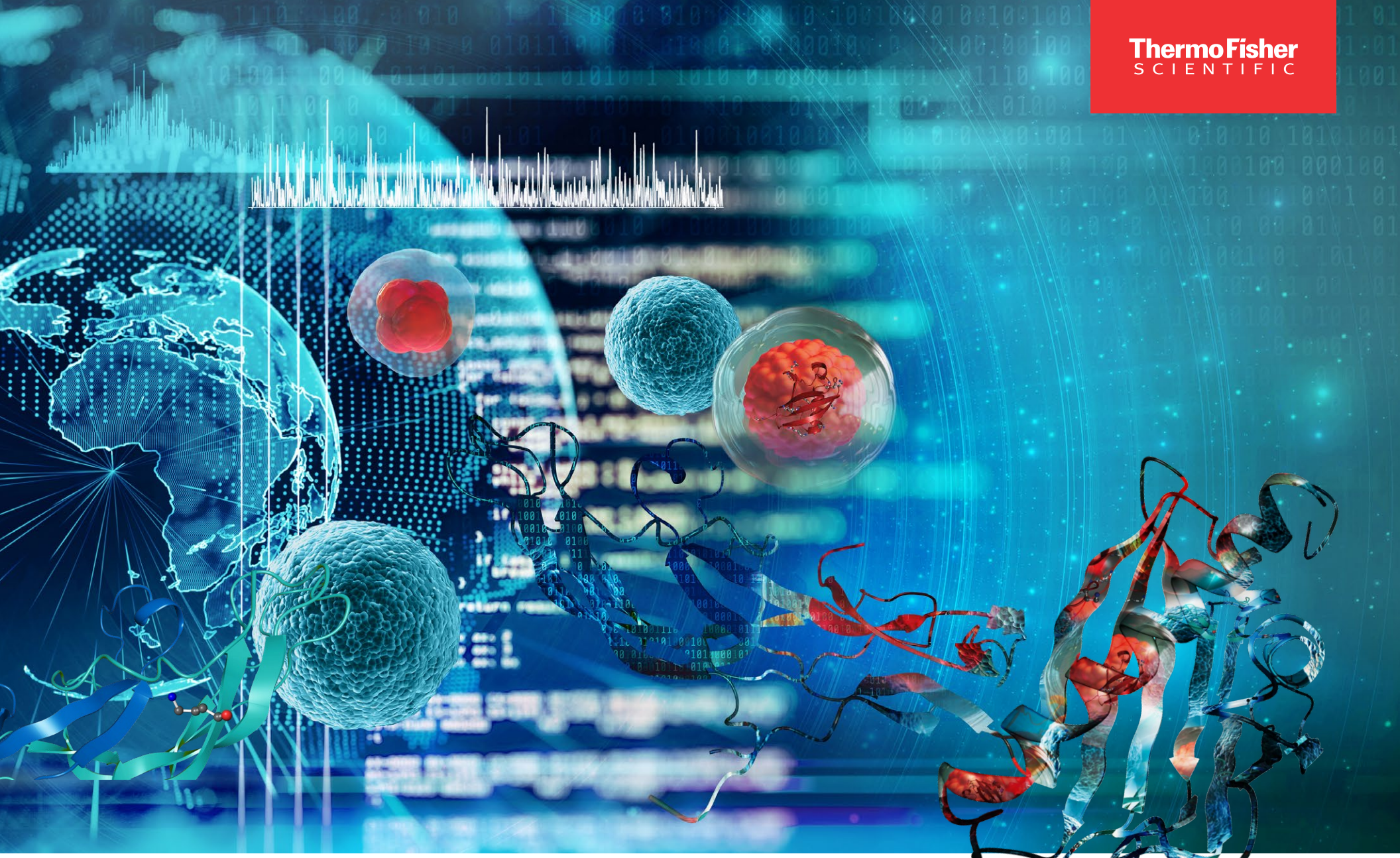


The intelligent protein informatics platform

Thermo Scientific Proteome Discoverer software

thermo scientific

Transform proteomics mass spectrometry data into insights

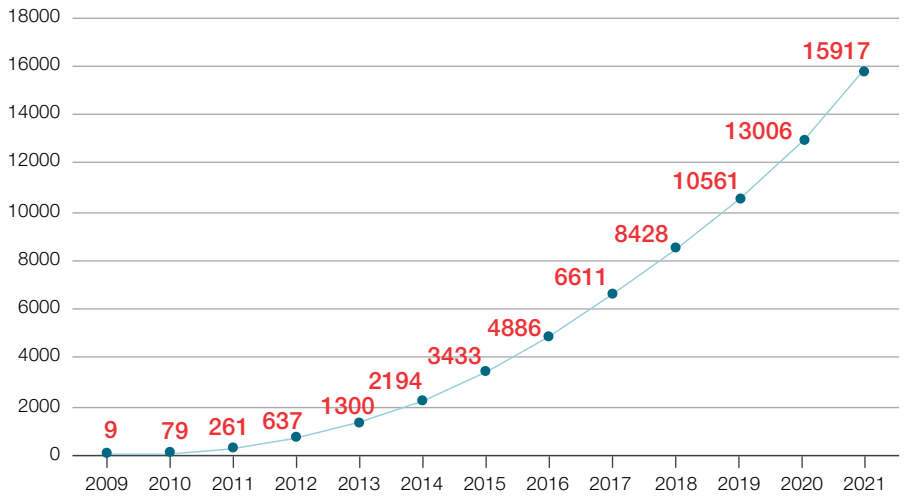
Thermo Scientific™ Proteome Discoverer™ software enables comprehensive proteomics data processing workflows empowered by artificial intelligence.

- **Powerful and flexible framework:** Optimized analysis for data generated by the sophisticated acquisition methods of Thermo Scientific™ Orbitrap™ mass spectrometers. Easily analyze MS^2 and MS^n data or multiple fragmentation modes in the same analysis. Dig deeper into data using artificial intelligence with INFERYS™ Rescoring and the CHIMERYS™ intelligent search algorithm. Mix and match multiple search engines in series or in parallel to maximize insights with built-for-purpose workflows. Use third party-developed nodes within an extensible framework.
- **Study management:** Integrated management of raw data files, search results, and quantification methods. Map study factors to quantification channels to set up replicates and statistical analyses to easily match results to experimental variables.
- **Biological annotation:** Match identified proteins with gene ontology, gene symbol, and pathway terms. Further enhance biological interpretation using enrichment analysis to help transform data into biological insights.
- **Powerful and flexible result visualization and interpretation:** Results are presented hierarchically, with dynamic links between proteins, peptides, peptide spectrum matches (PSMs), and quantification tables. Easily filter data to view only those results of primary interest. Use interactive data plots that directly link to their source tables to dynamically interact with results. Find which proteins or peptides are up- or down-regulated through the volcano plot, which samples are related via a PCA plot, or visualize abundance patterns across files using the heat map.

Since releasing in 2009, Proteome Discoverer software has been featured in thousands of publications across multiple proteomics applications including:

- Bottom-up peptide and protein Identification
- Label-Free Quantification (LFQ)
- Isobaric Mass Tagging (including Thermo Scientific™ TMT™ and Thermo Scientific™ TMTpro™) Peptide Quantification
- Single-cell proteomics
- Immunopeptidomics
- Metaproteomics
- Top-down Proteoform Identification and Quantification with Thermo Scientific™ ProSightPD™ software
- Crosslinked peptide identification using XlinkX software
- Post-Translational Modification (PTM) Analysis

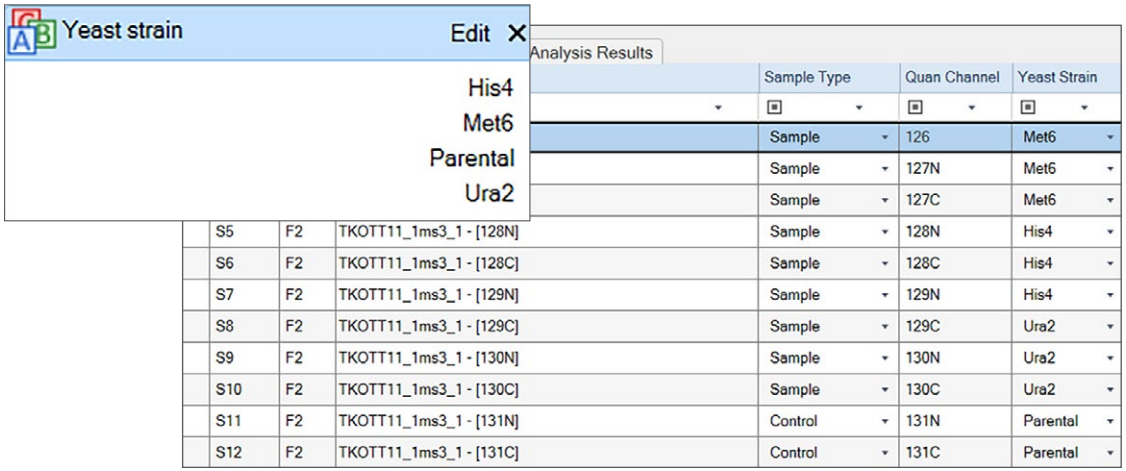
Proteome Discoverer software in peer reviewed articles



Transform proteomics mass spectrometry data into insights

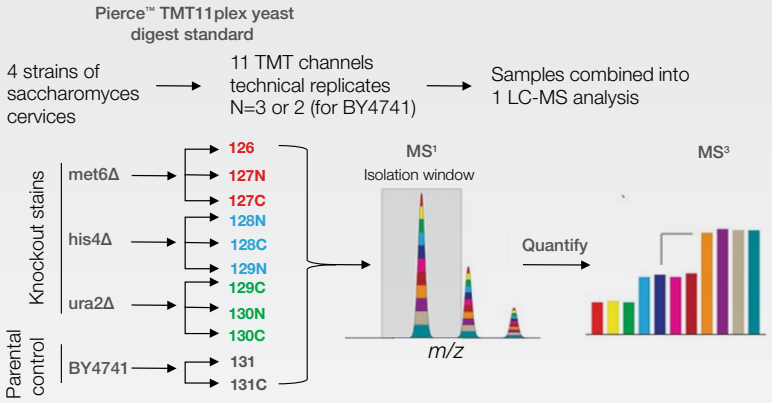
Integrated study management

Performing complex proteomics experiments requires the ability to map data files to experimental study conditions to organize and make sense of results. Proteome Discoverer software uses a built-in study design feature to assign qualitative and quantitative variables. The study allows selection of data files, association of the data files with a quantification method, the creation of experimental variables as study factors, and the mapping of study factors to data files and quantification channels. Each data file or collection of fractionated data files can be mapped to one or more study factors and then used to perform comparative and statistical analyses.

