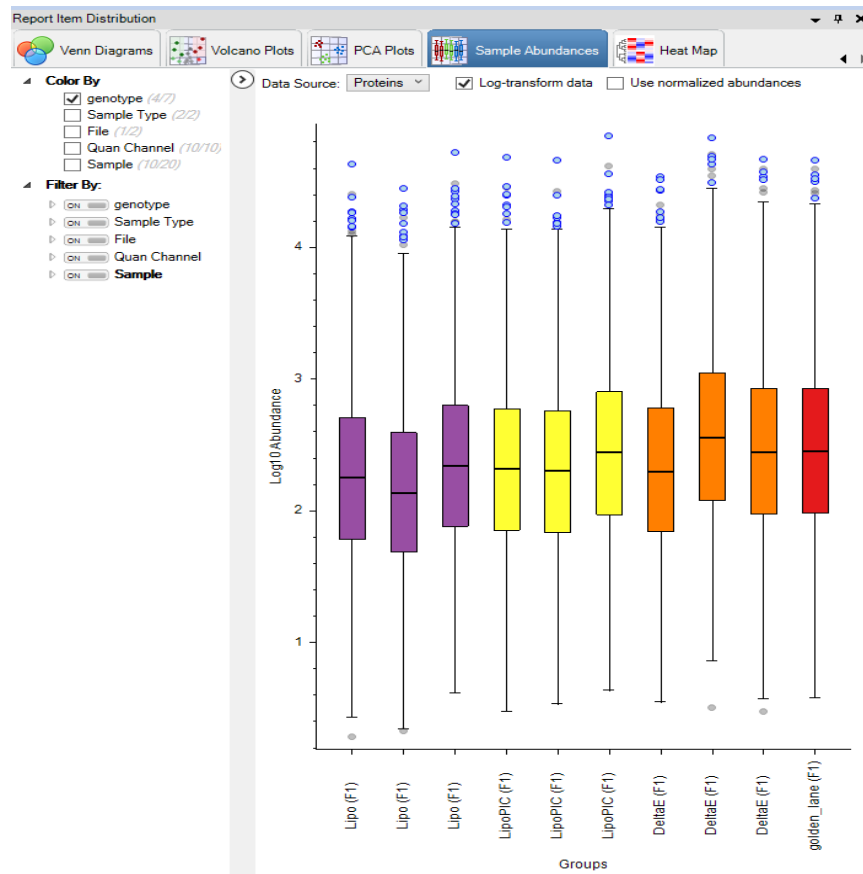
Applied Normalization Values:

- 1.015: F1: 128N, Sample, his4
- 1.006: F1: 128C, Sample, his4
- 1.015: F1: 129N, Sample, his4
- 1.063: F1: 126, Sample, met6
- 1.067: F1: 127N, Sample, met6
- 1.006: F1: 127C, Sample, met6
- 1.011: F1: 129C, Sample, ura2
- 1.000: F1: 130N, Sample, ura2
- 1.045: F1: 130C, Sample, ura2
- 1.063: F1: 131, Sample, wildtype

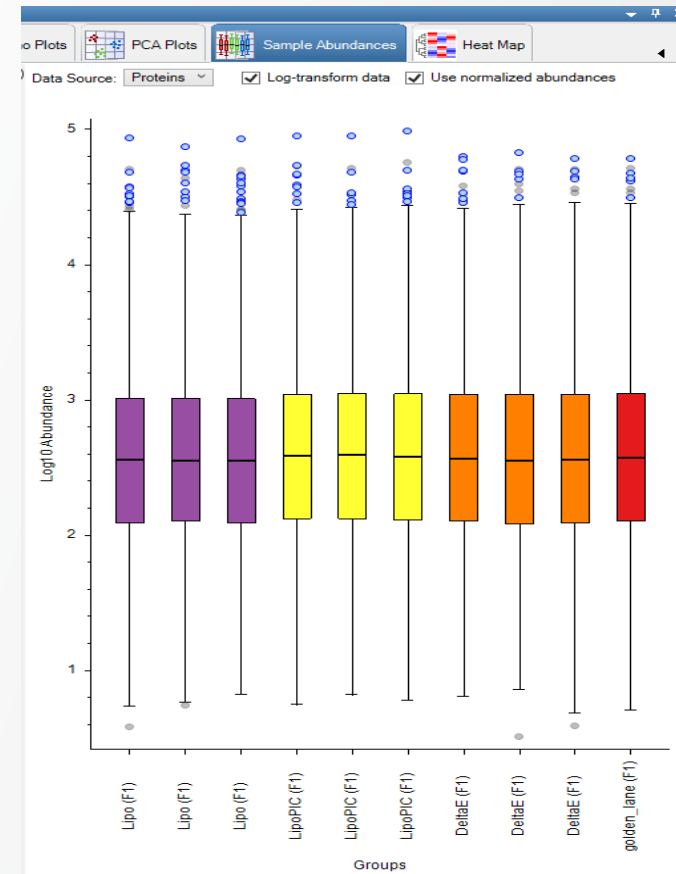
Normalization

- Check how normalization worked
- View → Distribution Charts → Sample Abundances

BEFORE

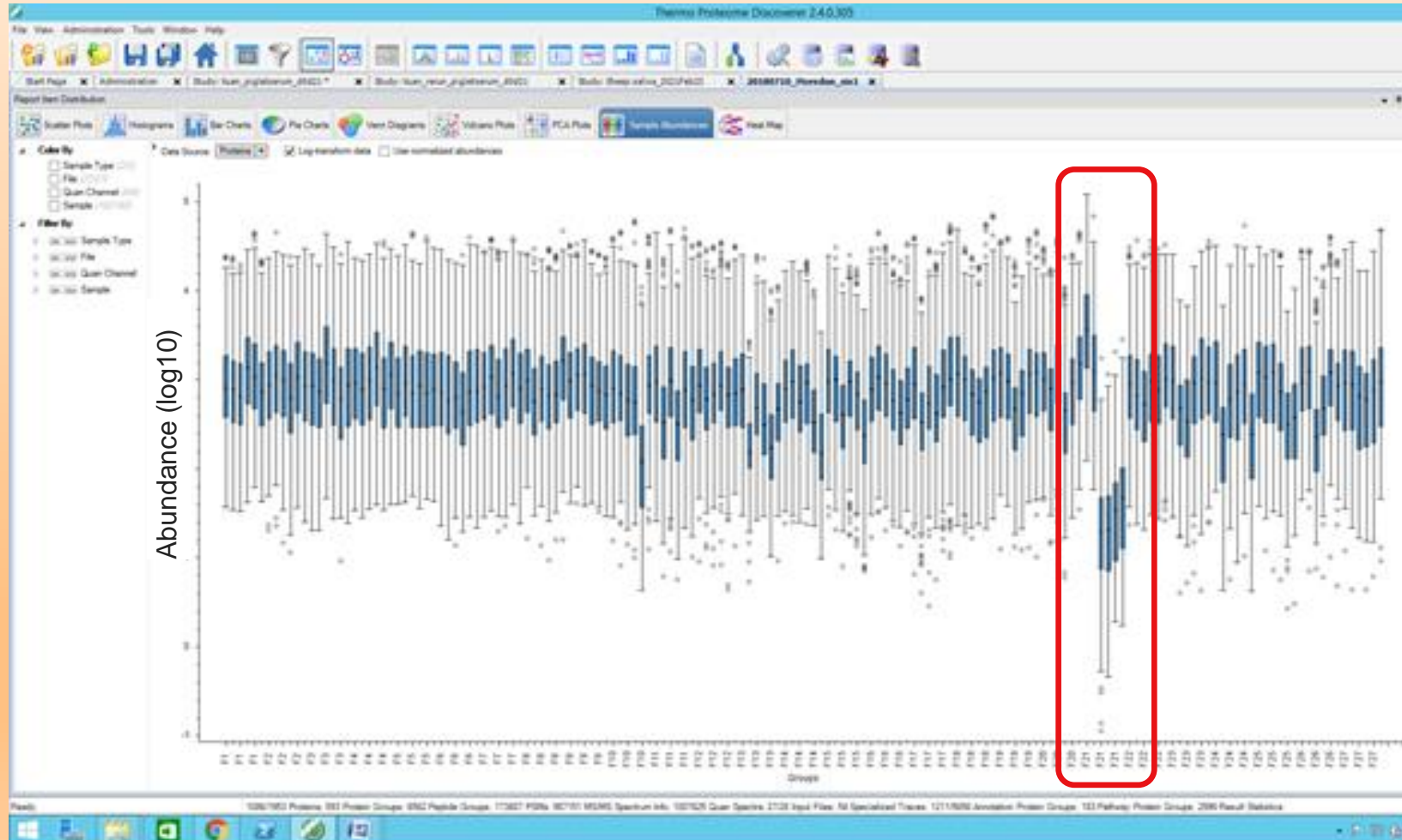


AFTER



Q: What went wrong ?

- View → Distribution Charts → Sample Abundances
- Only 2-3 fold difference in sample load acceptable



Multiplexed TMT Experiment

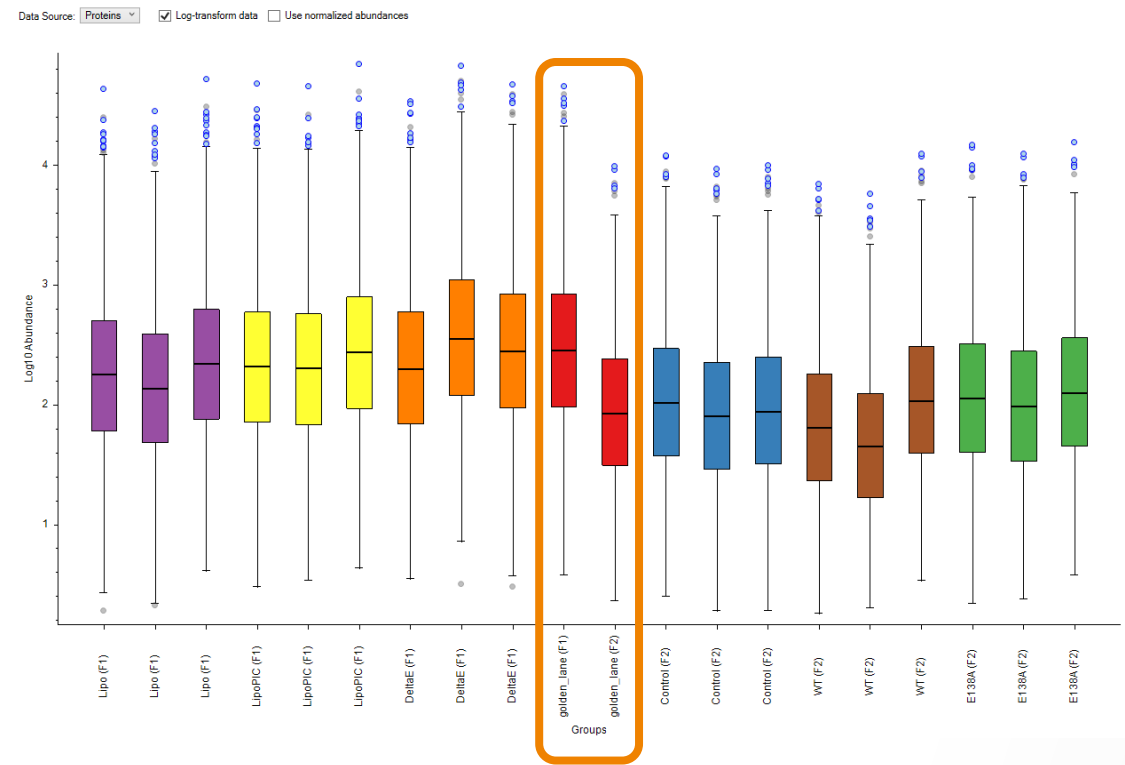
- The number of samples exceeds the number of available channels
 - The same label will be used for two or more different samples
 - Those will be analyzed in separate “sets”
- *Example of a multiplexed experiment:*
 - You have 6 conditions with 3 samples each (total 18 samples)
 - You have TMT 10plex kit available
 - Split sample to two batches; 9 samples each
 - Prepare a bridge sample (e.g., “pooled” sample, run as the 10th channel)
 - Run 2 TMT 10plex sets
 - *Both runs should be using the same quan method!*

Normalization of Multiplexed TMT Experiment

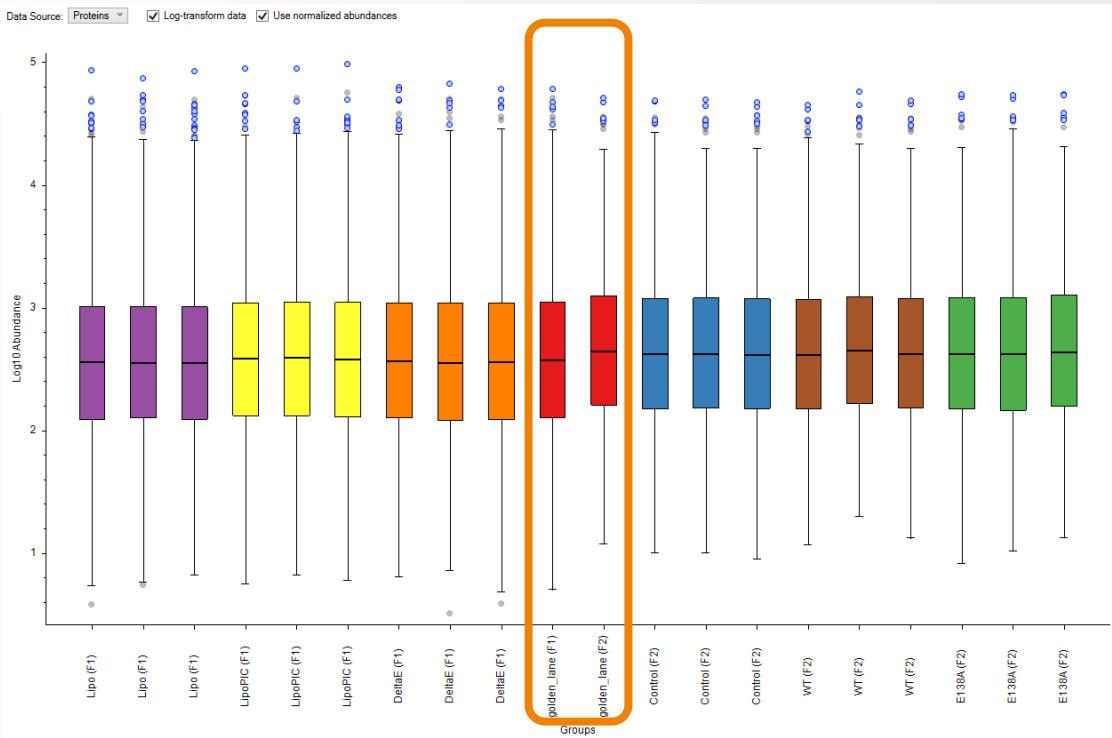


- 2 sets of TMT 10plex

BEFORE

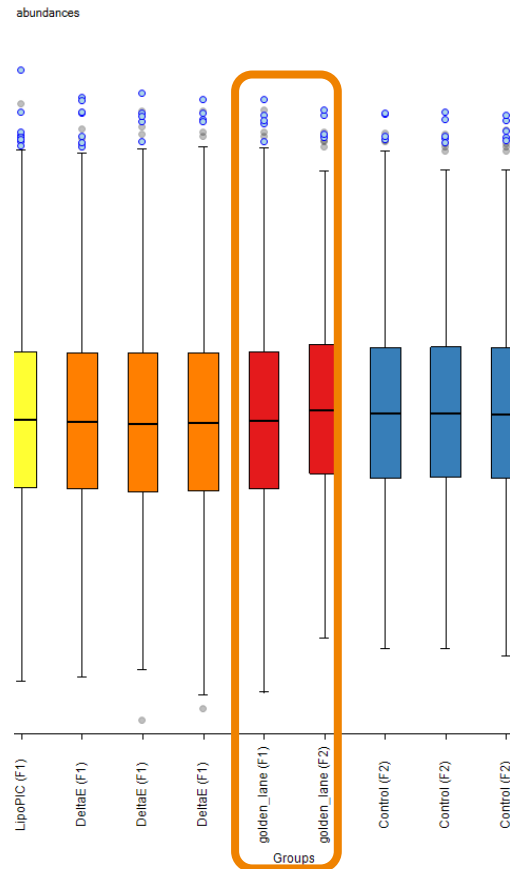


AFTER



Normalization of Multiplexed TMT Experiment

- *Note:* the median abundance of the two bridge samples is not identical
- The normalization works only within the same set, not across different sets



