# PEAKS®

## Online XPRO

HIGH-THROUGHPUT, MULTI-USER, PROTEOMICS LC-MS/MS ANALYSIS SOFTWARE

**PEAKS** is a specialized tool that offers unrivalled peptide identification rates driven by deep learning technology and integrated with library search, direct database search and *de novo* sequencing.

As a vendor-neutral computing platform, it is capable of directly loading raw mass spectrometry data and standard data formats. Deploy the PEAKS Xpro workflows, from *de novo* sequencing, PEAKS DB (database search) identification, PEAKS LIB (spectral library search), PEAKS PTM (post translational modification) analysis, and SPIDER homology search to identify the presence of peptides and proteins in your project. Quantification analysis by labeling and label-free quantification (LFQ) can also be performed using the PEAKS Q addon module. Intuitive result visualization tools are provided at every stage of analysis and results can be exported. PEAKS Online Xpro supports data-dependent and data-independent acquisition analyses (DDA and DIA respectively) and ion mobility mass spectrometry (IMS-MS).

## EXAMPLE CASE STUDY

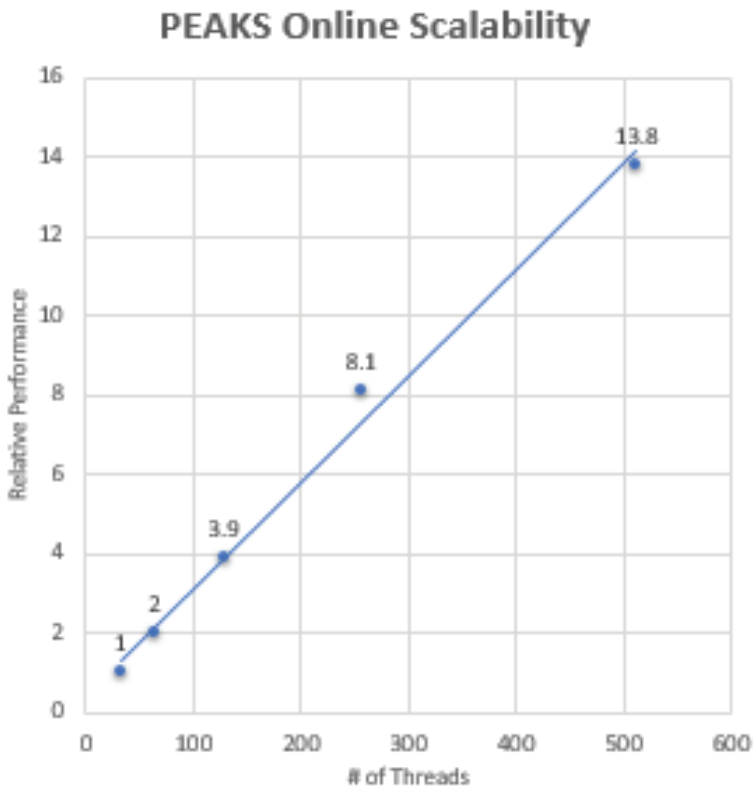
| Data size | |
|---|---|
| # of Samples | 56 |
| Total MS runs (180 min/run) | 672 |
| MS | 5,106,542 |
| MS/MS | 28,858,408 |

| Time to complete analysis of Data | CPU Cores |
|---|---|
| 17h11m | 512 |
| 1 day 4h32m | 256 |
| 2 days 1h24m | 128 |
| 4 days 19h55m | 64 |
| 9 days 15h10m | 32 |

Use PEAKS Online Xpro to take advantage of powerful and shared computing resources to perform LC-MS/MS protein & peptide identification and quantification analyses. The restructured platform allows large datasets to be processed efficiently by multiple users at the same time; with the ability to run on any cluster, multi-CPU machine, or cloud server.

# PEAKS SERVER-BASED SOLUTION KEY FEATURES

Re-running analyses has never been so easy. With PEAKS Online, the algorithm will determine the quickest way to reanalyze your results and use existing results, if applicable, instead of running the whole dataset again.




Data monitoring for acquisition machine allows real-time processing of data as it comes off the mass spectrometer.

Allow concurrent access from multiple users to process multiple projects in parallel.

High performance and advanced algorithms to provide a complete, vendor-neutral solution for discovery proteomics, including protein/peptide *de novo* sequencing, identification and quantification.

Centralized configuration and monitoring system to easily maintain, prioritize and share all PEAKS Online data analyses.

## Advanced System Architecture

Built on top of latest technologies to fully utilize the computing power of your hardware to provide:

High throughput solution: Allows concurrent access from multiple users to support parallelism at project and data level.

Distributed database: Yields higher I/O performance and better fault tolerance.