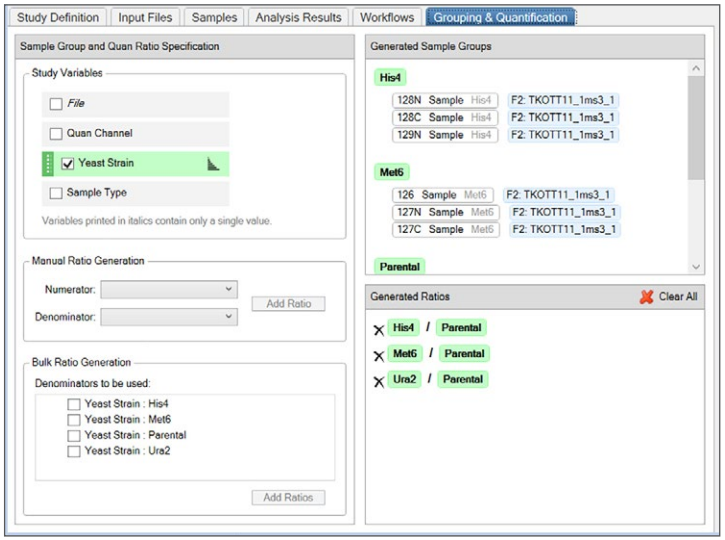


			Sample Type	Quan Channel	Yeast Strain
			Sample	126	Met6
			Sample	127N	Met6
			Sample	127C	Met6
S5	F2	TKOTT11_1ms3_1 - [128N]	Sample	128N	His4
S6	F2	TKOTT11_1ms3_1 - [128C]	Sample	128C	His4
S7	F2	TKOTT11_1ms3_1 - [129N]	Sample	129N	His4
S8	F2	TKOTT11_1ms3_1 - [129C]	Sample	129C	Ura2
S9	F2	TKOTT11_1ms3_1 - [130N]	Sample	130N	Ura2
S10	F2	TKOTT11_1ms3_1 - [130C]	Sample	130C	Ura2
S11	F2	TKOTT11_1ms3_1 - [131N]	Control	131N	Parental
S12	F2	TKOTT11_1ms3_1 - [131C]	Control	131C	Parental

Proteome Discoverer software enables the creation and mapping of sample types, study factors, and quantification channels to data files for simplified and integrated study management.



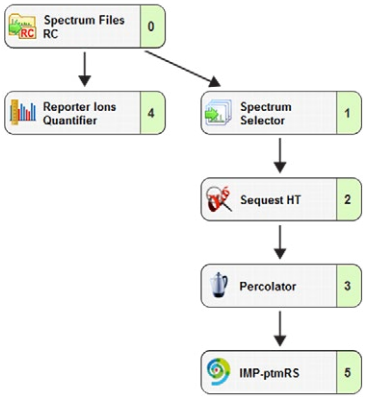
Schematic representation of Pierce TMT11plex Yeast Digest Standard and replicates as an example of study management with Proteome Discoverer software. The standard is composed of four Saccharomyces cerevisiae strains: three lines respectively lacking the non-essential proteins Met6, His4, or Ura2, and the parental strain BY4741 for reference.[1]



Generate ratios from study variables or sample types for comparative and statistical analyses.

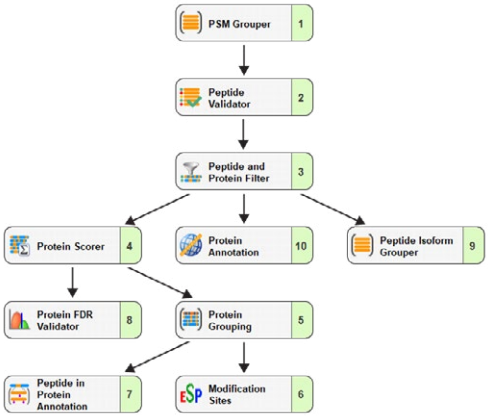
Pre-defined and customizable processing workflows

The Proteome Discoverer software workflow system is composed of a node-based architecture that can connect in linear or branched pathways to maximize the amount of information that can be gleaned from a proteomics dataset. The workflow system allows for new users to quickly and easily deploy pre-defined workflows for TMT, LFQ, PTM identification, and crosslinked peptide identification and offers flexibility for expert users to assemble advanced workflows to analyze the most complex proteomics data sets. Spectra are introduced into the top of the workflow and pass through nodes that perform filtering, peptide identification, quantification, false discovery rate (FDR) calculations, PTM site localization, bioinformatic analysis, and more.



The workflow system in Proteome Discoverer allows for customizable workflows to analyze complex proteomics data. The workflow system is composed of nodes that can connect in linear or branched pathways. While these workflows are highly customizable, pre-defined workflows for common workflows such as TMT, LFQ, PTM identification and cross-linked peptide ID are included.

The consensus workflow adds additional grouping and validation steps along with biological interpretation. The Protein Annotation node adds result tables to group proteins with pathway annotation and GO and Pfam annotations for built in bioinformatic analysis, such as enrichment. Workflows can also be customized using the Scripting Node to run executables to integrate results from custom scripts directly into Proteome Discoverer software results.



Spectra are streamed from the top of the workflow and pass through nodes that perform filtering, peptide ID, quantification, FDR calculation, PTM site localization, bioinformatic analysis, and more.

Powerful data processing workflows tailored to fit advanced instrument acquisition strategies

Proteomics experiments can require multiple fragmentation types to generate the necessary information for a given study. Proteome Discoverer software supports all Orbitrap mass spectrometer fragmentation methods for MS^2 or MS^n data including higher energy collisional dissociation (HCD), collision-induced dissociation (CID), electron transfer dissociation (ETD), electron transfer higher energy collisional dissociation (EThcD), ultraviolet photodissociation (UVPD), proton transfer charge reduction (PTCR) and combinations thereof. Uncover more peptides and proteins using the Thermo Scientific™ FAIMS Pro Duo interface for identification and quantitation. Customize workflows for multiple MS^n modes with alternating CID and ETD fragmentation using the Scan Event Filter and multiple search engines to improve annotation. Deploy the most confident TMT quantification strategies with synchronous precursor selection MS^3 (SPS MS^3) and Real-Time Search-Aided TMT SPS MS^3 with Sequest™ HT and INFERYS Rescoring, the CHIMERYS intelligent search algorithm, or Comet and use the Reporter Ions Quantifier node to quantify. Use MS^2 and MS^3 information to identify and quantify crosslinked peptides with the XlinkX node. Perform top-down proteomics with the best sequence coverage with a combination of UVPD, ETD, EThcD, HCD, CID, and PTCR with ProSightPD software. Identify the most PTMs with multi-stage activation and EThcD using Sequest HT and IMP-ptmRS.

Workflow	Pre-built Proteome Discoverer software workflows	Additional benefit
Real-time Search Aided TMT SPS MS^3	PWF_Tribrid_TMT_Quan_SPS_MS3_SequestHT_Percolator, PWF_Tribrid_TMTpro_Quan_SPS_MS3_SequestHT_Percolator	Can also use Comet search engine to more closely match the RTS results
LFQ	PWF_Hybrid_LFQ_INFERYS_Rescoring_SequestHT Percolator_2stage, PWF_Tribrid_Precursor_Quan_and_LFQ_IT HCD_SequestHT_Percolator, PWF_Hybrid_LFQ_CHIMERYS	Also supports FAIMS LFQ
PTMs	PWF_Hybrid_SequestHT_MSAmanda_Percolator_ptmRS.pdProcessingWF	Identify and localize PTMs and can be used with quan nodes
Immunopep-tidomics	PWF_Hybrid_INFERYS_Rescoring_SequestHT_Percolator	INFERYS Rescoring helps confirm peptides that have low Xcorr but high correlation to the prediction spectrum for the identified peptide
MS^2-MS^3 analysis of cleavable crosslinking	DSSO_MS2_MS3_TMT6quan	Proteome Discoverer software can quantify crosslinked peptide data using either TMT or LFQ
Top down analysis	PSPD 1 percent FDR LFQ for Med HI data, PSPD Truncation Subsequense Search with FDR, PSPD Single Proteoform	Enables identification and quantification of large (>50 kDa) proteoforms. Also supports PTCR.

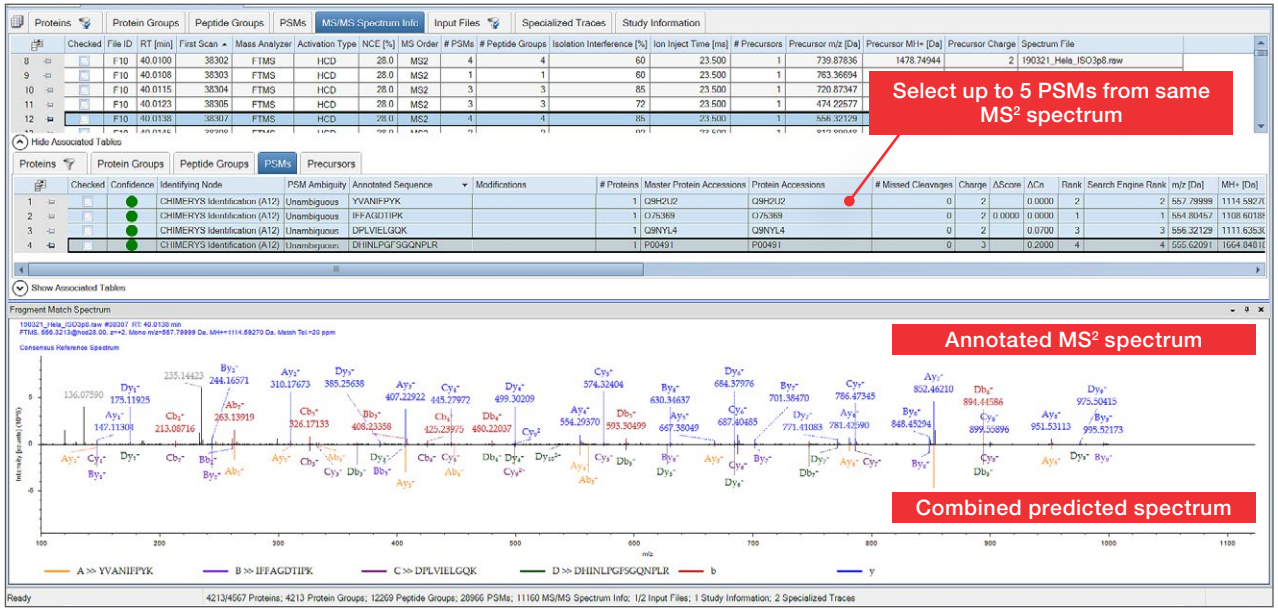
Utilize pre-built Proteome Discoverer software workflows to easily turn data into meaningful results.