Scaling

Use the bridge channels to normalize across multiple sets

- Set Scaling **On controls average**

Parameters of 'Reporter Ions Quantifier'	
Hide Advanced Parameters	
1. General Quantification Settings	
Peptides to Use	Unique + Razor
Consider Protein Groups for Peptide Unique	True
Use Shared Quan Results	True
Reject Quan Results with Missing Channels	False
2. Reporter Quantification	
Reporter Abundance Based On	Automatic
Apply Quan Value Corrections	False
Co-Isolation Threshold	50
Average Reporter S/N Threshold	10
SPS Mass Matches [%] Threshold	65
Minimum Channel Occupancy [%] Threshold	0
3. Normalization and Scaling	
Normalization Mode	Total Peptide Amount
Proteins For Normalization	
Scaling Mode	On Controls Average

Define the “controls” in Samples table

Add Files

Add Fractions

Remove Files

Open Containing Folder

New Analysis

Study Definition

Input Files

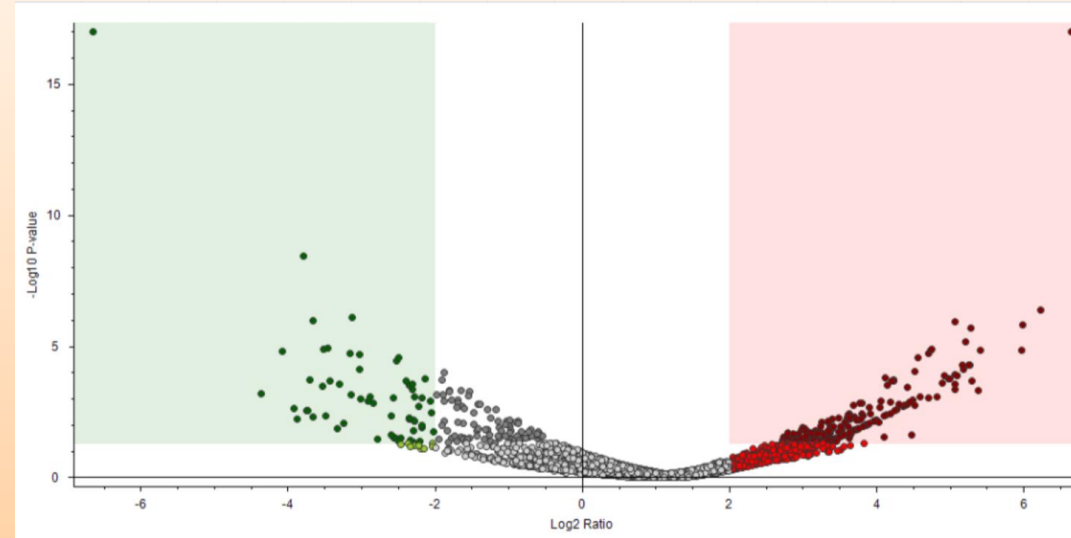
Samples

Analysis Results

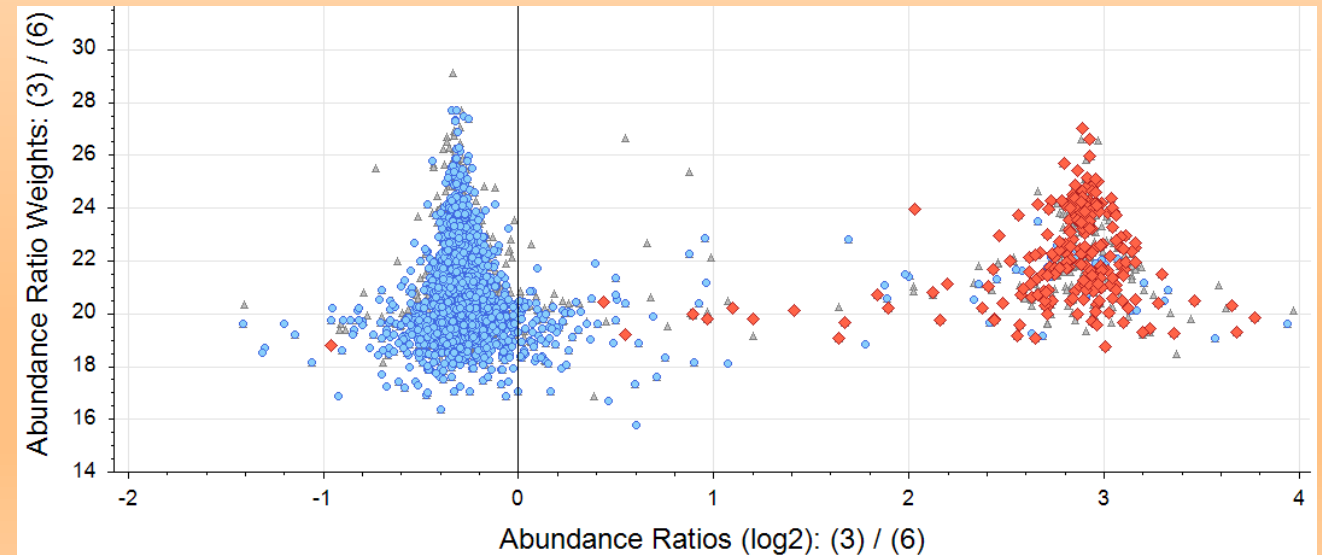
Error	Sample	File	Sample Identifier	Sample Type	Quan Chan
	S1	F1	BDL_TMT_SLD_MY_TD_01 - [113]	Control	113
	S2	F2	BDL_TMT_SLD_MY_TD_02 - [113]	Control	113
	S3	F3	BDL_TMT_SLD_MY_TD_03 - [113]	Control	113
	S4	F4	BDL_TMT_SLD_MY_TD_04 - [113]	Control	113
	S5	F1	BDL_TMT_SLD_MY_TD_01 - [114]	Sample	114
	S12	F2	BDL_TMT_SLD_MY_TD_02 - [114]	Sample	114
	S19	F3	BDL_TMT_SLD_MY_TD_03 - [114]	Sample	114
	S26	F4	BDL_TMT_SLD_MY_TD_04 - [114]	Sample	114
	S6	F1	BDL_TMT_SLD_MY_TD_01 - [115]	Sample	115
	S13	F2	BDL_TMT_SLD_MY_TD_02 - [115]	Sample	115
	S20	F3	BDL_TMT_SLD_MY_TD_03 - [115]	Sample	115
	S27	F4	BDL_TMT_SLD_MY_TD_04 - [115]	Sample	115

Q: Have I selected a correct normalization method ?

Check the volcano plot

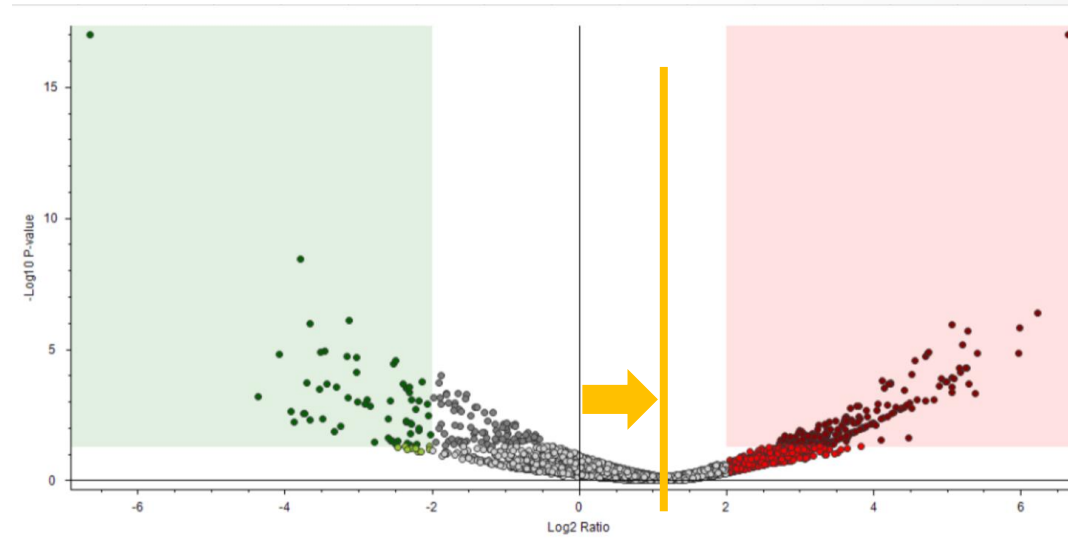


or even better, the plot of
Abundance Ratio Weights x
Abundance ratios (log2)

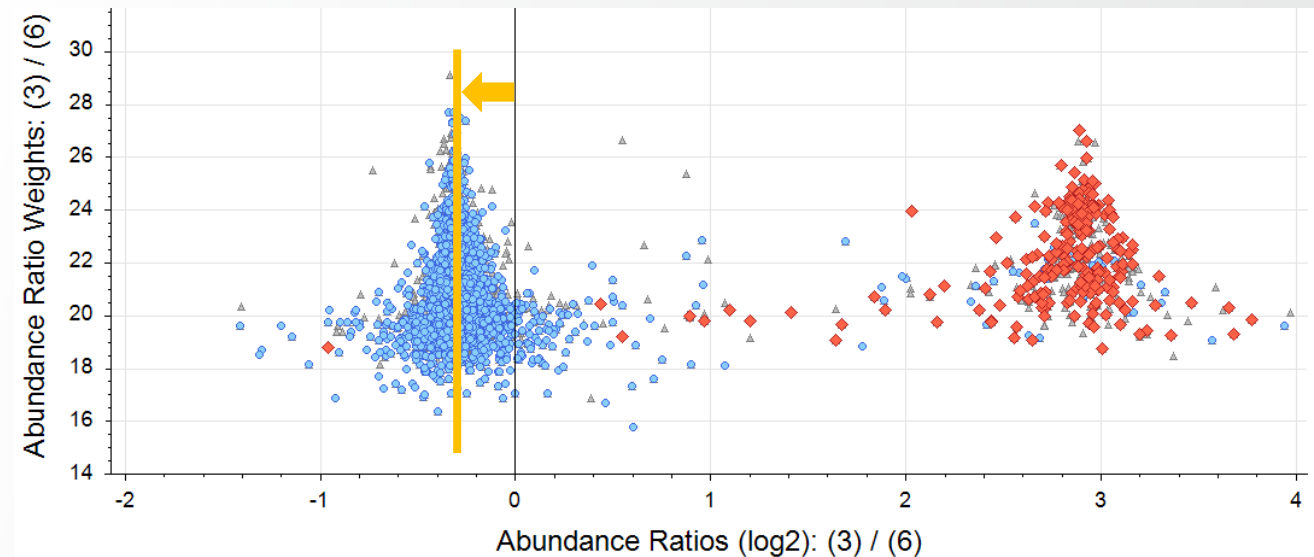


Q: Have I selected a correct normalization method ?

Check the volcano plot



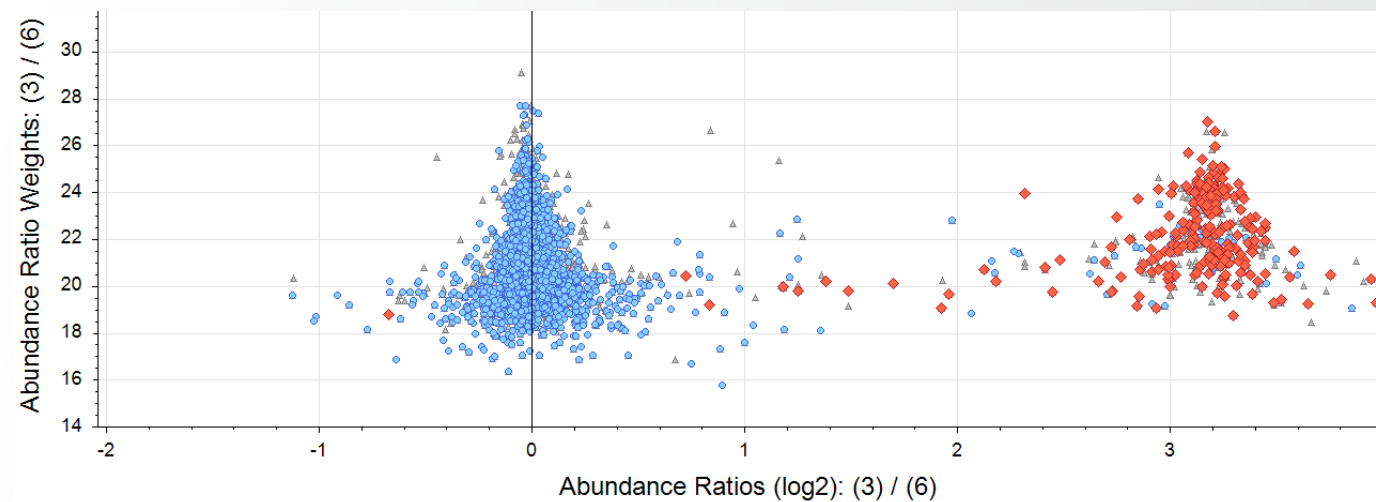
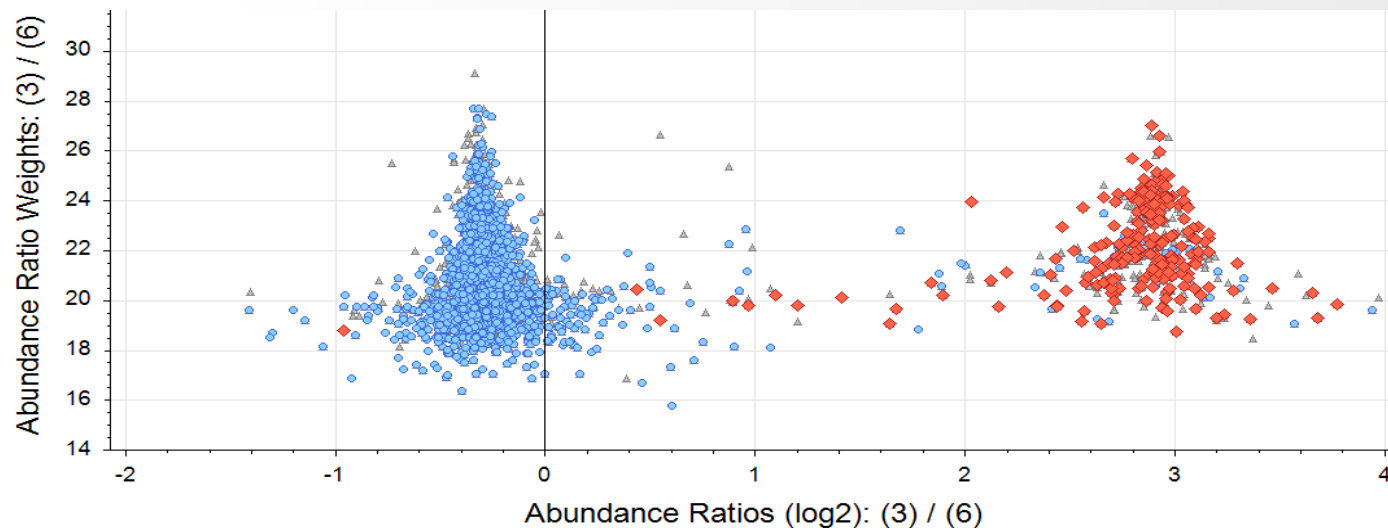
or even better, the plot of
Abundance Ratio Weights x
Abundance ratios (log2)



Different Proteins Used for Normalization

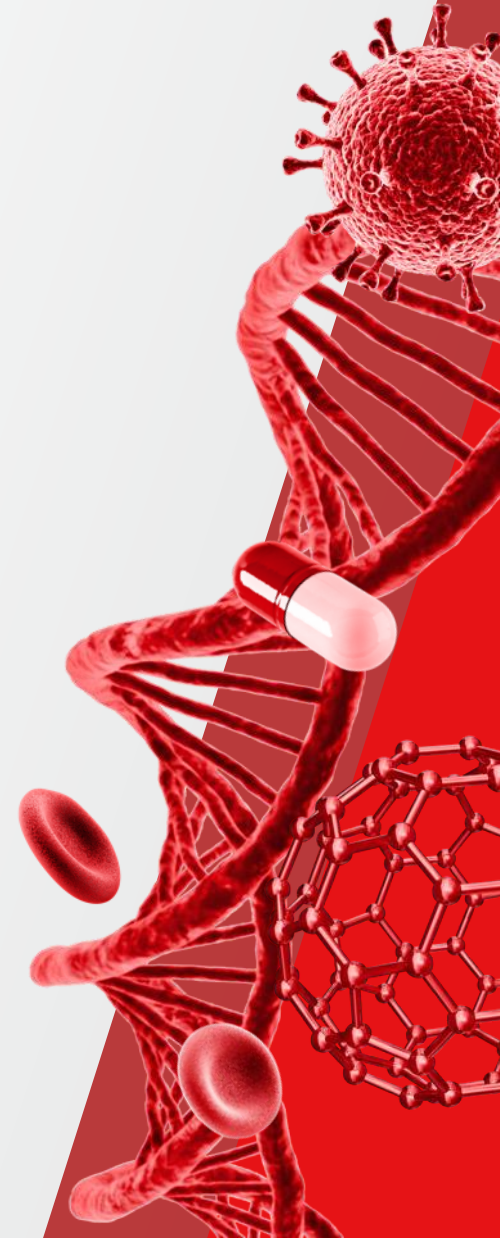
Mixed proteomes dilution experiment: identical human background; different bacterial spike-in

- Normalized **on total peptide amount**
 - The background protein distribution (blue = human) is not centered at $\log_2=0$
 - The total protein amount in the two samples is different
- Normalized on **specific protein amount**
 - “human.FASTA” db used
 - Normalized on all human peptides (amount in both samples identical)



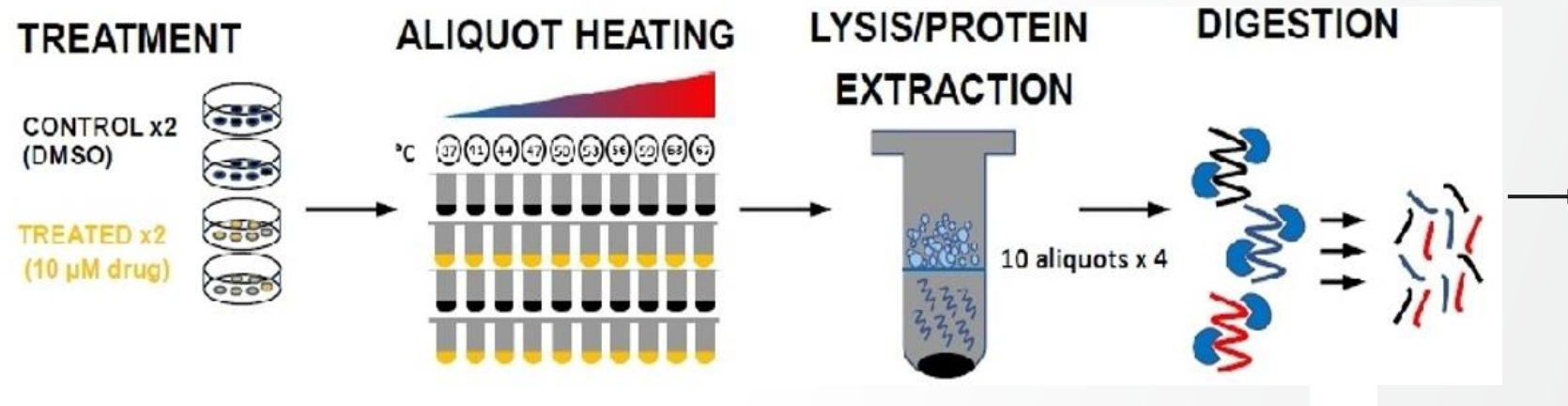
Multiplexed TMT - Experiment Setup

 The world leader in serving science



Multiplexed TMT11plex Experiment

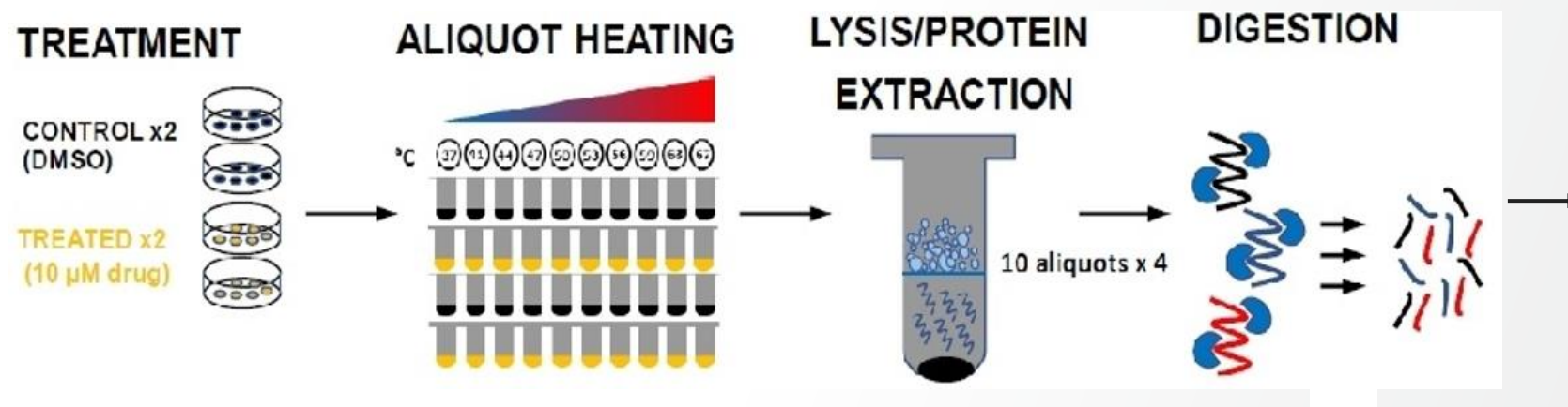
- 2 treatments (DMSO; drug)
- 2 biological replicates
- 10 temperature points (channels 126-131N)
- 1 “pool” sample (bridge channel 131C)



T (°C)	Channel
37	126
41	127N
44	127C
47	128N
50	128C
53	129N
56	129C
59	130N
63	130C
67	131N
Pool	131C

Multiplexed TMT11plex Experiment

- To accommodate all samples, we need **4x TMT11plex sets**
- Each set **fractionated** (2 fractions each)
- Each fraction **run in triplicate** **=> Total 24 .raw files**

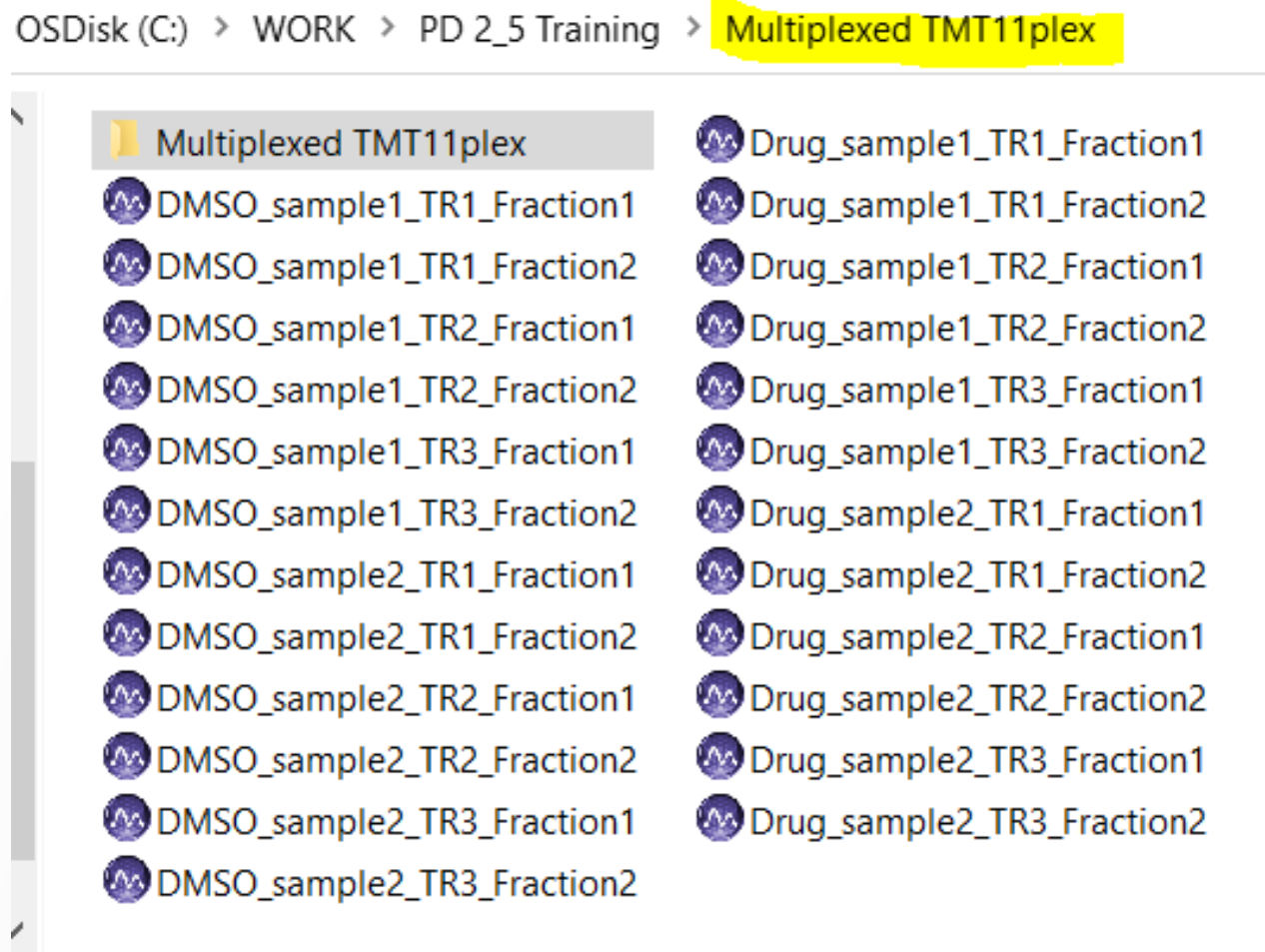


T (°C)	Channel
37	126
41	127N
44	127C
47	128N
50	128C
53	129N
56	129C
59	130N
63	130C
67	131N
Pool	131C