Comprehensive search capabilities empowered by artificial intelligence

A deeper mining of proteomics data by reimagining the analysis of tandem mass spectra

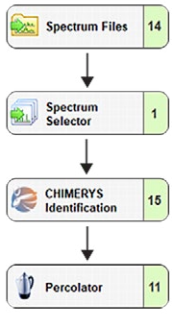
Deconvolution of chimeric spectra with artificial intelligence

The new CHIMERYS intelligent search algorithm utilizes artificial intelligence to rethink the analysis of tandem mass spectra from the ground up. CHIMERYS overcomes the challenges of mixed peptide MS² spectra by providing accurate peptide spectrum and retention time predictions through deep learning combined with a powerful deconvolution algorithm. This revolutionary approach leads to a deeper mining of data and substantially increases the number of PSMs found in data-dependent acquisition data. In comparison to previous strategies, CHIMERYS finds more PSMs per tandem mass spectrum and markedly improves the identification rate, with fewer spectra returning no PSMs and many spectra returning three or more PSMs. This results in a substantial improvement in the number of unique peptides and proteins identified and quantified in typical proteomics datasets compared to other approaches.

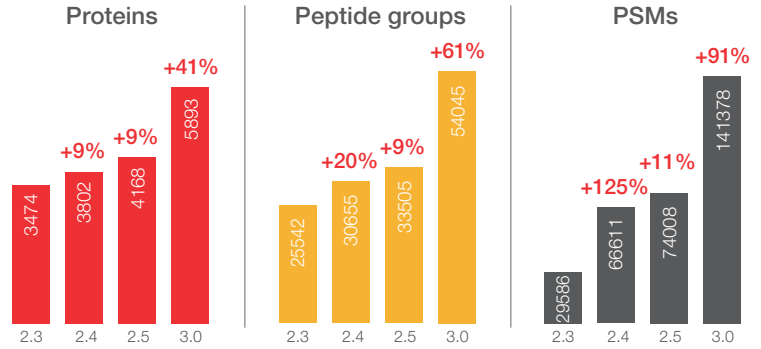


Deconvolution and visualization of chimeric spectra. The experimental and predicted spectrum can be visualized by selecting up to 5 PSMs at a time to display a mirror plot.

CHIMERYS intelligent search algorithm workflow



Continuous innovation to dig deeper into data

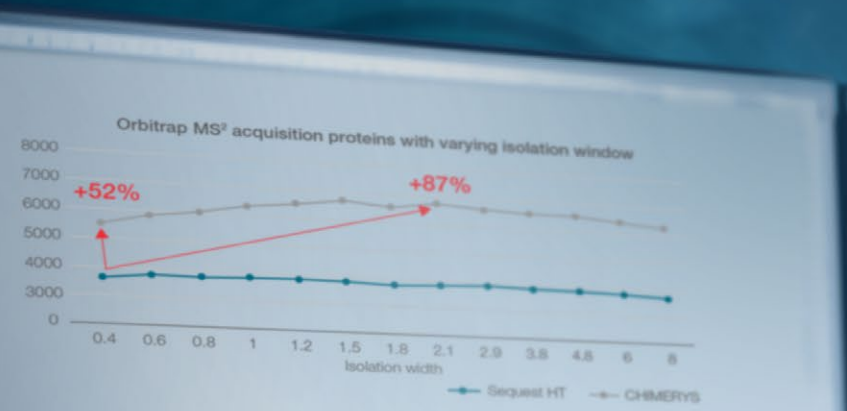


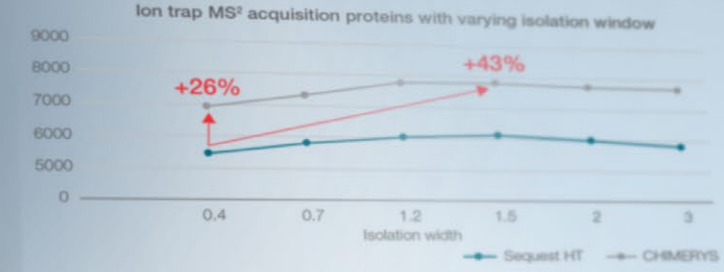
The CHIMERYS intelligent search algorithm workflow produces 70% more proteins, more than 2× more unique peptides, and almost 5× as many PSMs as Proteome Discoverer 2.3 software using the same data from a typical 90-minute run.

Not just software—wide window acquisition

Optimize instrument acquisition methods with Proteome Discoverer software with CHIMERYS

The data-dependent acquisition of proteomics samples traditionally utilizes narrow isolation windows to reduce the co-isolation of multiple peptides due to limitations in the ability of search engines to deconvolute chimeric spectra. In contrast to these traditional constraints, the CHIMERYS intelligent search algorithm utilizes artificial intelligence to rethink the analysis of tandem mass spectra from the ground up to enable the deconvolution of complex chimeric spectra and enables new strategies of data acquisition. Employing wider precursor isolation windows for MS^2 data acquisition paired with data processing using CHIMERYS allows for even larger increases in performance for both Orbitrap/Orbitrap and Orbitrap/Ion Trap acquisition methods. Acquiring MS^2 data with the Orbitrap provides the biggest improvements with isolation widths between 2–4 Th and Ion Trap data acquisition plateaus around 1.5 Th. The pairing of optimized data acquisition and processing strategies provides synergistic benefits by increasing instrument utilization through the simultaneous fragmentation of more peptides per tandem mass spectrum collected and the deconvolution of the resulting chimeric spectra to accurately interpret the resulting data complexity with CHIMERYS.



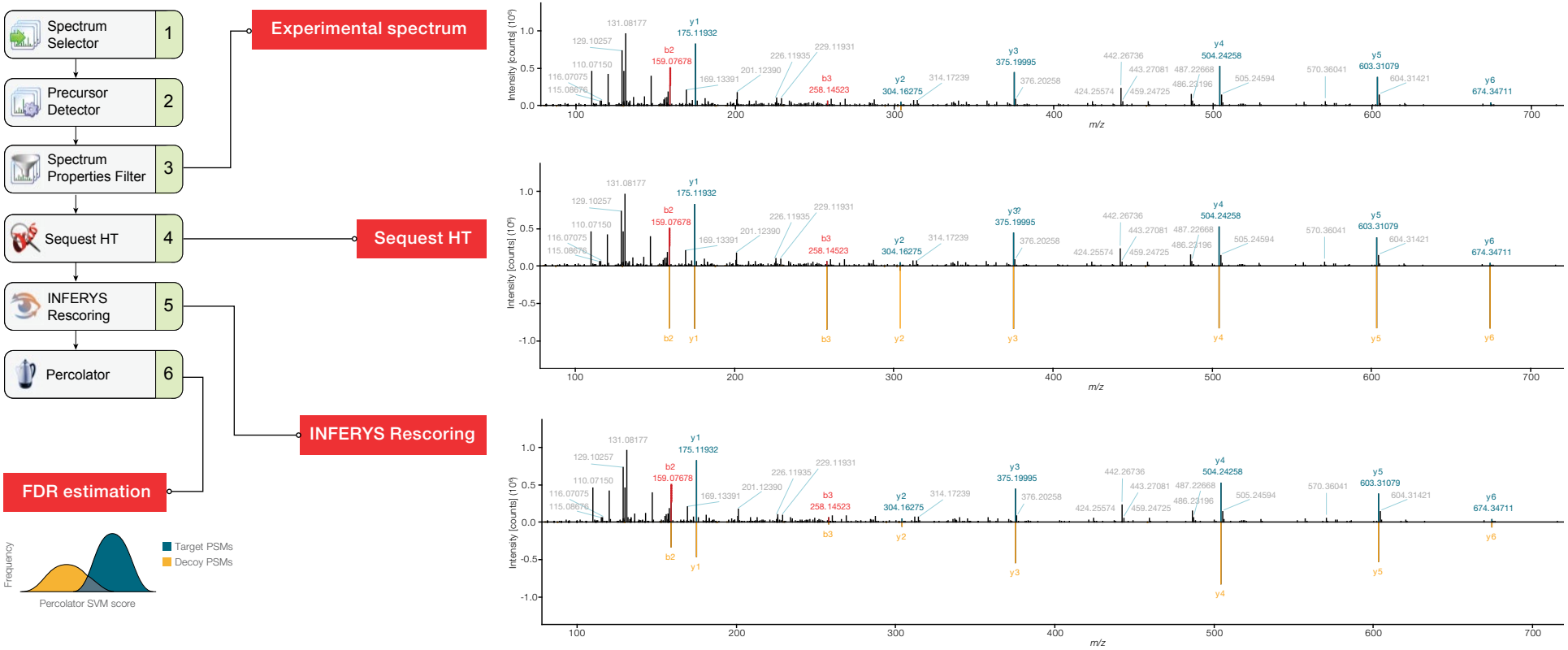


Improving confidence in search results using INFERYS Rescoring

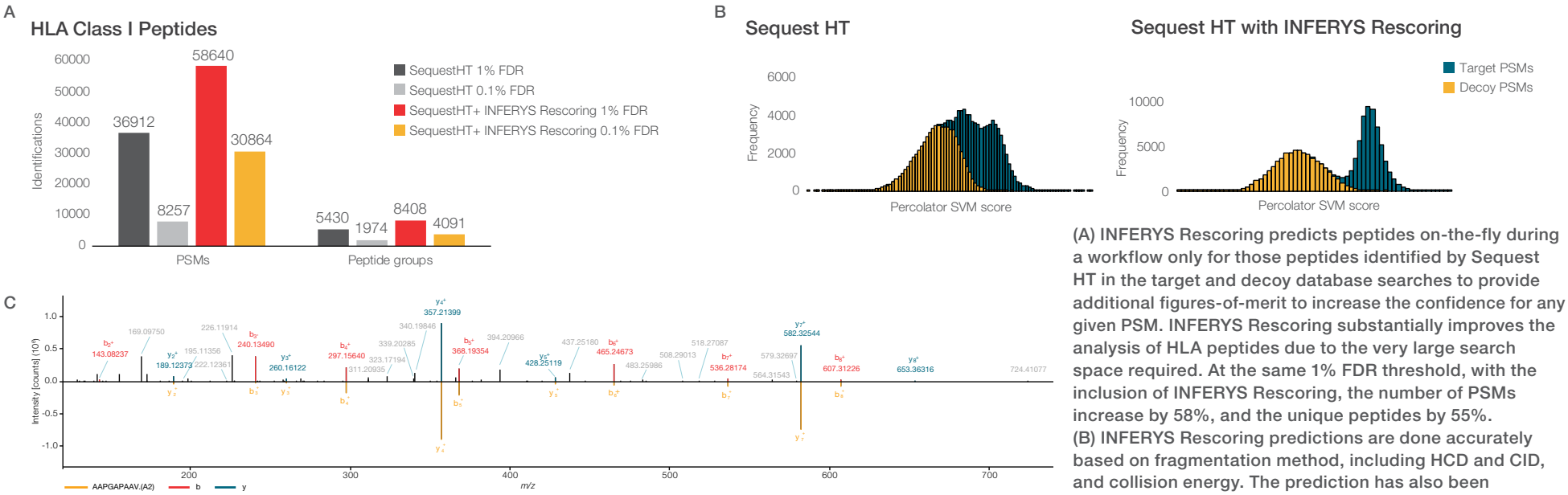
To improve the confidence in peptide identification for Sequest HT, the INFERYS Rescoring deep learning-based model calculates all predicted spectra from all target and decoy hits and calculates figures-of-merit for the correlation of predicted and measured spectra. This extra information generated by INFERYS Rescoring is added into the Percolator calculation for FDR thresholds. This results in an improvement in the number of unique peptides and proteins identified and quantified in typical proteomics datasets compared to other approaches.