Multiplexed TMT11plex

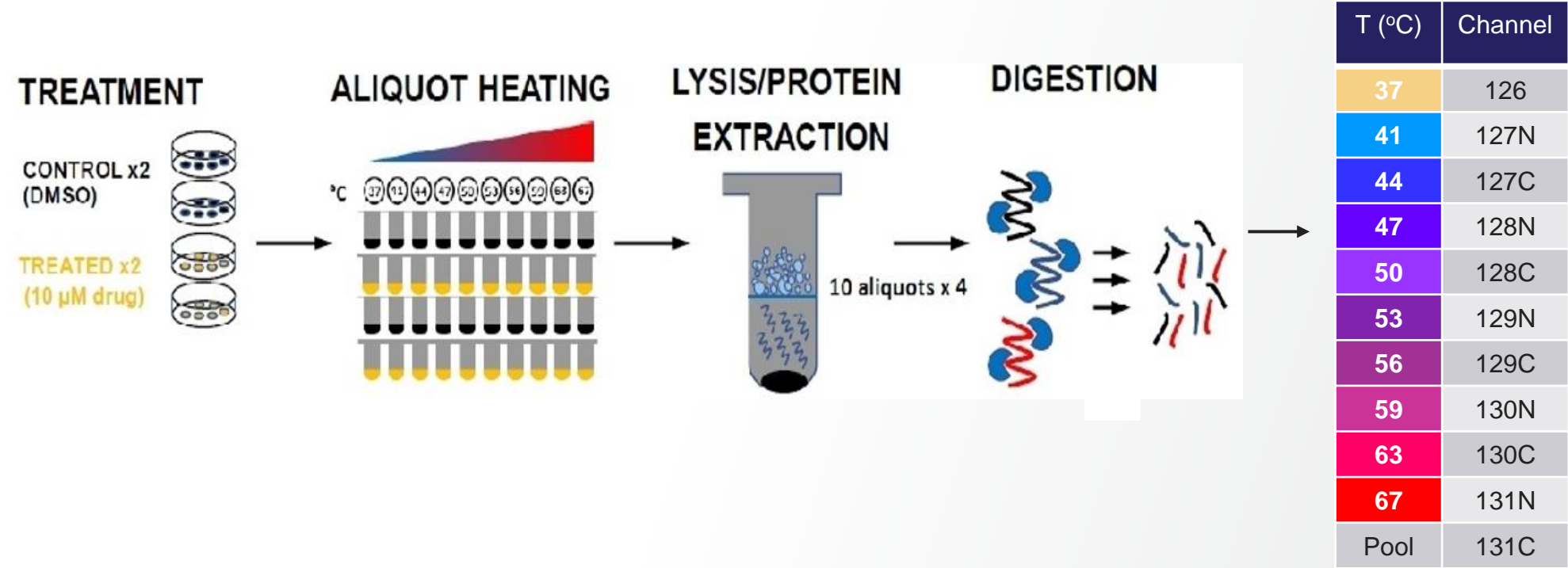
Data files

- Does not contain real data
- To be used just to set up a Study / Analysis
- Do not process the data!

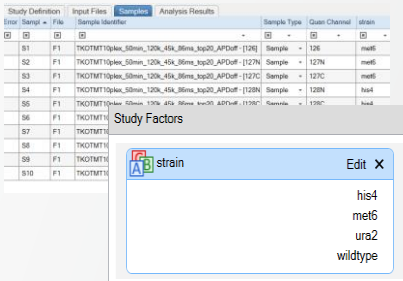
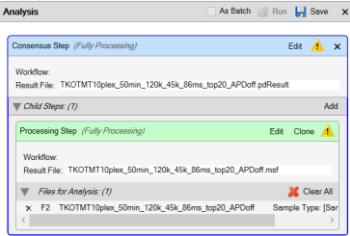
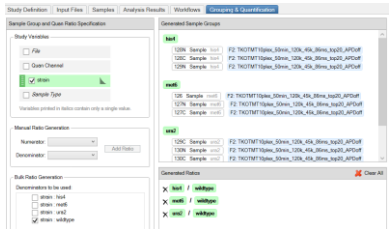
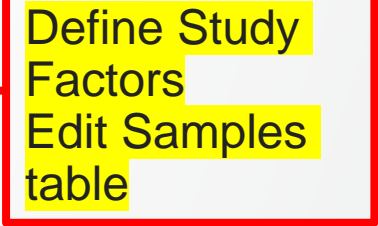
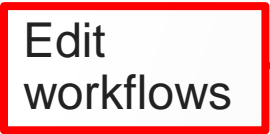
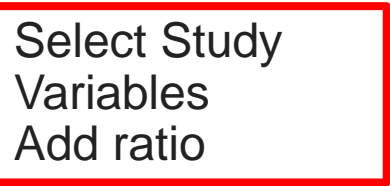
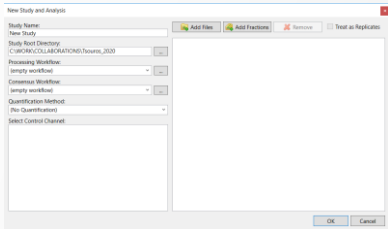
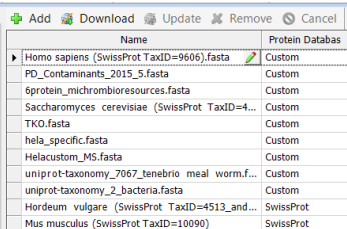
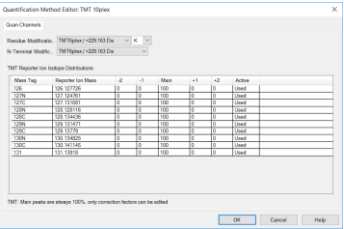
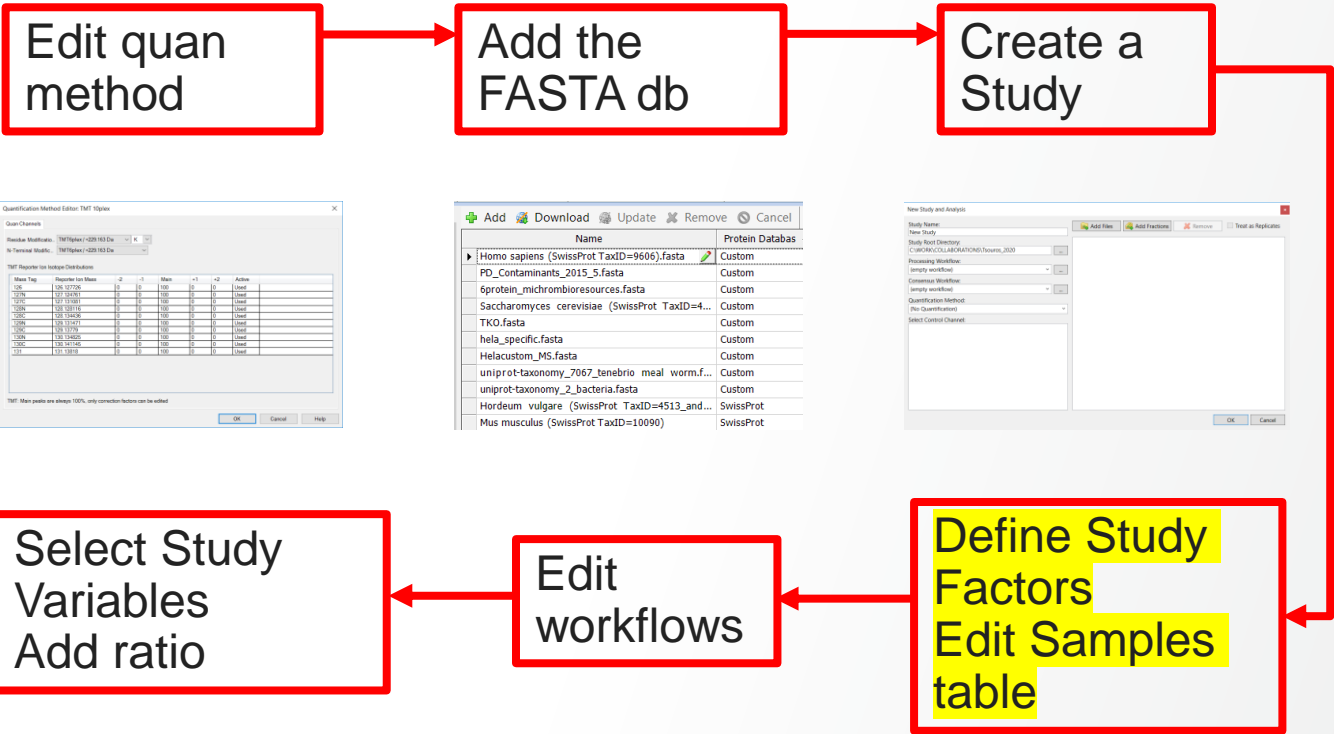


Questions being asked?

- Effect of the drug treatment on protein expression at various temperatures



General Quan Experiment Flowchart



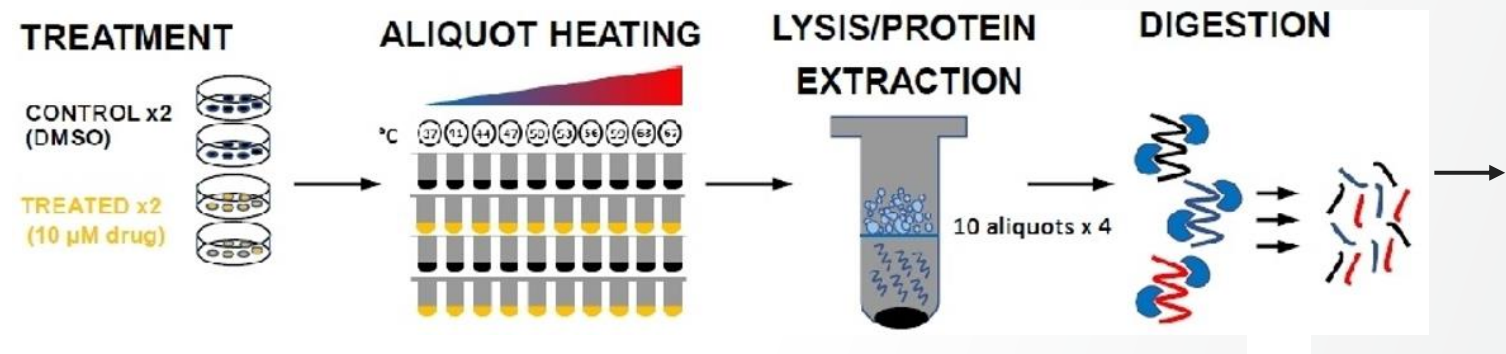
Nested vs Non-nested?

- Are the differences between the two samplings of the same treatment important?

Should NOT be



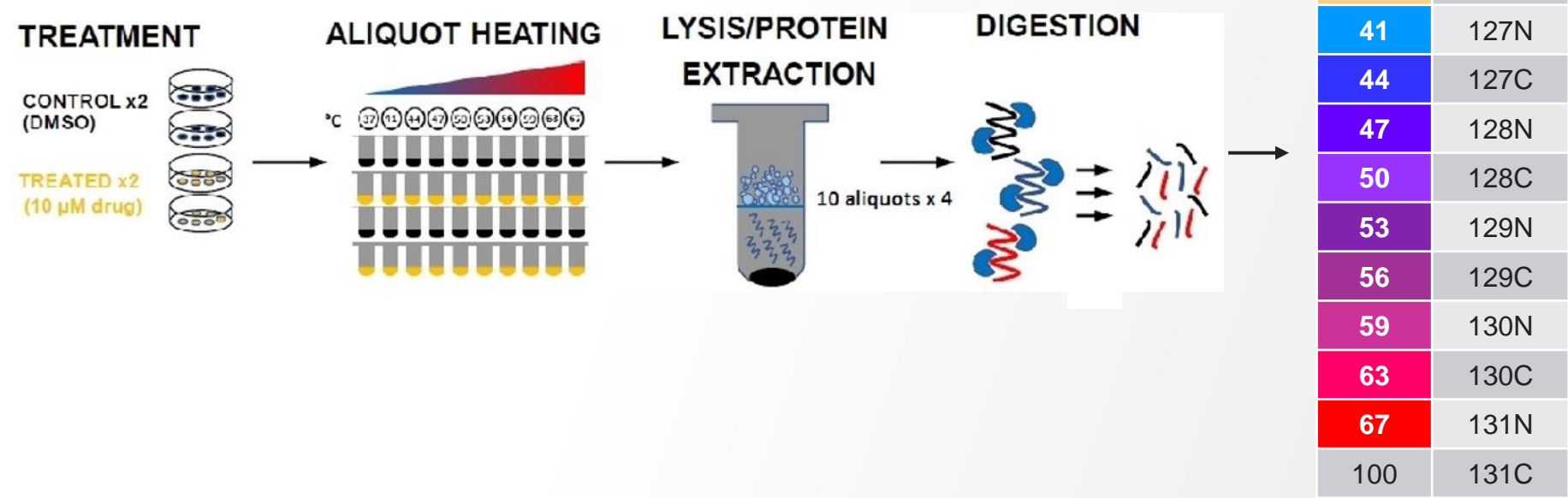
Non-nested design



T (°C)	Channel
37	126
41	127N
44	127C
47	128N
50	128C
53	129N
56	129C
59	130N
63	130C
67	131N
Pool	131C

Study factors

- **Treatment** (DMSO; drug; pool) – *categorical* study factor
- **Biological replicate** (1; 2) – *categorical* study factor
- **Temperature** (37-67; pool) – *numerical* study factor
 - Fractions – use “add fractions”
 - Technical replicate – no dedicated factor needed



Study factors

- Treatment (DMSO; drug; pool)
 - *categorical* study factor
- Biological replicate (1; 2)
 - *Categorical* study factor
- Temperature (37-67)
 - *Numerical* study factor
- The order/hierarchy of study factors can be changed later in Grouping&Quantification table