While the additional confidence provided by INFERYS Rescoring increases the number of identified and quantified proteins for all proteomics applications, it provides specific benefits for single-cell proteomics, metaproteomics, and immunopeptidomics. By including a metric for the fragment ion intensities, INFERYS Rescoring can identify targets that are hard to separate from decoys based on classical scores.

Accurate predictions of fragment ion intensities enable the calculation of intensity-based scores, which increases the confidence in results. The additional confidence provided by INFERYS Rescoring is particularly important for immunopeptidomics, where there is a huge search space and many very similar peptides with similar properties. For example, in a HLA Class I data set from a patient derived melanoma cell line immunopeptidomics study, INFERYS Rescoring provides a 58% increase in PSMs and 55% increase in peptide identifications compared to Sequest HT alone. INFERYS Rescoring also increases the confidence in identifications when very low protein loads are used leading to low abundance MS^2 spectra, such as for single-cell proteomics. By adding extra figures of merit, peptides which may otherwise be missed can be rescued to improve the coverage and identify and quantify more proteins and peptides per sample. This results in 15–20% more peptide IDs versus Sequest HT alone.



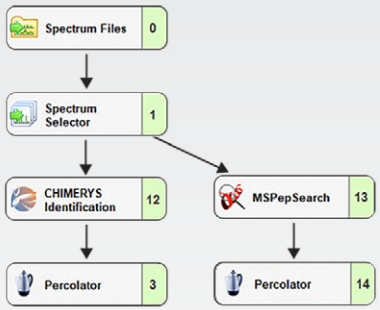
The INFERYS Rescoring workflow. Experimental MS^2 spectra are scored by Sequest HT by matching all theoretical b and y ions, each weighted equally by the search engine. The INFERYS Rescoring node subsequently reads the peptide sequences identified by Sequest HT, predicts their fragment intensities, and correlates the predicted spectrum to the original. Values corresponding to this correlation are included as features in Percolator, helping increase confidence for PSMs whose fragment abundances closely match those from the prediction.



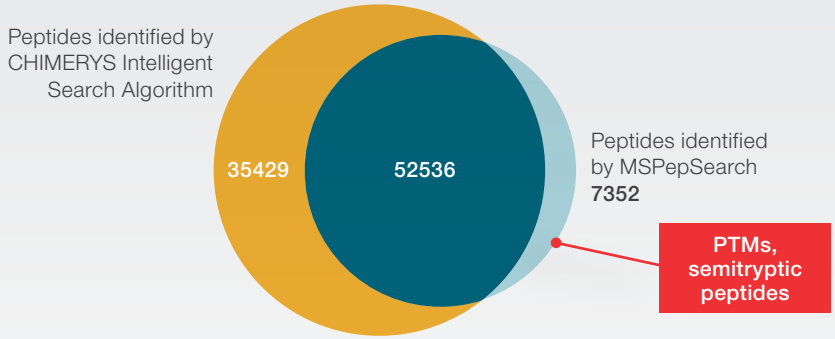
(A) INFERYS Rescoring predicts peptides on-the-fly during a workflow only for those peptides identified by Sequest HT in the target and decoy database searches to provide additional figures-of-merit to increase the confidence for any given PSM. INFERYS Rescoring substantially improves the analysis of HLA peptides due to the very large search space required. At the same 1% FDR threshold, with the inclusion of INFERYS Rescoring, the number of PSMs increase by 58%, and the unique peptides by 55%. (B) INFERYS Rescoring predictions are done accurately based on fragmentation method, including HCD and CID, and collision energy. The prediction has also been improved to include TMT and TMTpro. This results in an increased Percolator SVM score for well matched Target PSMs. (C) The mirror plot shows how close the predicted spectrum below matches to the experimental MS/MS spectrum above for the selected PSM.

Mix and match search engines with an extensible framework

The extensible framework of Proteome Discoverer software allows for flexible workflows and the deployment of multiple search engines. Sequest HT is the default search engine and is used to identify peptides in tandem mass spectra. Sequest HT evaluates protein sequences from the provided database and calculates a list of peptides that could be present given the precursor information. Sequest HT calculates a theoretical tandem mass spectrum for this list of candidate peptides and compares the theoretical spectrum to the experimental spectrum to yield the Xcorr score for ranking. In addition to Sequest HT, a series of third party search engines are available to use in Proteome Discoverer software.



The flexible workflows of Proteome Discoverer software allows for search engines to be used independently, in series, or in parallel to significantly increase the number of peptide and protein identifications.

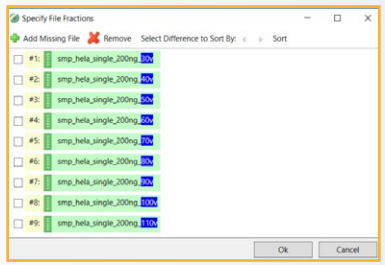


The extensible framework of Proteome Discoverer software workflows allows for the use of multiple search engines to glean more results from your raw data. Tailor search strategies to increase peptides even further by using library search or other search engines specializing in PTMs, such as MSPepSearch.

Label-free quantification workflows

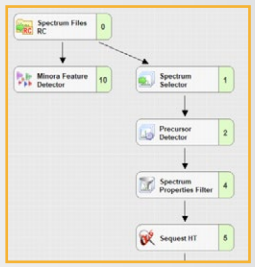
Support for single or multidimensional separations

- “Add Files” or “Add Fractions”
- For multidimensional data, RT alignment happens for the same fraction



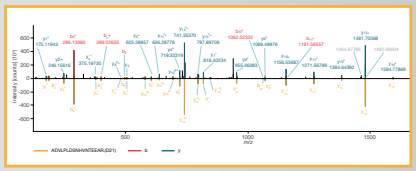
Predefined LFQ workflows

- Support for all Thermo Scientific Orbitrap mass spectrometers
- FAIMS Pro Duo Interface support



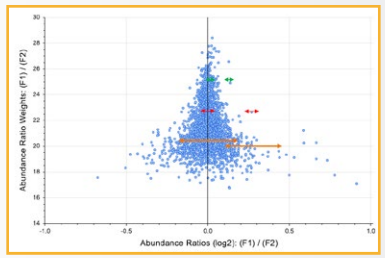
Peptide identification

- Multiple search engines
- Precursor Detector
- INFERYS Rescoring
- CHIMERYS Intelligent Search Algorithm



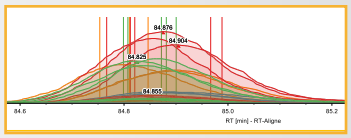
Unique approach for ratio and p-value calculation

- Protein ratios calculated as the median of peptide ratios including replicates
- “Background-based” p-value uses distribution of background ratios at different abundance levels to calculate p-value



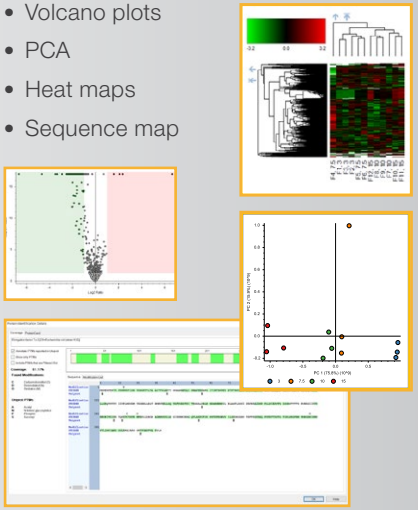
LFQ visualization tools

- Chromatograms, aligned and unaligned
- Abundance plots
- Extracted ion chromatograms



Statistical analysis and visualization tools

- Volcano plots
- PCA
- Heat maps
- Sequence map

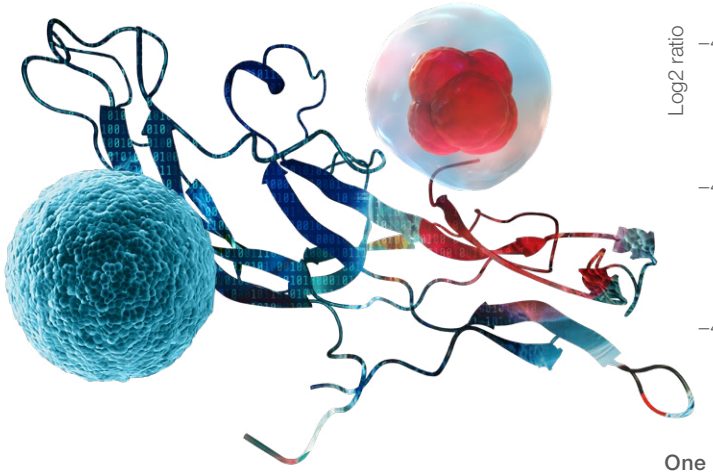


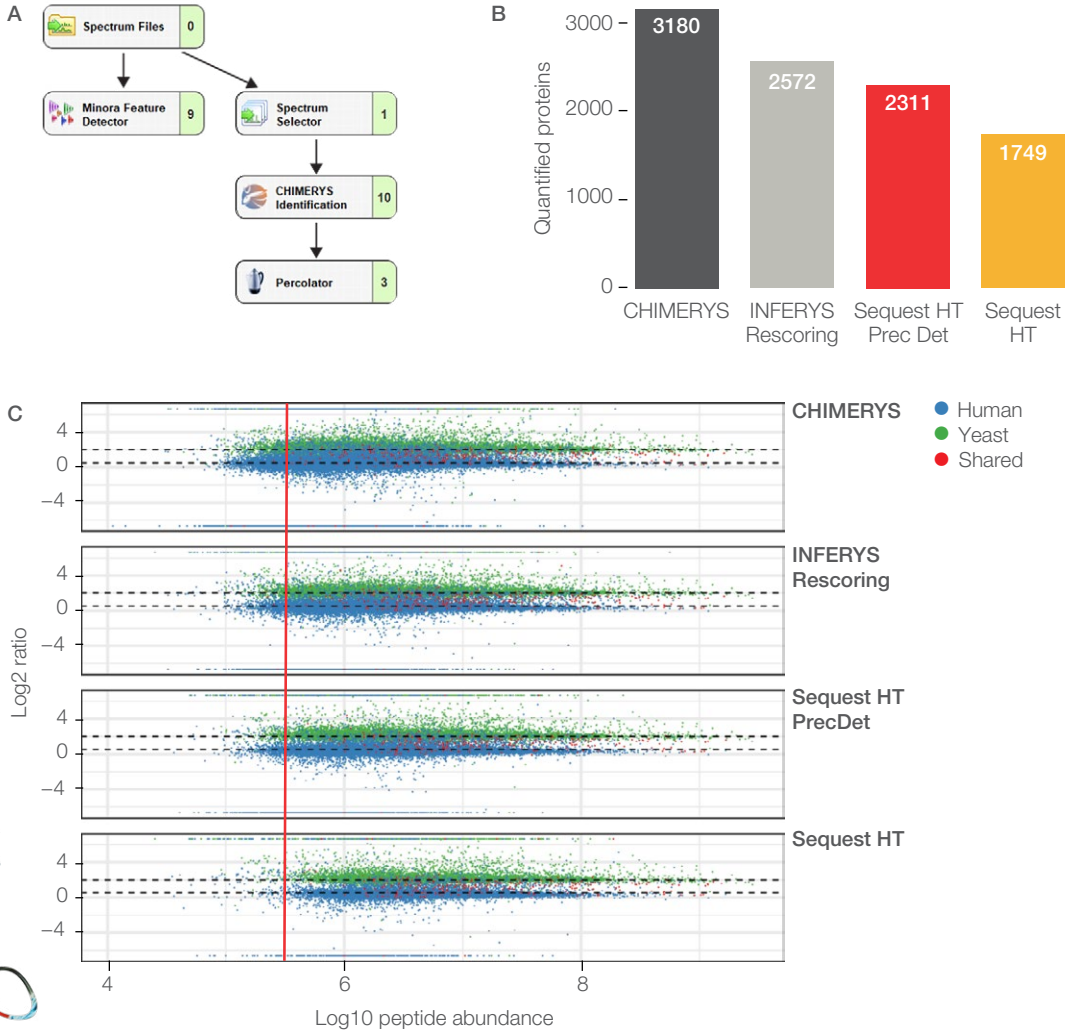
Comprehensive quantification capabilities

Proteome Discoverer software enables LFQ analyses with comprehensive capabilities. Both single or multidimensional separations can be utilized through the addition of files or fractions. Predefined LFQ workflows provide support for all Orbitrap mass spectrometers as well as support for data generated with the FAIMS Pro Duo interface. The Minora Feature Detector node detects and quantifies the isotopic clusters in the MS[1] data and then maps the clusters to PSMs from the search engines used in the workflow across all data files. Proteome Discoverer software also uses a unique approach for ratio and p-value calculations by using a pairwise ratio method to produce accurate ratios and a p-value calculation that uses the distribution of “background” proteins to determine whether or not any single protein is differentially expressed. To aid in the visualization and filtering of LFQ results the software also includes visualization and statistical analysis tools, such as abundance plots and heat maps, with direct connectivity to the results for dynamic filtering and plotting.

Identify and quantify more proteins and peptides with artificial intelligence

LFQ workflows can utilize the CHIMERYS intelligent search algorithm or Sequest HT with INFERYS Rescoring to identify and quantify more proteins. By digging deeper into the data to identify more peptides, CHIMERYS paired with the Minora Feature Detector quantifies more peptides and proteins. Similarly, adding INFERYS Rescoring into workflows using Sequest HT adds figures-of-merit from the comparison of the predicted spectra from INFERYS Rescoring for each target and decoy hit to the experimental fragmentation spectrum that produced those PSMs. Percolator uses this information to further separate the distribution of true matches from random ones to produce more identified and quantified peptides and proteins.





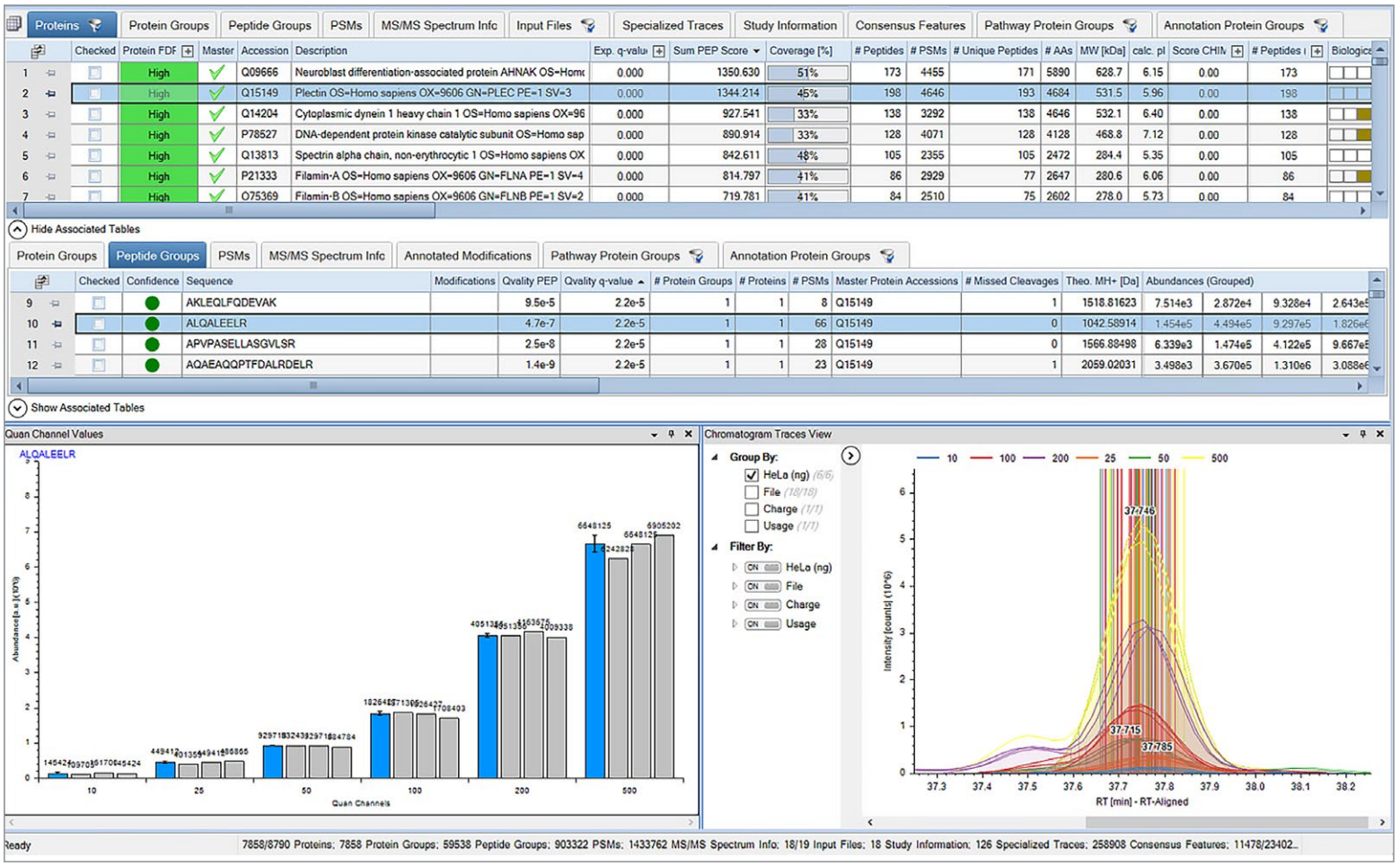
One hour data-dependent acquisition of 250 ng HeLa/125 ng Yeast lysate analyzed in triplicate on an Thermo Scientific™ Orbitrap Exploris™ 480 mass spectrometer. (A) A typical LFQ workflow using Minora Feature Detector and CHIMERYS Identification. (B) The CHIMERYS intelligent search algorithm increases the number of quantified proteins compared to other strategies for typical LFQ analyses. (C) The addition of the CHIMERYS intelligent search algorithm allows for increased identification and quantification, particularly for lower abundance peptides, resulting in improved sensitivity and a lower limit of quantification.

Integrated statistical analysis and visualization tools

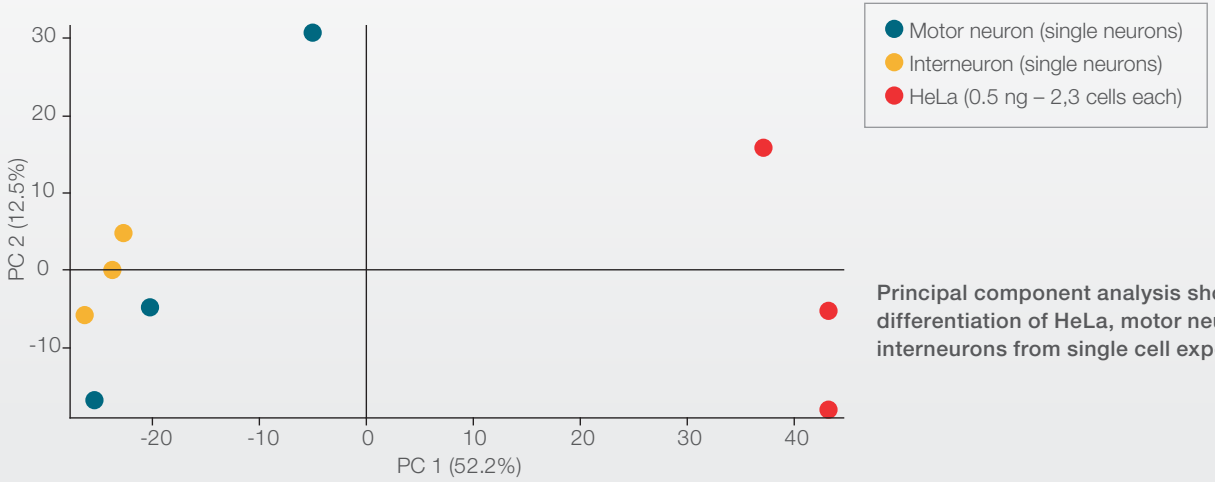
The statistical analysis and visualization tools built into Proteome Discoverer software allow for rapid filtering, visualization, and interpretation of results. LFQ results include chromatograms, Principal Component Analysis (PCA) plots, heat maps, quantification channel charts, and more. Quantification channel charts can quickly display changes for a given peptide or protein group across different experimental variables in tandem with chromatogram traces to visualize the detected peaks across all conditions.