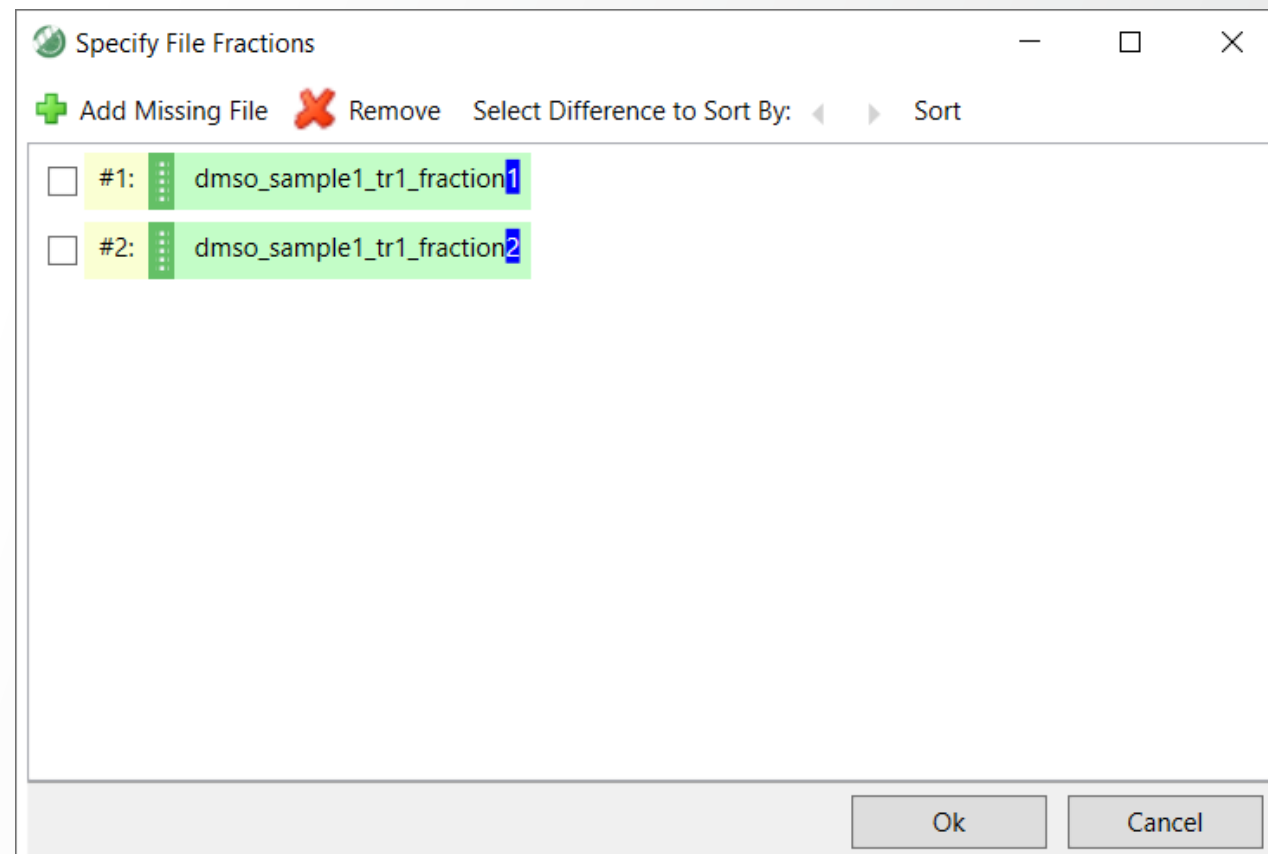
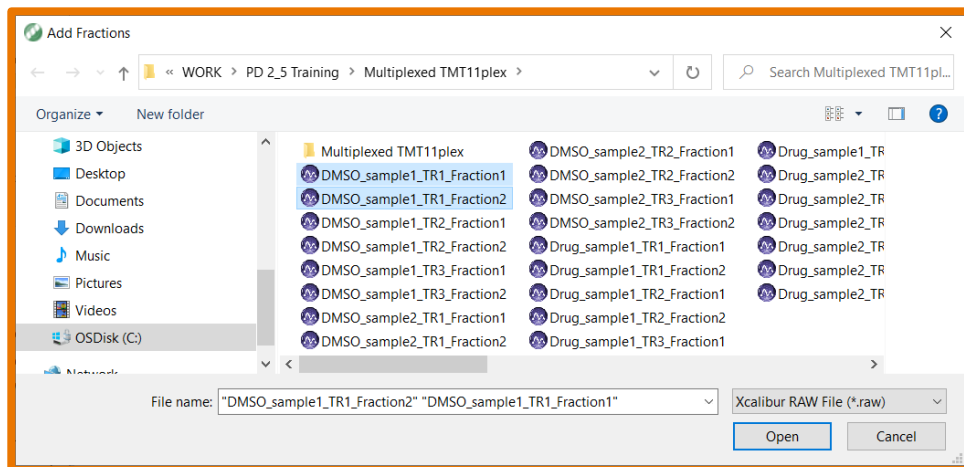
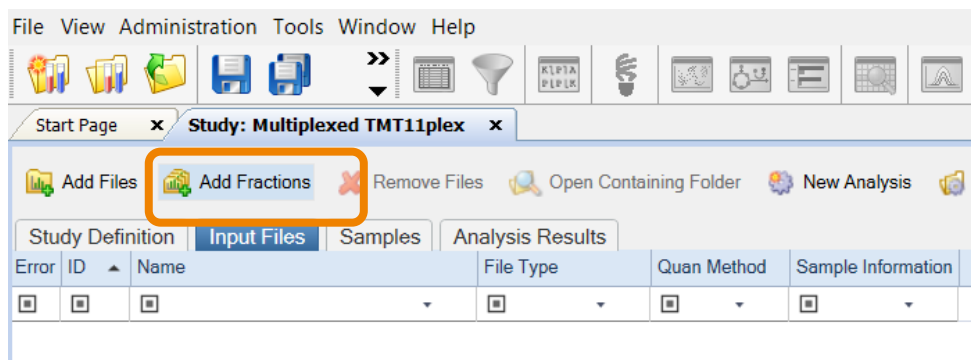
treatment		Edit X
	DMSO	
	Drug	
	Pool	

Biological replicate		Edit X
	1	
	2	

Temperature		Edit X
	37 Celsius	
	41 Celsius	
	44 Celsius	
	47 Celsius	
	50 Celsius	
	53 Celsius	
	56 Celsius	
	59 Celsius	
	63 Celsius	
	67 Celsius	

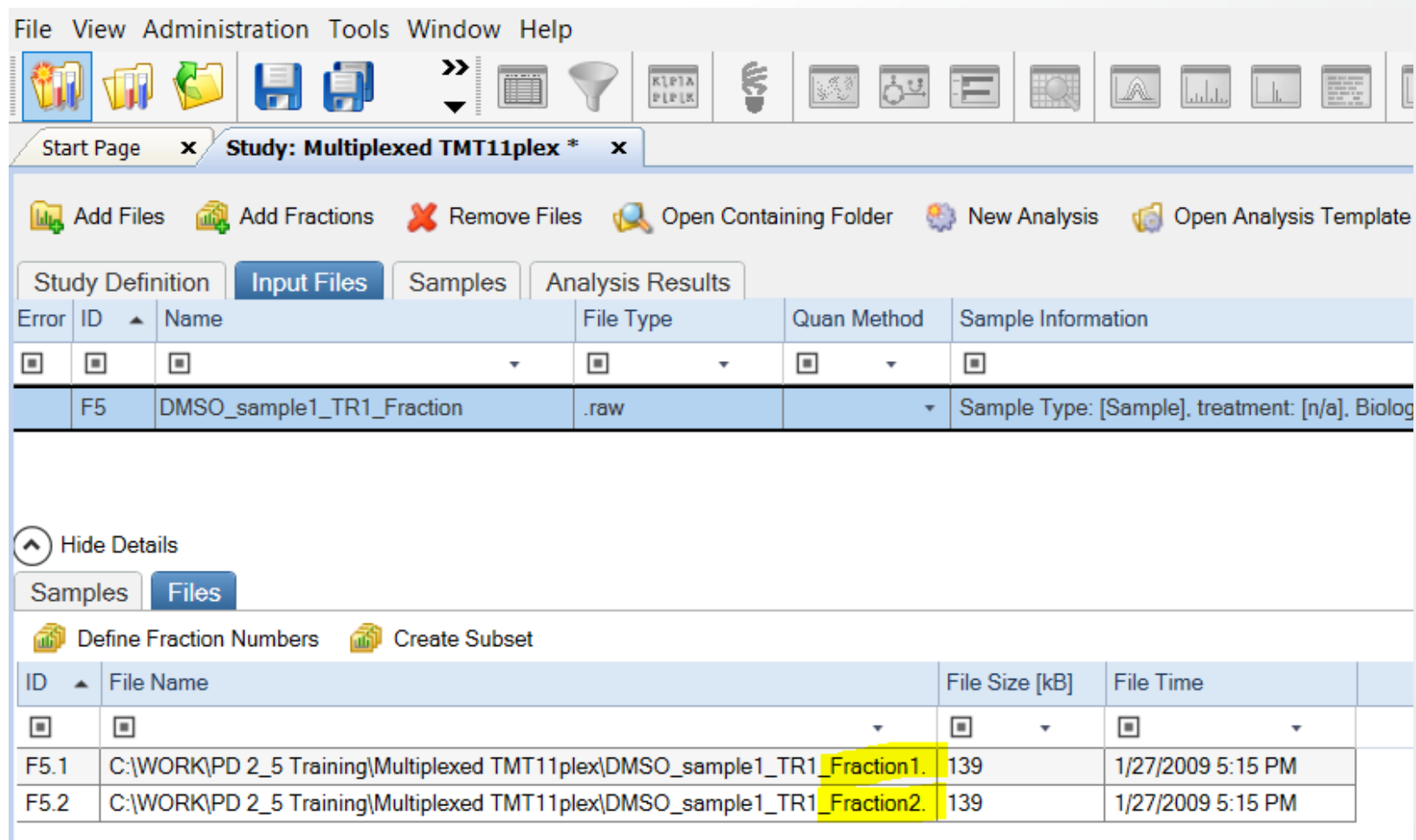
Handling Fractions

- Input Files table
- Add Fractions



Handling Fractions

- **Input Files** table
- View details of added fractions; then add the rest of the files “as fractions”



The screenshot displays the ThermoFisher software interface. The top menu bar includes File, View, Administration, Tools, Window, and Help. Below the menu is a toolbar with various icons. The main window has a tab labeled 'Study: Multiplexed TMT11plex *'. Under this tab, there are buttons for 'Add Files', 'Add Fractions', 'Remove Files', 'Open Containing Folder', 'New Analysis', and 'Open Analysis Template'. The 'Input Files' tab is selected, showing a table with columns: Error, ID, Name, File Type, Quan Method, and Sample Information. The table contains one row with ID 'F5' and Name 'DMSO_sample1_TR1_Fraction'. Below this table, there is a 'Hide Details' button and a 'Samples' tab. The 'Files' tab is selected, showing a table with columns: ID, File Name, File Size [kB], and File Time. This table contains two rows, F5.1 and F5.2, both with a file size of 139 kB and a file time of 1/27/2009 5:15 PM. The file names are 'C:\WORK\PD 2_5 Training\Multiplexed TMT11plex\DMSO_sample1_TR1_Fraction1.' and 'C:\WORK\PD 2_5 Training\Multiplexed TMT11plex\DMSO_sample1_TR1_Fraction2.' respectively. A large orange arrow points from the 'Input Files' table to the 'Files' table.

Error	ID	Name	File Type	Quan Method	Sample Information
	F5	DMSO_sample1_TR1_Fraction	.raw		Sample Type: [Sample], treatment: [n/a], Biolog

Hide Details

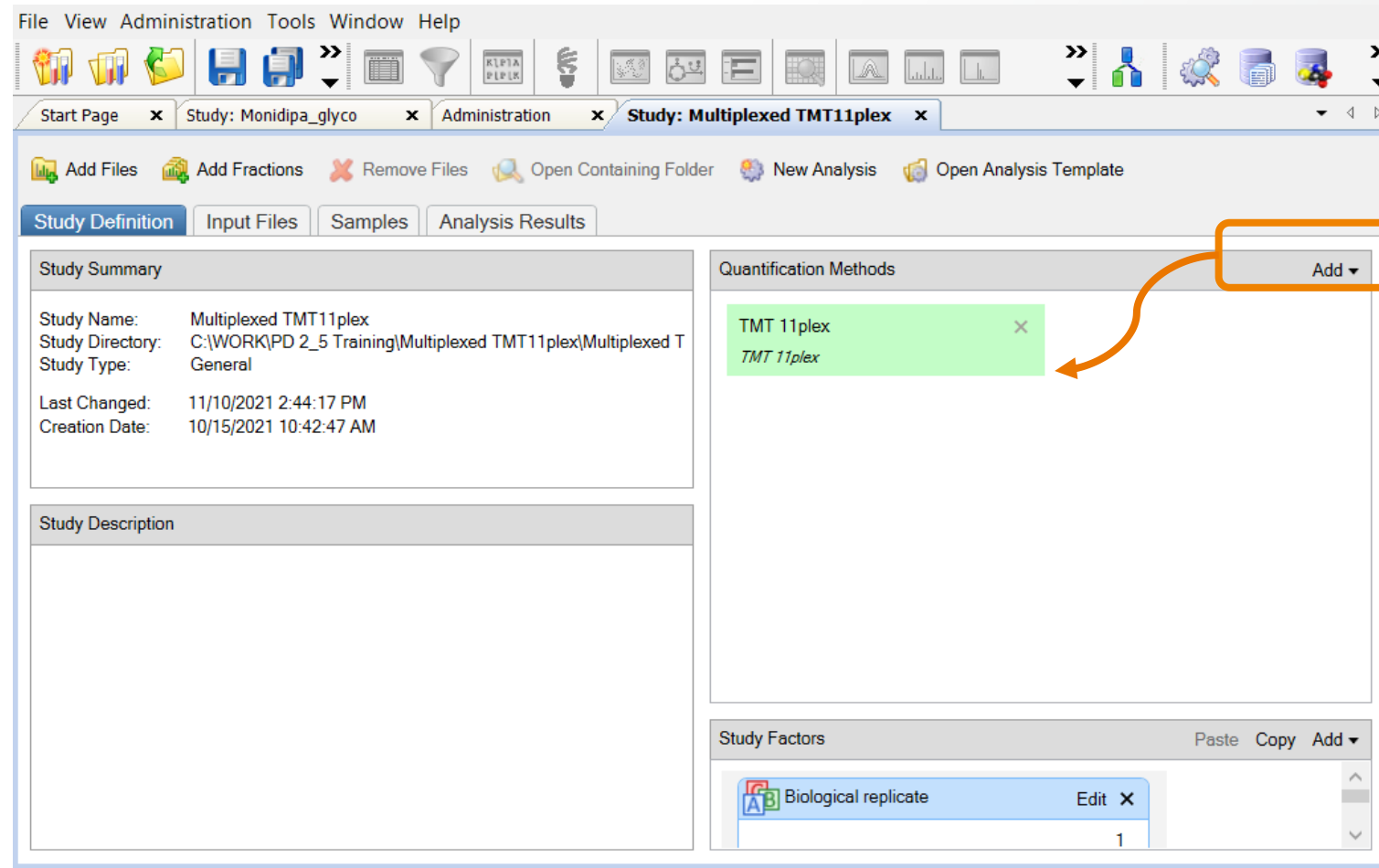
Samples Files

Define Fraction Numbers Create Subset

ID	File Name	File Size [kB]	File Time
F5.1	C:\WORK\PD 2_5 Training\Multiplexed TMT11plex\DMSO_sample1_TR1_Fraction1.	139	1/27/2009 5:15 PM
F5.2	C:\WORK\PD 2_5 Training\Multiplexed TMT11plex\DMSO_sample1_TR1_Fraction2.	139	1/27/2009 5:15 PM

Select quan method

- Using default TMT11plex



Associate files with quan method

- **Input Files** table
- Highlight all the rows
 - Right mouse click → Set quan method to → TMT11plex

The screenshot shows the Thermo Proteome Discoverer 2.5.0.400 interface. The 'Input Files' tab is active, displaying a table with columns: Error, ID, Name, File Type, Quan Method, and Sample Information. The table contains 12 rows of data, with rows F5 through F16 highlighted. A right-click context menu is open over the table, showing options such as 'Copy With Headers', 'Copy', 'Clear Selection', 'Cell Selection Mode', 'Enable Row Grouping', 'Set Quan Method to', 'Set as Input File', and 'Find missing files'. The 'Set Quan Method to' option is selected, and a sub-menu is visible showing 'TMT 11plex' as the chosen method.

Error	ID	Name	File Type	Quan Method	Sample Information
	F5	DMSO_sample1_TR1_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F6	DMSO_sample1_TR2_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F7	DMSO_sample1_TR3_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F8	DMSO_sample2_TR1_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F9	DMSO_sample2_TR2_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F10	DMSO_sample2_TR3_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F11	Drug_sample1_TR1_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F12	Drug_sample1_TR2_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F13	Drug_sample1_TR3_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F14	Drug_sample2_TR1_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F15	Drug_sample2_TR2_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]
	F16	Drug_sample2_TR3_Fraction	.raw	TMT 11plex	Sample Type: [Sample], treatment: [n/a], Biological replicate: [n/a], Temperature: [n/a]

Edit Samples table

- Samples table
- Use sample identifier to select samples of the same group

The screenshot displays the Thermo Proteome Discoverer 2.5.0.400 software interface. The 'Samples' tab is active, showing a table with columns: Error, Samp, File, Sample Identifier, Sample Type, Quan Channel, treatment, Biological replic, and Temperature [C]. The 'sample1' identifier in the 'Sample Identifier' column is highlighted with an orange box. A context menu is open over this cell, listing options: Copy With Headers (Ctrl+C), Copy, Clear Selection, Cell Selection Mode, Enable Row Grouping, Set Sample Type to, Set treatment to, Set Biological replicate to, Set Temperature to, and Set as Input File. The 'Set Biological replicate to' and 'Set Temperature to' options have sub-menus open, showing values 'n/a' and '1' respectively.

Error	Samp	File	Sample Identifier	Sample Type	Quan Channel	treatment	Biological replic	Temperature [C]
			sample1					
	S4	F5	DMSO_sample1_TR1_Fraction - [126]	Sample	126	n/a	n/a	n/a
	S5	F6	DMSO_sample1_TR2_Fraction - [126]				n/a	n/a
	S6	F7	DMSO_sample1_TR3_Fraction - [126]				n/a	n/a
	S10	F11	Drug_sample1_TR1_Fraction - [126]				n/a	n/a
	S11	F12	Drug_sample1_TR2_Fraction - [126]				n/a	n/a
	S12	F13	Drug_sample1_TR3_Fraction - [126]				n/a	n/a
	S16	F5	DMSO_sample1_TR1_Fraction - [127N]				n/a	n/a
	S17	F5	DMSO_sample1_TR1_Fraction - [127C]				n/a	n/a
	S18	F5	DMSO_sample1_TR1_Fraction - [128N]				n/a	n/a
	S19	F5	DMSO_sample1_TR1_Fraction - [128C]				n/a	n/a
	S20	F5	DMSO_sample1_TR1_Fraction - [129N]				n/a	n/a
	S21	F5	DMSO_sample1_TR1_Fraction - [129C]				n/a	n/a
	S22	F5	DMSO_sample1_TR1_Fraction - [130N]	Sample	130N	n/a	n/a	n/a
	S23	F5	DMSO_sample1_TR1_Fraction - [130C]	Sample	130C	n/a	n/a	n/a

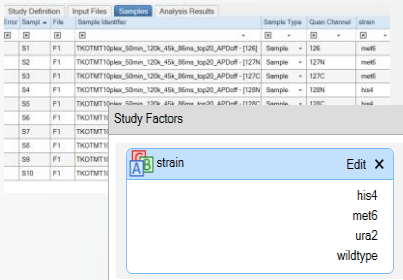
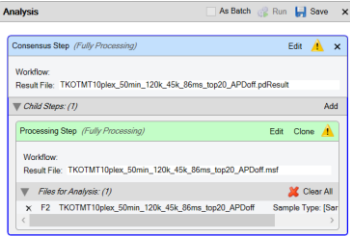
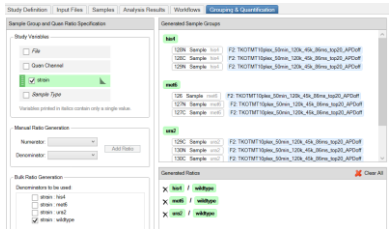
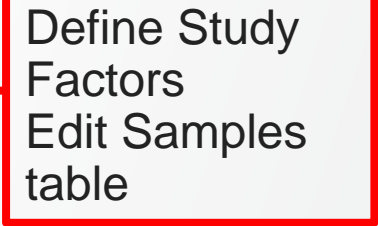
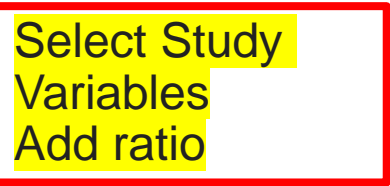
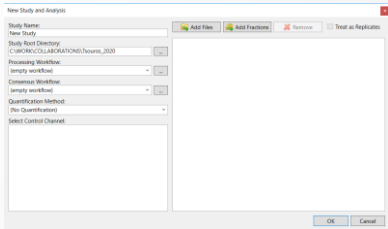
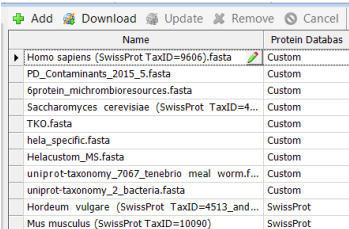
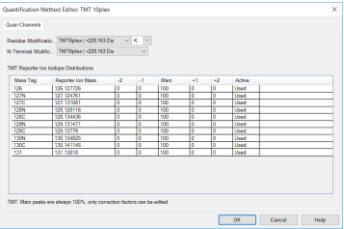
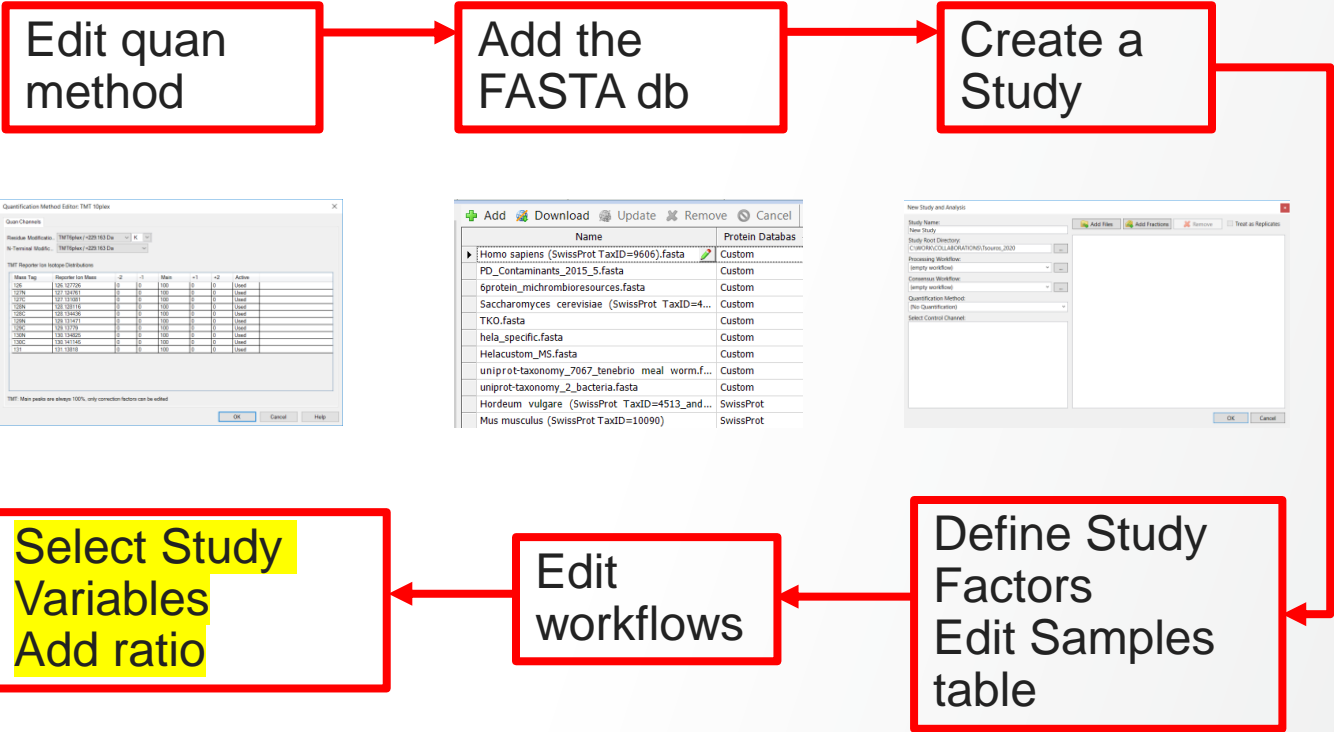
Edit Samples table

- Samples table
- Use sample identifier to select samples of the same group

The screenshot shows the Thermo Proteome Discoverer 2.5.0.400 software interface. The 'Samples' tab is active, displaying a table with columns: Error, Samp, File, Sample Identifier, Sample Type, Quan Channel, treatment, Biological replic, and Temperature [C]. The 'Sample Identifier' column is highlighted with an orange box, and a context menu is open over it. The menu options include: Copy With Headers (Ctrl+C), Copy, Clear Selection, Cell Selection Mode, Enable Row Grouping, Set Sample Type to, Set treatment to, Set Biological replicate to, Set Temperature to, and Set as Input File. The 'Set Temperature to' option is expanded, showing a list of temperatures: n/a, 37 Celsius, 41 Celsius, 44 Celsius, 47 Celsius, and 50 Celsius.

Error	Samp	File	Sample Identifier	Sample Type	Quan Channel	treatment	Biological replic	Temperature [C]
			126					
	S4	F5	DMSO_sample1_TR1_Fraction - [126]	Sample	126	n/a	1	n/a
	S5	F6	DMSO_sample1_TR2_Fraction - [126]			n/a	1	n/a
	S6	F7	DMSO_sample1_TR3_Fraction - [126]			n/a	1	n/a
	S7	F8	DMSO_sample2_TR1_Fraction - [126]			n/a	2	n/a
	S8	F9	DMSO_sample2_TR2_Fraction - [126]			n/a	2	n/a
	S9	F10	DMSO_sample2_TR3_Fraction - [126]			n/a	2	n/a
	S10	F11	Drug_sample1_TR1_Fraction - [126]			n/a	1	n/a
	S11	F12	Drug_sample1_TR2_Fraction - [126]			n/a	1	n/a
	S12	F13	Drug_sample1_TR3_Fraction - [126]			n/a	1	n/a
	S13	F14	Drug_sample2_TR1_Fraction - [126]			n/a		n/a
	S14	F15	Drug_sample2_TR2_Fraction - [126]					n/a
	S15	F16	Drug_sample2_TR3_Fraction - [126]	Sample	126	41 Celsius		n/a

General Quan Experiment Flowchart



Unset Variables

- PD will not allow you to proceed as some samples have an unassigned variable

File View Administration Tools Window Help

Start Page x Study: Multiplexed TMT11plex 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
- ☐ Biological replicate
- ☒ Temperature !
- ☒ treatment
- ☐ Sample Type

Manual Ratio Generation

Numerator:

Denominator:

Add Ratio

Generated Sample Groups

! There are samples with unset study variables selected for grouping.