Complex datasets in applications with samples containing diverse protein expression, such as single cell proteomics, can be quickly visualized using multivariate tools such as PCA plots and hierarchical clustering heat maps. These plots are highly interactive, making it easy to select and filter sets of peptides or proteins that differentiate groups. For example, the heat map and PCA plots can be used to classify different cell types in single cell proteomics data.[2]

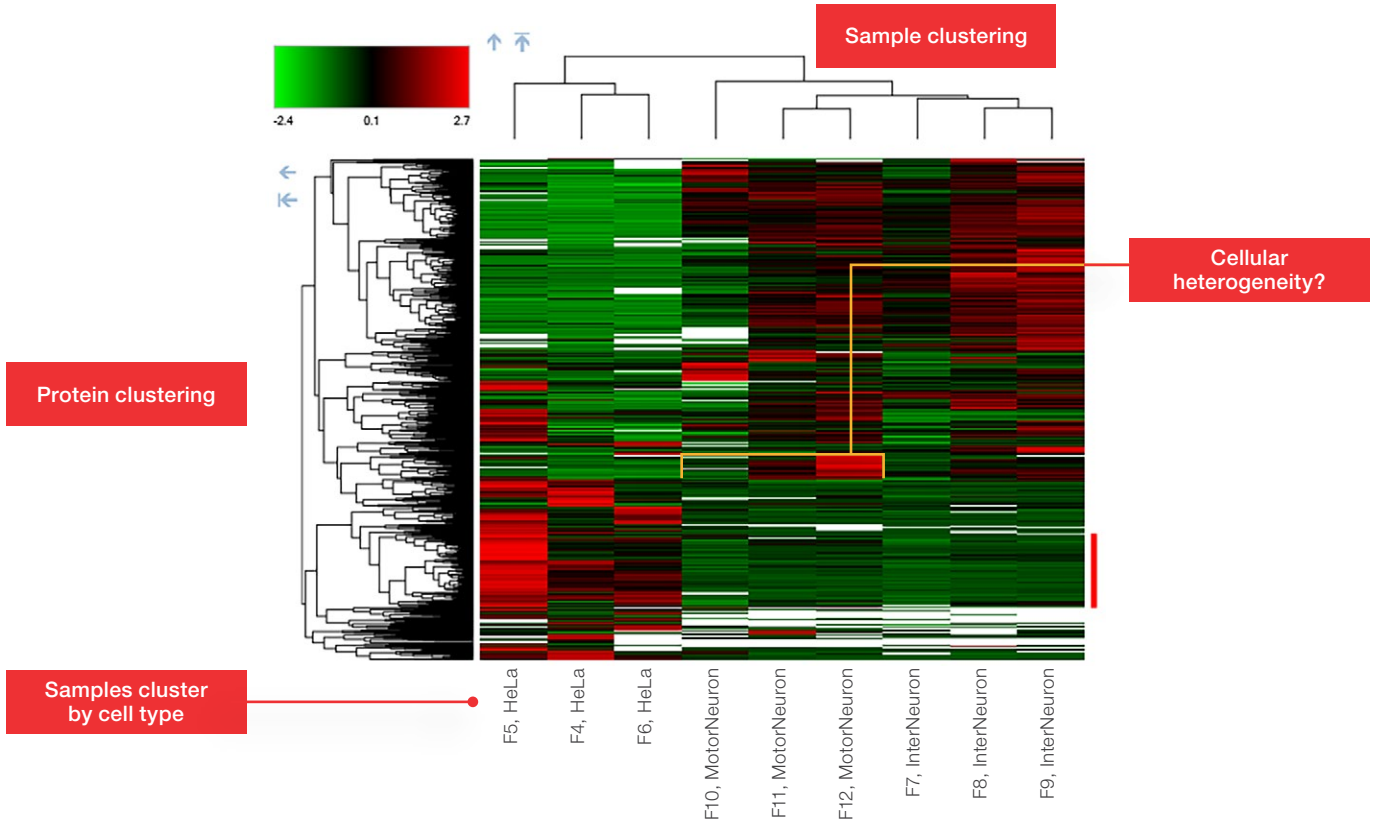
LFQ results presentation



LFQ results can be visualized with chromatograms, principal component analysis plots, heat maps, quantitation channel charts, and more within the Proteome Discoverer software framework, allowing for quick and easy interpretation of results.



Principal component analysis showing the differentiation of HeLa, motor neurons, and interneurons from single cell experiments.



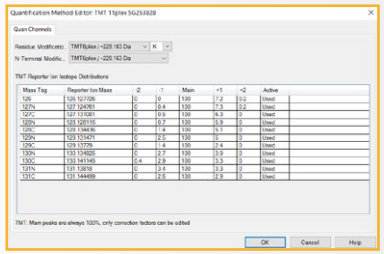
Heat map analysis showing the differentiation of HeLa, motor neurons, and interneurons from single cell experiments.

Isobaric Mass Tagging (TMT and TMTpro) Labeled Peptide Quantification

Key features/benefits of TMT quantification workflows

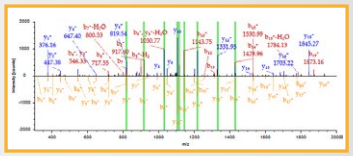
Default TMT methods

- TMT 2plex, 6plex, iodo TMT 6plex, 10plex, 11plex, and TMTpro 16plex and 18plex
- Correction factors



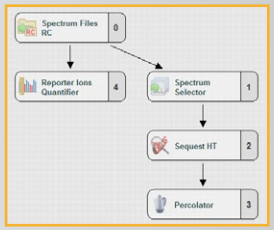
Support for advanced acquisition methods

- FAIMS
- TurboTMT
- TMT SPS MS^3



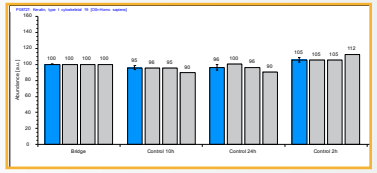
Predefined workflows

- Support for ion trap, hybrid and tribrid instruments
- MS^2 and MS^3-based quantification
- Customize to increase IDs



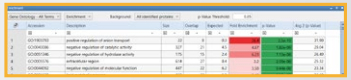
Unique approach for TMT protein quantification

- S/N-based
- Peptide filters for improved accuracy
- Summed protein abundances
- Scale to pool/control channel for larger datasets



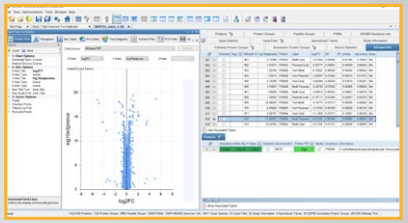
Powerful visualization and analytical tools

- PCA plots
- Volcano plots
- Heat maps
- Enrichment analysis



Extensibility

- Connect to external tools via the Scripting Node
- Integration with MSstatsTMT via custom script

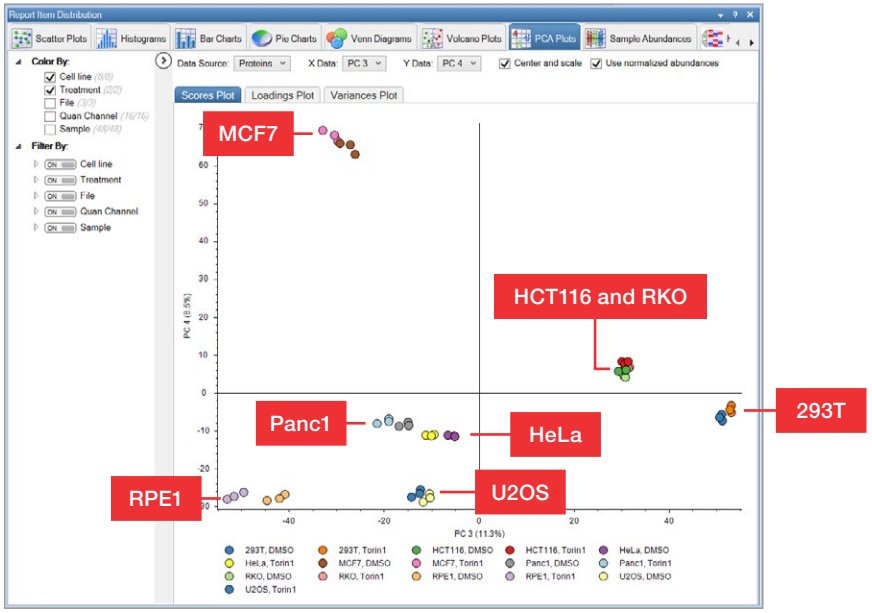


Integrate intelligent data acquisition strategies with sophisticated data analysis to maximize insights

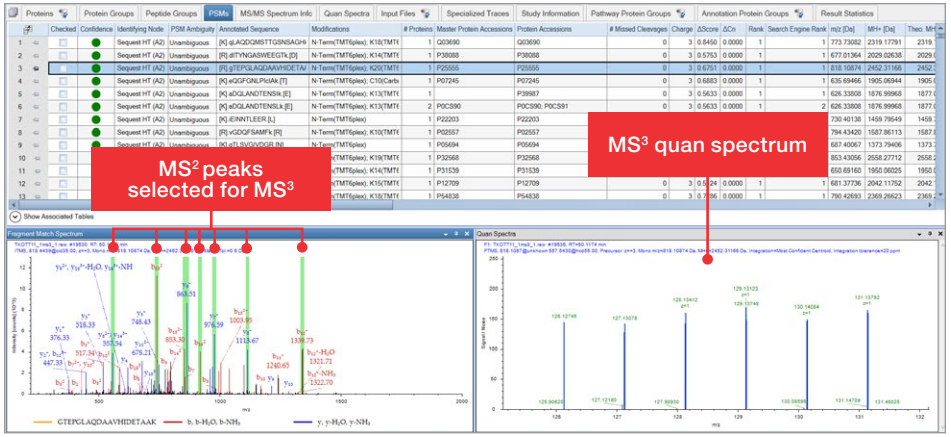
TMT and TMTpro reagents allow for the multiplexing of proteomics samples with robust and efficient derivatization for better quantitative accuracy and precision. Proteome Discoverer software offers default TMT methods to match the various reagents to provide the most accurate and precise results. Proteome Discoverer software offers predefined workflows to support all Orbitrap mass spectrometers with the capability to do both MS^2 and MS^3-based quantification, along with the ability to customize workflows to increase the number of identified and quantified peptides and proteins. TMT protein quantification is performed using signal to noise summed protein abundances from the peptide to protein level values for improved confidence in quantification accuracy and precision. Larger datasets with multiple runs are supported using scaling with pools or control channels. Results are easily visualized and filtered and can be customized using the Scripting Node, for example with integration of the MSstatsTMT R libraries.

Visualization of TMT and TMTpro results

Multiplexing samples across different groups can provide extensive information but can require the use of multivariate visualization to identify similarities and differences. For example, the comparison of eight different human cell lines with and without torin stimulation yields clustering by similarity that can be visualized with the built in PCA and Heat Map plots. Tying the study factors and samples together via study management paired with the built-in visualization provides clear differentiation between groups, with HCT116 and RKO as well as HeLa and Panc1 clustering together due to the similarity of the cell lines.



The multivariate analysis of different groups, such as multiple cancer cells lines as seen in this PCA plot, is integrated directly into the results visualization. Three replicates for eight different human cell lines each with and without torin stimulation that are quantified with TMTpro reagent labeling leads to clustering by cell lines that are similar and a slight shift in response to torin stimulation.[3]



TMT accuracy and precision is enhanced by utilizing TMT SPS MS³ methods and the number of identified peptides is improved with Real-Time Search-Aided methods. Proteome Discoverer software results allow for the visualization of MS² fragment ions that are selected for MS³ and the visualization of the MS³ quantitation spectrum for maximum confidence in results.

Match Real-Time Search-Aided TMT SPS MS³ results with the Comet search engine

The Real-Time Search-Aided TMT SPS MS³ functionality uses the Comet search engine for intelligent triggering of MS³ scans based on the identification of a PSM in the preceding MS² scan to improve accuracy, increase the number of quantified proteins, and provide interference free spectra. The Comet search engine has been added into the Proteome Discoverer software framework to more closely match the peptides identified during acquisition and can be used to process data alone or in conjunction with other search engines.

Improved identification and confidence for TMT workflows

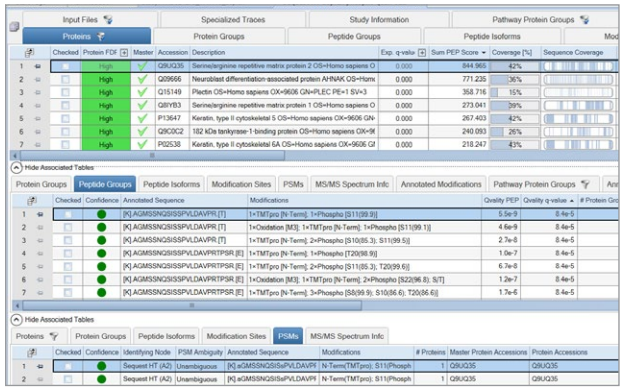
Adding artificial intelligence capabilities to TMT workflows allows for additional identifications and more confident quantification. Using the CHIMERYS intelligent search algorithm in MS²-based TMT workflows will provide confidence that peptides used for quantification are not derived from chimeric spectra, which would result in errors in quantification. Both CHIMERYS and INFERYS Rescoring can also help improve the number of identified and quantified peptides and proteins by digging deeper into the data.

Data visualization

Proteome Discoverer software offers an intuitive interface for results interpretation, including exceptional connectivity between data tables and plots to allow rapid interpretation, filtering, and visualization for a faster time to results.

Peptide views

Results are presented using a nested table where PSMs are directly linked to unique peptides, proteins, and annotated MS^n spectra in an interactive fashion. This allows for dynamic filtering and connectivity across tables, graphical charts, spectra, and chromatograms for more thorough investigation and interpretation.

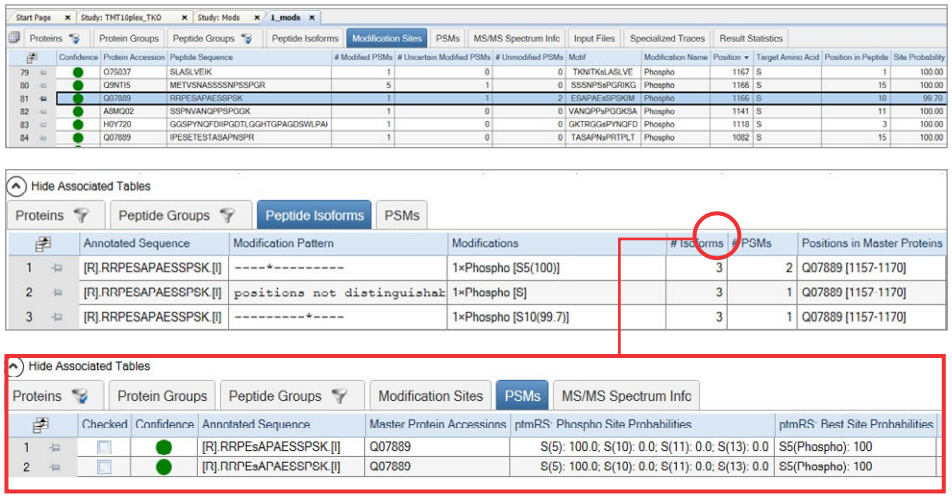


The hierarchical nesting of tables helps organize results to link proteins, peptides, and PSMs together to simplify the navigation through complex results.

Post translational modification identification and visualization

Understanding the function of proteins requires the identification of not only unmodified peptide sequences but also the localization and presence of PTMs. The Proteome Discoverer software framework includes the IMP-ptmRS search node, which uses the peptide sequence determined by the main search engine, calculates the masses of residues considering all possible positional isoforms, identifies and detects the diagnostic fragment ions, scores the hits, and calculates a probability of localization for each PTM.[3] In addition, Proteome Discoverer software includes a chemical modifications table and modifications dialog to easily annotate additional peaks in tandem mass spectra for neutral losses and/or diagnostic ions. This unique capability streamlines the analysis and interpretation for PTMs of interest. Further insights can be gleaned by using an error tolerant search through the Mascot node. The linked Modification Sites, Peptide Isoforms, and PSMs tables can be used to find differentially expressed PTMs in complex proteomics data.

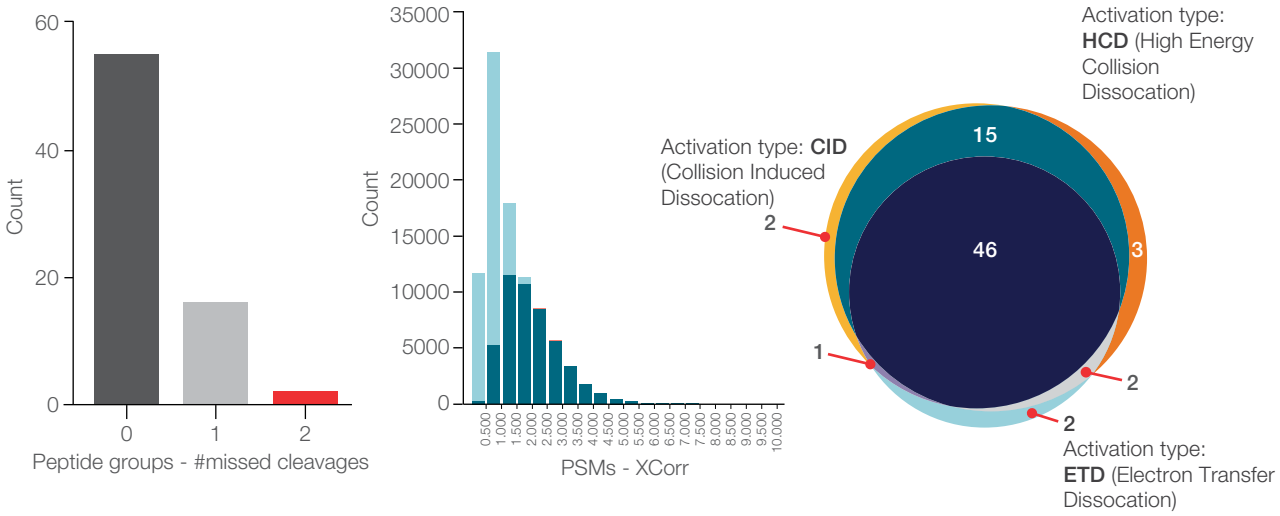
The protein sequence coverage view shows the location of identified peptides as well as identified and localized post-translational modifications. If the Protein Annotation node is used in the Consensus workflow, the PTMs annotated in UniProt for the selected protein will be displayed.



PTMs can be easily identified and visualized through the modification sites table, which allows for the selection of peptide sequences and the display of peptide isoforms. Within each peptide isoform the modification pattern can be identified and annotated through the use of ptmRS site probabilities and best site probabilities score for site specificity and increased confidence in annotation.

Deeper insights into data

Proteome Discoverer software offers multiple plotting capabilities related to both raw data and data processing results. Information about raw data can be used to optimize sample preparation or acquisition strategies and can also be used to troubleshoot. Distribution charts can display any metric with scatter plots, histograms, and bar charts. For example, the number of missed cleavages can be displayed in a bar chart to confirm full protein digestion or a histogram chart of Xcorr values can be used to show which PSMs pass the FDR threshold filter, or a comparison of identified peptides using different fragmentation techniques can be seen with a Venn diagram. These retrospective analyses help offer a more complete understanding of the generation and processing of data to move forward with informed decision making to deploy optimized sample preparation, acquisition, and processing strategies to achieve the best data quality and results

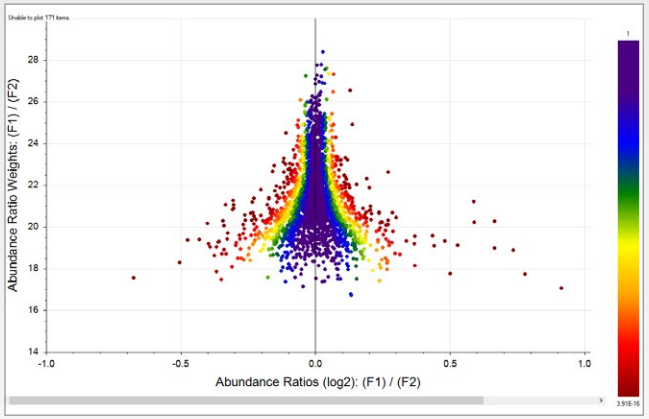


Distribution charts can be utilized to visualize, interpret, and optimize experimental procedures, data acquisition settings, and results. For example, a bar graph can display the number of missed cleavages, a histogram can show the Xcorr values for PSMs passing the pre-set FDR filter (dark red) and excluded hits (salmon), and Venn diagrams can display peptide sequences identified by different fragmentation techniques.

Integrated statistical analysis tools

Interpreting the results of large proteomics experiments requires the ability to map study factors and samples to relevant statistical tests and ratio generation with flexible filtering and visualization capabilities to determine the relevant experimental results. Proteome Discoverer software allows for a hierarchy of study factor variables to handle biological factors, experimental categorical factors, and technical replicates. In order to provide the best quantitative and qualitative results, normalization and scaling can be performed along with protein ratio calculations based on either protein abundance or pairwise ratios. Statistical analyses can be performed with a t-test. Results can also be visualized using volcano plots to show pairwise comparisons of proteins or peptides with adjustable P-value and log2 fold change values to select and filter based on desired changes.