37 Celsius DMSO


126	Sample	DMSO	1	37 Celsius	F5: DMSO_sample1_TR1_Fraction
126	Sample	DMSO	1	37 Celsius	F6: DMSO_sample1_TR2_Fraction
126	Sample	DMSO	1	37 Celsius	F7: DMSO_sample1_TR3_Fraction
126	Sample	DMSO	2	37 Celsius	F8: DMSO_sample2_TR1_Fraction
126	Sample	DMSO	2	37 Celsius	F9: DMSO_sample2_TR2_Fraction
126	Sample	DMSO	2	37 Celsius	F10: DMSO_sample2_TR3_Fraction

37 Celsius Drug

126	Sample	Drug	1	37 Celsius	F11: Drug_sample1_TR1_Fraction
126	Sample	Drug	1	37 Celsius	F12: Drug_sample1_TR2_Fraction
126	Sample	Drug	1	37 Celsius	F13: Drug_sample1_TR3_Fraction
126	Sample	Drug	2	37 Celsius	F14: Drug_sample2_TR1_Fraction
126	Sample	Drug	2	37 Celsius	F15: Drug_sample2_TR2_Fraction
126	Sample	Drug	2	37 Celsius	F16: Drug_sample2_TR3_Fraction

Study factors

- Treatment (DMSO; drug; pool)
 - *categorical* study factor
- Biological replicate (1; 2)
 - *Categorical* study factor
- Temperature (37-67; *value 100 is for “pool” which won’t figure in any protein ratio calculation*)
 - *Numerical* study factor
- The order/hierarchy of study factors can be changed later in Grouping&Quantification table


 treatment

Edit X

DMSO

Drug


Pool

 Biological replicate

Edit X

1

2

 Temperature

Edit X

37 Celsius

41 Celsius

44 Celsius

47 Celsius

50 Celsius

53 Celsius

56 Celsius

59 Celsius

63 Celsius

67 Celsius

100 Celsius

Define Pool sample

- Set channel 131C as “treatment = pool”; “temperature = 100”

Start Page x Study: Multiplexed TMT11plex * 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

Error	Samp	File	Sample Identifier	Sample Type	Quan Channel	treatment	Biological replic	Temperature [C]
			131C					
	S25	F5	DMSO_sample1_TR1_Fraction - [131C]	Control	131C	Pool	1	100
	S35	F6	DMSO_sample1_TR2_Fraction - [131C]	Control	131C	Pool	1	100
	S45	F7	DMSO_sample1_TR3_Fraction - [131C]	Control	131C	Pool	1	100
	S55	F8	DMSO_sample2_TR1_Fraction - [131C]	Control	131C	Pool	2	100
	S65	F9	DMSO_sample2_TR2_Fraction - [131C]	Control	131C	Pool	2	100
	S75	F10	DMSO_sample2_TR3_Fraction - [131C]	Control	131C	Pool	2	100
	S85	F11	Drug_sample1_TR1_Fraction - [131C]	Control	131C	Pool	1	100
	S95	F12	Drug_sample1_TR2_Fraction - [131C]	Control	131C	Pool	1	100
	S105	F13	Drug_sample1_TR3_Fraction - [131C]	Control	131C	Pool	1	100
	S115	F14	Drug_sample2_TR1_Fraction - [131C]	Control	131C	Pool	2	100
	S125	F15	Drug_sample2_TR2_Fraction - [131C]	Control	131C	Pool	2	100
	S135	F16	Drug_sample2_TR3_Fraction - [131C]	Control	131C	Pool	2	100

Define Pool sample

- Set channel 131C as “sample type = control”
 - These samples will be used for normalization across multiple sets

Error	Samp	File	Sample Identifier	Sample Type	Quan Channel	treatment	Biological replic	Temperature [C]
			131C					
	S25	F5	DMSO_sample1_TR1_Fraction - [131C]	Control	131C	Pool	1	100
	S35	F6	DMSO_sample1_TR2_Fraction - [131C]	Control	131C	Pool	1	100
	S45	F7	DMSO_sample1_TR3_Fraction - [131C]	Control	131C	Pool	1	100
	S55	F8	DMSO_sample2_TR1_Fraction - [131C]	Control	131C	Pool	2	100
	S65	F9	DMSO_sample2_TR2_Fraction - [131C]	Control	131C	Pool	2	100
	S75	F10	DMSO_sample2_TR3_Fraction - [131C]	Control	131C	Pool	2	100
	S85	F11	Drug_sample1_TR1_Fraction - [131C]	Control	131C	Pool	1	100
	S95	F12	Drug_sample1_TR2_Fraction - [131C]	Control	131C	Pool	1	100
	S105	F13	Drug_sample1_TR3_Fraction - [131C]	Control	131C	Pool	1	100
	S115	F14	Drug_sample2_TR1_Fraction - [131C]	Control	131C	Pool	2	100
	S125	F15	Drug_sample2_TR2_Fraction - [131C]	Control	131C	Pool	2	100
	S135	F16	Drug_sample2_TR3_Fraction - [131C]	Control	131C	Pool	2	100

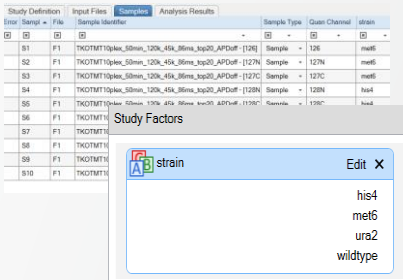
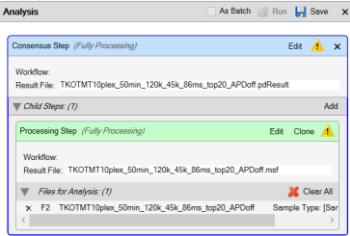
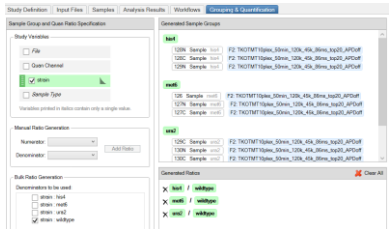
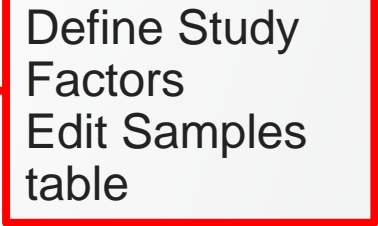
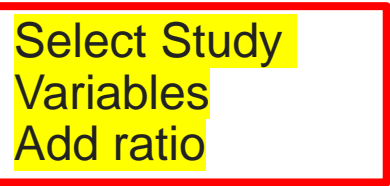
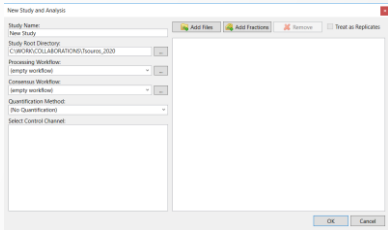
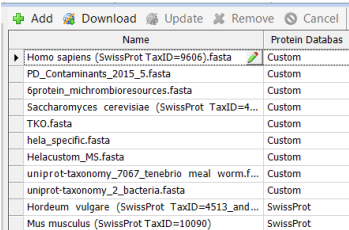
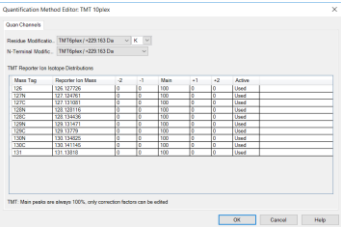
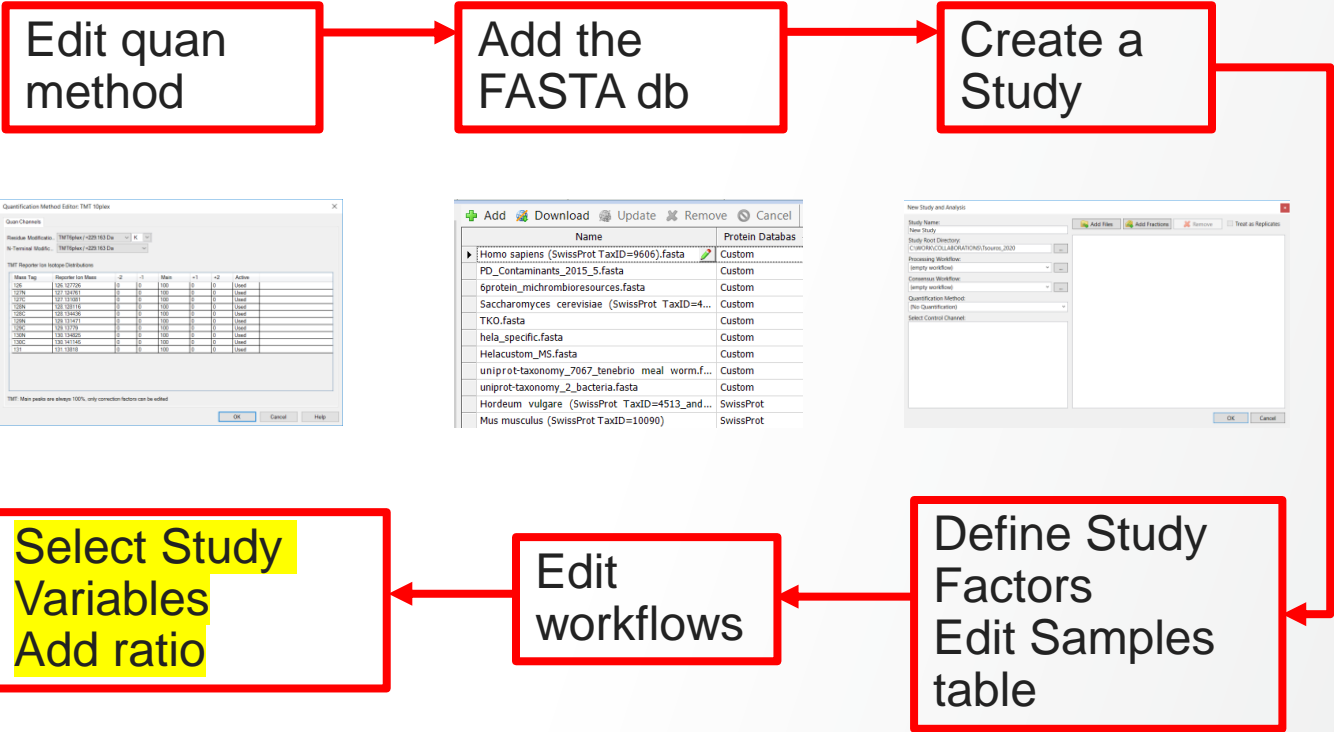
Reporter Ions Quantifier node

Define scaling “on controls average”

The screenshot displays the ThermoFisher Reporter Ions Quantifier node configuration window. The interface is divided into several sections:

- Study Definition:** Shows the study name "Study: Multiplexed TMT11plex".
- Parameters of 'Reporter Ions Quantifier':** A list of parameters organized into five sections:
 - 1. General Quantification Settings:** Includes "Peptides to Use" (Unique + Razor), "Consider Protein Groups for" (True), "Use Shared Quan Results" (True), and "Reject Quan Results with Mi" (False).
 - 2. Reporter Quantification:** Includes "Reporter Abundance Base" (Automatic), "Apply Quan Value Correctio" (False), "Co-Isolation Threshold" (50), "Average Reporter S/N Thre" (10), "SPS Mass Matches [%] Thr" (65), and "Minimum Channel Occupan" (0).
 - 3. Normalization and Scaling:** Includes "Normalization Mode" (Total Peptide Amount), "Proteins For Normalization" (On Controls Average), and "Scaling Mode" (On Controls Average).
 - 4. Exclude Peptides from Protein Quantification:** Includes "For Normalization" (Use All Peptides), "For Protein Roll-Up" (Use All Peptides), "For Pairwise Ratios" (Exclude Modified), and "1. Considered Peptide Mod" (None).
 - 5. Quan Rollup and Hypothesis Testing:** Includes "Protein Ratio Calculation" (Protein Abundance Based), "Maximum Allowed Fold Cha" (100), "Imputation Mode" (None), and "Hypothesis Test" (ANOVA (Individual Proteins)).
- Workflow Tree:** A diagram showing the workflow steps: "MSF Files" (0) → "PSM Grouper" (1) → "Peptide Validator" (2) → "Peptide and Protein Filter" (3). The "Reporter Ions Quantifier" node (10) is highlighted with a purple border and a green status indicator.

General Quan Experiment Flowchart



Q: Is biological variance obscuring the effect of treatment?

Technical replicates presumed to have the lowest variability

- Set up sample groups containing just a single sample
- Define some ratios (*same sample* as denominator throughout)
- Run the Analysis
- Check the **PCA plot**

- Do samples cluster by channel (=temperature)?

O.K.

- Or do samples cluster by file (=biological replicate)?

?

The screenshot displays the 'Grouping & Quantification' tab of a software interface. It is divided into three main sections:

- Sample Group and Quan Ratio Specification:**
 - Study Variables:** A list of variables with checkboxes. 'Quan Channel' and 'File' are checked and highlighted in green. 'treatment', 'Temperature', 'Biological replicate', and 'Sample Type' are unchecked.
 - Manual Ratio Generation:** Fields for 'Numerator' and 'Denominator' with a dropdown menu and an 'Add Ratio' button.
 - Bulk Ratio Generation:** A section titled 'Denominators to be used:' containing a tree view. Under 'Quan Channel : 126', 'File : F5' is selected with a checkmark. Other files (F6-F10) are listed with unchecked checkboxes. An 'Add Ratios' button is at the bottom.
- Generated Sample Groups:** A list of sample groups. Each entry shows a sample ID (126), a sample name (Sample), a treatment (DMSO), a temperature (37 Celsius), and a fraction (F5-F8). The fractions are highlighted in green.
- Generated Ratios:** A list of ratios. Each entry shows a sample ID (126), a sample name (Sample), a treatment (DMSO), a temperature (37 Celsius), and a fraction (F5-F15). The fractions are highlighted in green. A 'Clear All' button is at the top right.

Grouping & Quantification table

Effect of temperature at various treatment conditions

- Primary variable
“Treatment”
- Secondary variable
“Temperature”
- Change order of variables if needed

(median 41C DMSO) / (median 37C DMSO)
(median 44C DMSO) / (median 37C DMSO)
(median 47C DMSO) / (median 37C DMSO)
...

Study Definition | Input Files | Samples | Analysis Results | Workflows | **Grouping & Quantification**

Sample Group and Quan Ratio Specification

Study Variables

- ☐ File
- ☐ Quan Channel
- ☐ Biological replicate
- ☒ treatment
- ☒ Temperature
- ☐ Sample Type

Manual Ratio Generation

Numerator: Add Ratio

Denominator:

Bulk Ratio Generation

Denominators to be used:

- ☒ treatment : DMSO
- ☐ Temperature : 37 Celsius
- ☐ Temperature : 41 Celsius
- ☐ Temperature : 44 Celsius
- ☐ Temperature : 47 Celsius
- ☐ Temperature : 50 Celsius
- ☐ Temperature : 53 Celsius

Add Ratios

Generated Sample Groups

1 of 21 sample groups not used in any ratio definition.

DMSO 37 Celsius

126	Sample	DMSO	1	37 Celsius	F5: DMSO_sample1_TR1_Fraction
126	Sample	DMSO	1	37 Celsius	F6: DMSO_sample1_TR2_Fraction
126	Sample	DMSO	1	37 Celsius	F7: DMSO_sample1_TR3_Fraction
126	Sample	DMSO	2	37 Celsius	F8: DMSO_sample2_TR1_Fraction
126	Sample	DMSO	2	37 Celsius	F9: DMSO_sample2_TR2_Fraction
126	Sample	DMSO	2	37 Celsius	F10: DMSO_sample2_TR3_Fraction

DMSO 41 Celsius

127N	Sample	DMSO	1	41 Celsius	F5: DMSO_sample1_TR1_Fraction
127N	Sample	DMSO	1	41 Celsius	F6: DMSO_sample1_TR2_Fraction
127N	Sample	DMSO	1	41 Celsius	F7: DMSO_sample1_TR3_Fraction
127N	Sample	DMSO	2	41 Celsius	F8: DMSO_sample2_TR1_Fraction
127N	Sample	DMSO	2	41 Celsius	F9: DMSO_sample2_TR2_Fraction
127N	Sample	DMSO	2	41 Celsius	F10: DMSO_sample2_TR3_Fraction

Generated Ratios Clear All

X	DMSO	53 Celsius	/	DMSO	37 Celsius
X	DMSO	56 Celsius	/	DMSO	37 Celsius
X	DMSO	59 Celsius	/	DMSO	37 Celsius
X	DMSO	63 Celsius	/	DMSO	37 Celsius
X	DMSO	67 Celsius	/	DMSO	37 Celsius
X	Drug	41 Celsius	/	Drug	37 Celsius
X	Drug	44 Celsius	/	Drug	37 Celsius
X	Drug	47 Celsius	/	Drug	37 Celsius
X	Drug	50 Celsius	/	Drug	37 Celsius

Grouping&Quantification table

- One sample group not used in any ratios (the “pool”)

Study DefinitionInput FilesSamplesAnalysis ResultsWorkflowsGrouping & Quantification

Sample Group and Quan Ratio Specification

Study Variables

☐ File

☐ Quan Channel

☐ Biological replicate

☒ treatment

☒ Temperature

☐ Sample Type

Manual Ratio Generation

Numerator:

Denominator:

Add Ratio

Bulk Ratio Generation

Denominators to be used:

☐ Temperature : 53 Celsius

☐ Temperature : 56 Celsius

☐ Temperature : 59 Celsius

☐ Temperature : 63 Celsius

☐ Temperature : 67 Celsius

☒ treatment : Pool

☒ Temperature : 100 Celsius

Add Ratios

Generated Sample Groups

1 of 21 sample groups not used in any ratio definition.

DMSO37 Celsius

126 Sample DMSO 1 37 Celsius

126 Sample DMSO 1 37 Celsius

126 Sample DMSO 1 37 Celsius

126 Sample DMSO 2 37 Celsius

126 Sample DMSO 2 37 Celsius

126 Sample DMSO 2 37 Celsius

F5: DMSO_sample1_TR1_Fraction

F6: DMSO_sample1_TR2_Fraction

F7: DMSO_sample1_TR3_Fraction

F8: DMSO_sample2_TR1_Fraction

F9: DMSO_sample2_TR2_Fraction

F10: DMSO_sample2_TR3_Fraction

DMSO41 Celsius

127N Sample DMSO 1 41 Celsius

127N Sample DMSO 1 41 Celsius

127N Sample DMSO 1 41 Celsius

F5: DMSO_sample1_TR1_Fraction

F6: DMSO_sample1_TR2_Fraction

F7: DMSO_sample1_TR3_Fraction

Generated Ratios

Clear All

X DMSO 41 Celsius / DMSO 37 Celsius

X DMSO 44 Celsius / DMSO 37 Celsius

X DMSO 47 Celsius / DMSO 37 Celsius

X DMSO 50 Celsius / DMSO 37 Celsius

X DMSO 53 Celsius / DMSO 37 Celsius

X DMSO 56 Celsius / DMSO 37 Celsius

X DMSO 59 Celsius / DMSO 37 Celsius

X DMSO 63 Celsius / DMSO 37 Celsius

X DMSO 67 Celsius / DMSO 37 Celsius

X Drug 41 Celsius / Drug 37 Celsius

X Drug 44 Celsius / Drug 37 Celsius

Grouping & Quantification table

Drug treatment effect at various temperatures

- Primary variable
“Temperature”
- Secondary variable
“Treatment”

- Change order of variables if needed

(median 37C Drug) / (median 37C DMSO)

(median 41C Drug) / (median 41C DMSO)

(median 44C Drug) / (median 44C DMSO)

...

The screenshot shows the 'Grouping & Quantification' tab in the 'Study: Multiplexed TMT11plex' software. The 'Sample Group and Quan Ratio Specification' section on the left has 'Study Variables' where 'Temperature' and 'treatment' are checked. A green arrow points from the 'Change order of variables if needed' bullet point to the 'Temperature' checkbox. Below this is 'Manual Ratio Generation' and 'Bulk Ratio Generation'. The 'Generated Sample Groups' section on the right shows a list of sample groups, including '37 Celsius DMSO' and '37 Celsius Drug'. The 'Generated Ratios' section at the bottom shows a list of ratios, including '37 Celsius Drug / 37 Celsius DMSO'.

Grouping & Quantification table

Drug treatment effect at various temperatures

- Protein ratio calculation:
 - PD selects median value representing each sample group
 - In case of even number of values (6), it uses geometrical mean of the two middle values

Study: Multiplexed TMT11plex

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
- ☐ Biological replicate
- ☒ Temperature
- ☒ treatment
- ☐ Sample Type

Manual Ratio Generation

Numerator: Denominator:

Bulk Ratio Generation

Generated Sample Groups

1 of 21 sample groups not used in any ratio definition.

37 Celsius DMSO

- 126 Sample DMSO 1 37 Celsius F5: DMSO_sample1_TR1_Fraction
- 126 Sample DMSO 1 37 Celsius F6: DMSO_sample1_TR2_Fraction
- 126 Sample DMSO 1 37 Celsius F7: DMSO_sample1_TR3_Fraction
- 126 Sample DMSO 2 37 Celsius F8: DMSO_sample2_TR1_Fraction
- 126 Sample DMSO 2 37 Celsius F9: DMSO_sample2_TR2_Fraction
- 126 Sample DMSO 2 37 Celsius F10: DMSO_sample2_TR3_Fraction

37 Celsius Drug

- 126 Sample Drug 1 37 Celsius F11: Drug_sample1_TR1_Fraction
- 126 Sample Drug 1 37 Celsius F12: Drug_sample1_TR2_Fraction
- 126 Sample Drug 1 37 Celsius F13: Drug_sample1_TR3_Fraction

Generated Ratios

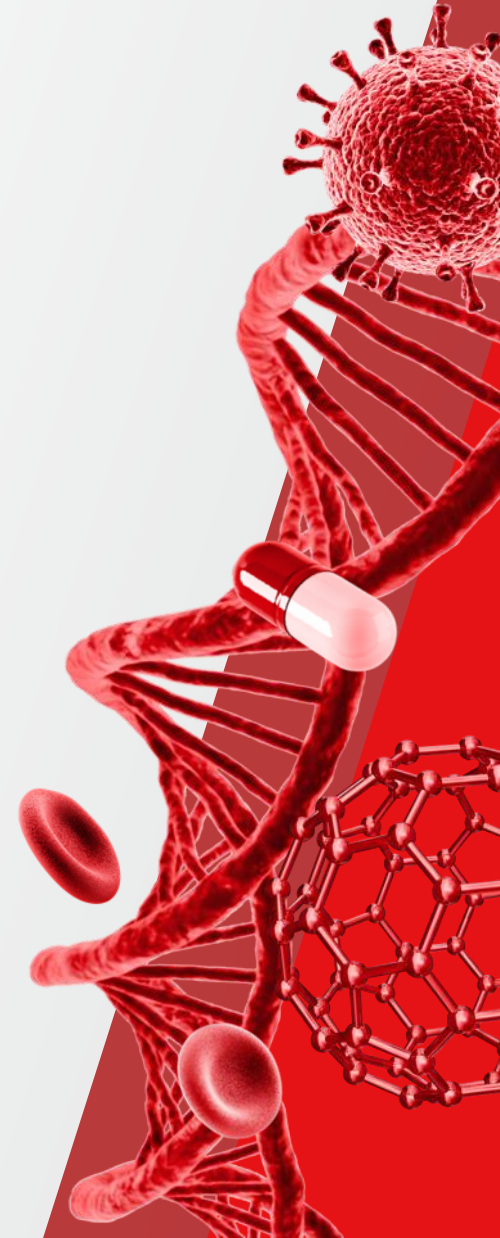
- 37 Celsius Drug / 37 Celsius DMSO
- 41 Celsius Drug / 41 Celsius DMSO

Median 37C DMSO

Median 37C Drug

Additional Info

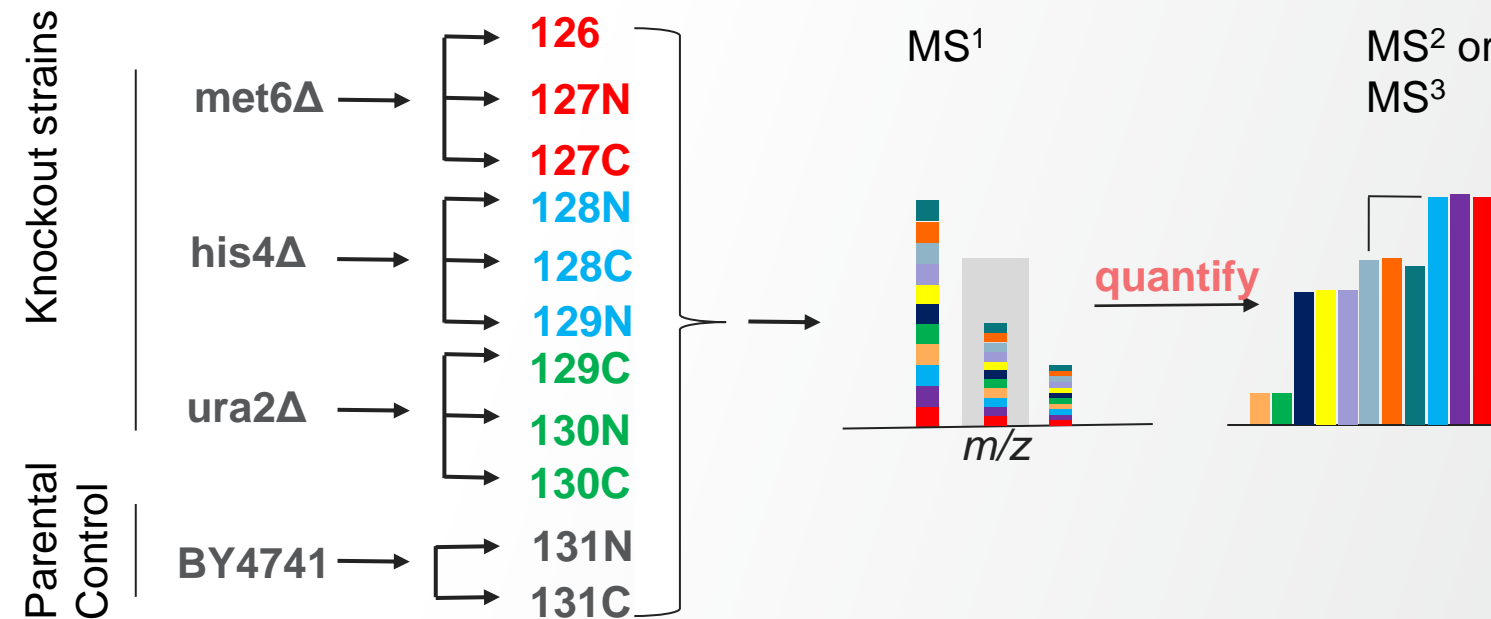
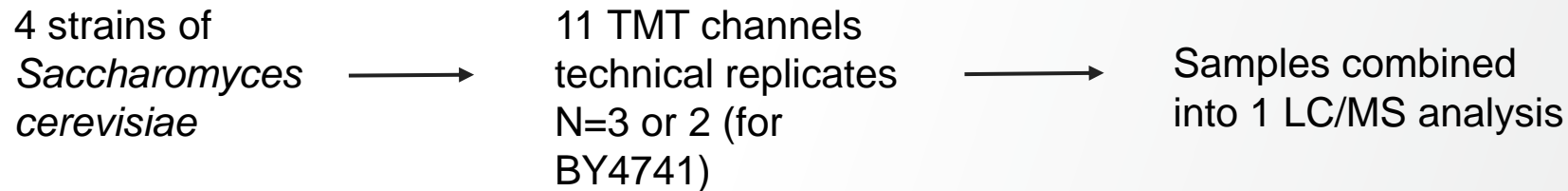
 The world leader in serving science



Standard Labeled Sample

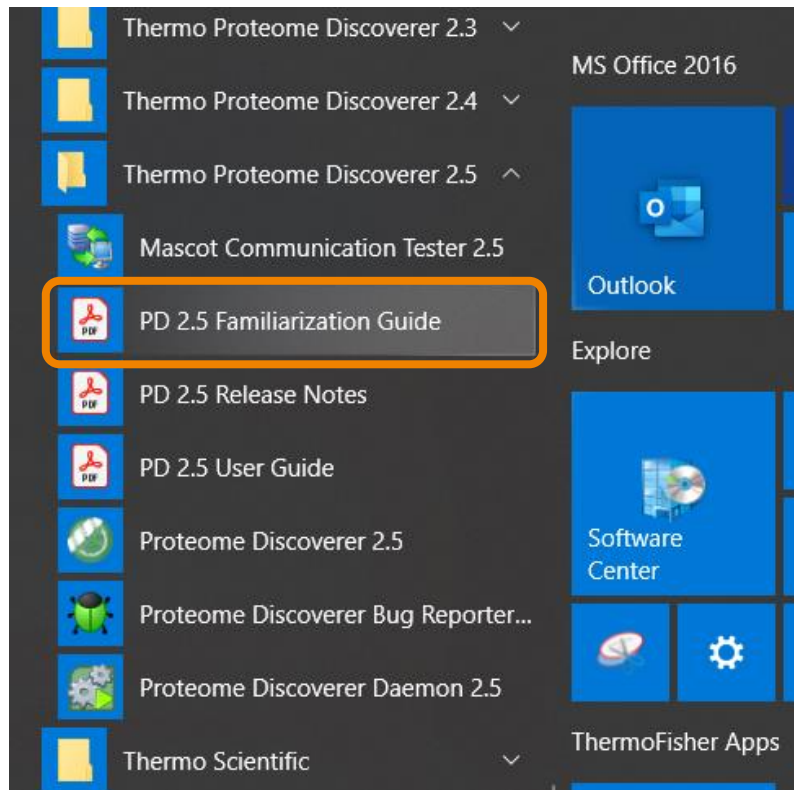
For instrument method development

Sample: Thermo Scientific™ Pierce™ TMT11plex Yeast Digest Standard



More TMT Quan Experiments

- Advanced level experiments in **Familiarization Guide**
- Windows → All Programs → Thermo PD 2.5



- Analysis of Yeast Gene Knockouts by Multiplexed Reporter Ion Quantification
- Reporter Ion Quantification with Phosphorylation
- Label-Free Quantitation with Spiked-In Proteins
- SILAC 3plex of E. coli Proteins Mixed with Known Ratios
- Analysis of Yeast Gene Knockouts by Multiplexed Reporter Ion Quantification with Two Replicates

- pd.support@thermofisher.com
- Provide the following information
 - Problem description
 - Screenshots illustrating the problem, highlighting an example
 - Bug report