Confidence interval chart with abundance ratios



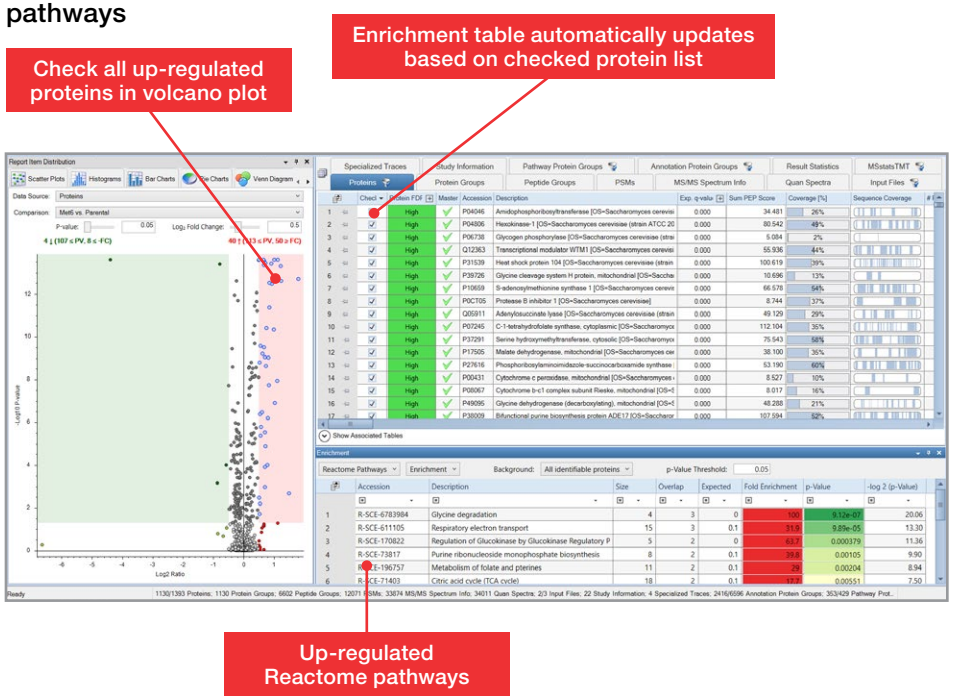
Volcano plot

Interactive abundance ratios comparing groups can be visualized with shading according to the confidence interval, allowing for a more confident interpretation of proteome changes across a large dynamic ratio range using a background-based p-value calculation. Volcano plots allow for the visualization and selection of proteins or peptides in comparisons with customizable p-value and log2 fold change values.

Annotation and pathway mapping

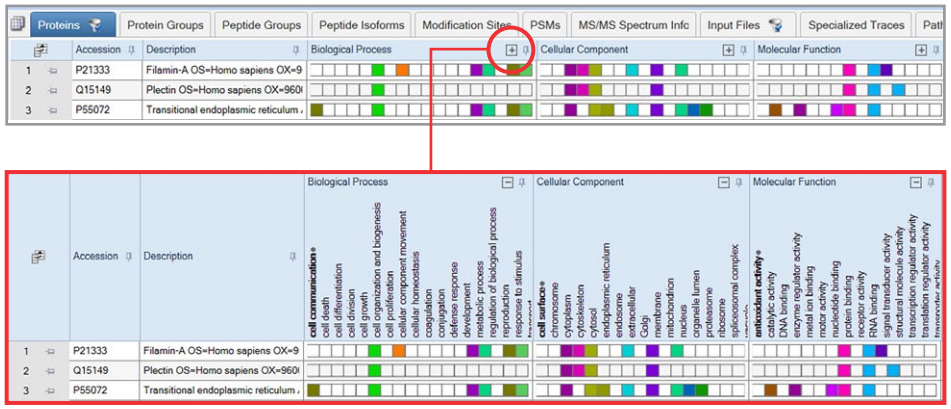
In addition to the identification and quantification of peptides and proteins, Proteome Discoverer software provides biological interpretation of results. The addition of the Protein Annotation node to the consensus workflow creates the Pathway Protein Groups and Annotation Protein Groups result tables. In addition, proteins can be selected from Volcano Plots, PCA plots, and Heat Maps to generate an Enrichment table to display overrepresented or underrepresented pathways, GO, and Pfam terms to quickly convert results into biological insights. The annotation service is regularly updated to provide the most up to date terms. The service also allows full proteome FASTA file downloads and allows for easy updates over time as entries are modified in the database.

Interactive enrichment table from graphical results to find overrepresented pathways

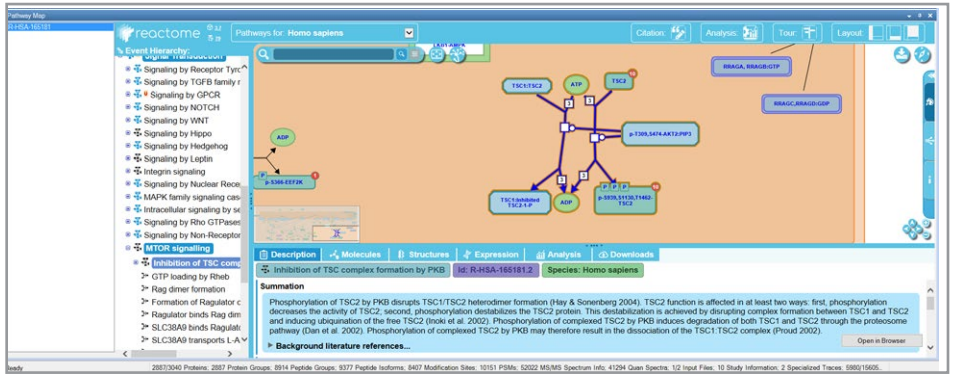


Interactive enrichment tables allow for the selection of up- or down-regulated proteins or peptides in any view, including volcano plots, PCA plots, and heat maps, and the visualization of corresponding Reactome pathways from the checked proteins to help derive biological interpretation of results. In addition to Reactome pathways, KEGG and Wikipathways as well as gene ontology terms and protein families can be visualized in the enrichment table.

GO Annotation: identify biological processes, cellular components, and molecular functions. Map results onto pathways to understand biological context.



An annotation service with frequent updates allows for the display of overrepresented pathways, GO, and Pfam terms associated with selected proteins.



Pathways can also be directly visualized, providing biological context and background literature references using pathway maps.

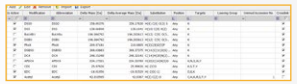
Protein crosslinking with XlinkX

Combine crosslinked peptide identification with TMT and LFQ analysis for quantitative structural biology insights

Key features for XlinkX for Proteome Discoverer software

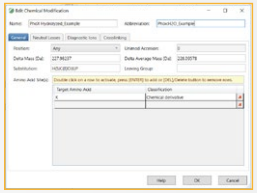
Crosslink modifications

- Predefined crosslinks
- Cleavable and non-cleavable
- Heterobifunctional crosslinkers



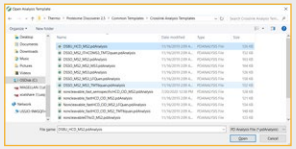
User interface for crosslink modification definition

- Users can add their own cross-linker



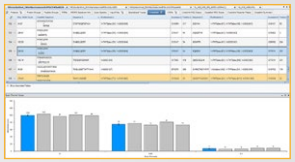
Predefined workflows

- Cleavable and non-cleavable
- Support for different acquisition modes



LFQ and TMT quantification

- Uses same nodes as for standard proteomics workflows

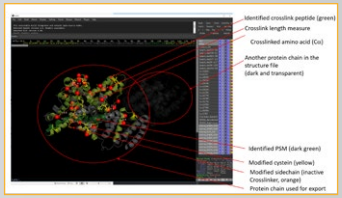


Annotated XlinkX spectra

- Chromatograms, aligned and unaligned
- Abundance plots
- Extracted ion chromatograms

Export to third party tools

- xiView
- PyMOL



Protein crosslinking enables structural proteomics studies to be performed using mass spectrometry to pinpoint residues in close proximity to the interaction interfaces between protein complexes and interfaces. The XlinkX nodes in Proteome Discoverer software allow for easy analysis of crosslinked peptide data, supporting both cleavable and non-cleavable crosslinkers. These nodes can also be combined with the TMT and the LFQ nodes in Proteome Discoverer software for quantification. * XlinkX developed in collaboration with Utrecht University.

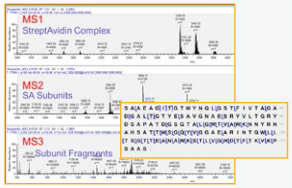
Top-down proteomics with ProSightPD software

Comprehensive top-down data analysis and visualization, supporting multiple fragmentation modes, FAIMS Pro Duo interface, and detailed proteoform analysis

Key features of ProSightPD software

Prebuilt and Flexible Top-down workflows

- Discovery LC-MS
- Native top-down
- Complex-down MS,[4] and more

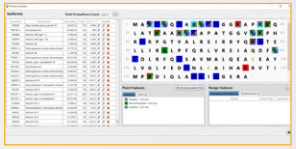


Supports advanced acquisition methods and quantification

- PTCR
- LFQ
- ETD, EThcD, UVPD
- FAIMS Pro Duo Interface

Powerful database management

- ProSight Annotator
  - Precise control of all proteoforms
  - Intuitive user interface



Versatile proteoform identification

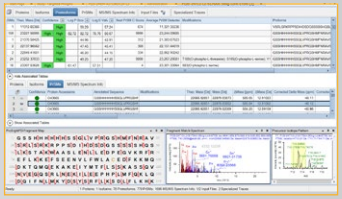
- Fast ProSight database search algorithms
  - Searches include truncated forms and unknown modifications via Δm mode.
  - Supports unresolved precursors

Unique top-down results visualization

- Visualizations from Proteome Discover software
- TD Validator
- Fragment Map View

Extensive reporting capabilities

- TD Viewer
- ProForma Annotated Sequences



ProSightPD software enables the identification, characterization, and quantification of Top-down proteomics datasets within the Proteome Discoverer software framework. ProSightPD software includes a Database Manager to import proteins of interest and add or remove isoforms or PTMs in tandem with workflows equipped to address the challenges of finding and annotating known and unknown proteoforms, including processing capabilities for all fragmentation modes such as HCD, CID, ETD, EThcD, UVPD, and PTCR, along with FAIMS Pro Duo interface data. Mix and match search types and deconvolution strategies to match acquisition and processing strategies to identify intact proteoforms and sub-sequences of proteoforms. Visualize and filter results with the capabilities of Proteome Discoverer software including automated discovery and characterization of proteoforms and instantly validate results with TD Validator.

Almanac web-based monitoring and management

Stay connected to your science. See how the Thermo Scientific™ Almanac™ application can help you get the most out of your instruments.

thermofisher.com/almanac

Technical and online support: peak performance for your instruments

Helping you keep your instruments running at peak performance is our goal. Whether you're looking for an instrument manual or spare parts, want to submit a repair request, or check on the status of your warranty or service contract, we have every support option you're looking for.

thermofisher.com/technicalresources

Stay up and running with unparalleled laboratory services

Unity™ Lab Services provides a single source for integrated instrument service and support as well as laboratory and supply management. Our customized service offerings and highly experienced world-class service experts have the flexibility and expertise to address your lab's unique needs. We provide a complete portfolio of services and support designed with the MAM workflow in mind to help you improve productivity, reduce total cost of ownership, and help ensure performance of your instruments and workflows.

unitylabservices.com/serviceadvantages

Thermo Scientific Orbitrap mass spectrometers
thermofisher.com/orbitrap

Thermo Scientific Vanquish Neo UHPLC system
thermofisher.com/vanquishneo

[1] Schweppe et al, Anal. Chem. (2019), 91, 6, 4010–4016.

[2] https://pubs.rsc.org/en/content/articlehtml/2021/sc/d0sc03636f

[3] Li et al, Nature methods https://doi.org/10/1038/s41592-020-0781-4

[4] Taus, T. et al, J Proteome Res., 10, 5354-5362 (2011). doi: 10.1021/pr200611n

Learn more at **thermofisher.com/proteomediscoverer**