# Don’t get left in the dark!
# Work together as a cohesive research group

PEAKS Online means high-throughput data processing on a shared resource. This server-designed proteomics software is fully parallelized with the ability to run on a cluster of multi-CPU machine or cloud server.

Users are able to run the same proven algorithms included in the PEAKS Studio solution, efficiently and on a larger scale.

By using a web interface client, users can send/retrieve data to/from the server and view the results, on any operating system, in an intuitive manner.

Ready to scale: Vertically and horizontally, add new worker node(s) or database node(s) to the system and see performance improvement right away.

Cross-platform deployment:
Deploy the server on any Windows or Linux systems.

Dual interfaces: The command line interface offers the abilty to automate dtaa analysis workflows and result exporting while the web interface provides a graphical user interface to visually configure workflows and easily assess results in detail.

# Super sensitive peptide identification rates driven by deep learning technology and integrated with library search, direct database search and *de novo* sequencing

Generally speaking, there are three common ways to interpret tandem mass spectrometry data: database search, *de novo* sequencing and spectral library search:

**DATABASE SEARCH:**
Given a spectrum and a protein sequence database, find a peptide in the database that has the best match with the spectrum.

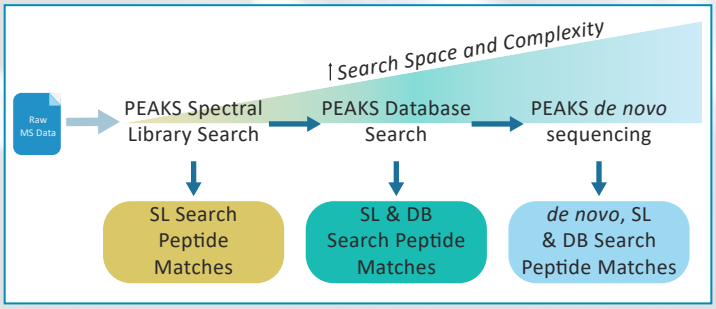
**SPECTRAL LIBRARY SEARCH:**
Given a spectrum and spectral library, find a peptide in the spectral library that has the best match with the spectrum.

***DE NOVO* SEQUENCING:**
Given a spectrum, find a peptide that has the best match with the spectrum.

1. When the peptides are believed to be in a protein sequence database, then a database search approach is preferred.
2. When studying a particular proteome, a peptide spectral library for the targeted biological system being studied can be used to focus your analysis.
3. However, when such a sequence database or spectral library is unavailable, *de novo* sequencing is needed to derive the peptide sequence directly from the spectrum.

PEAKS Online introduces a new identification workflow that combines spectral library search, database search and *de novo* sequencing to provide an in-depth analysis. This unique pipeline provides an innovative solution to focus the analysis on peptide spectral library to gather the targeted information for the research purpose. After performing a quick screen, users can go beyond the spectral library by enabling a database search and/or PEAKS' *de novo* sequencing in a single easy to use pipeline.

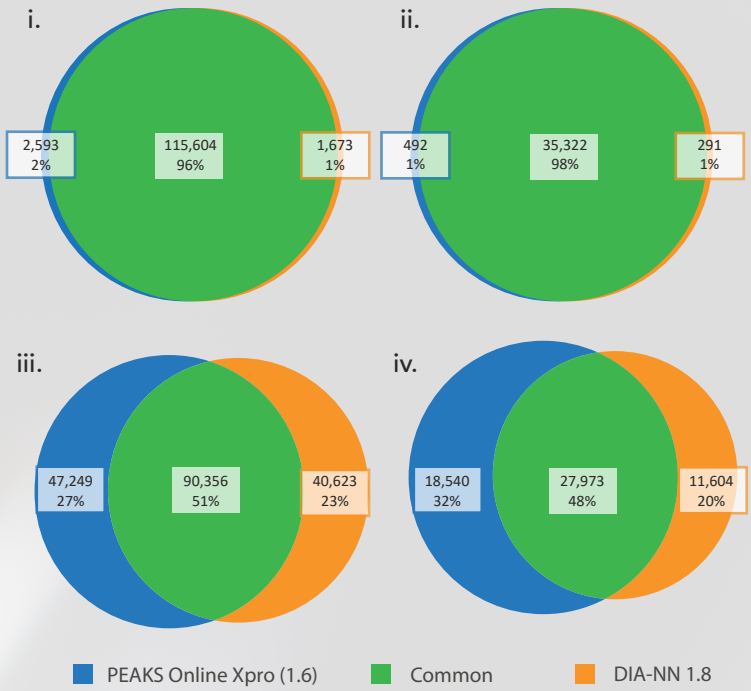
| Identification Workflow | 1% Precursor Count | 1% Peptide Count |
|---|---|---|
| POL LB | 118,197 | 35,814 |
| POL LB + DB | 177,445 | 49,875 |
| POL LB + DB + *de novo* | 208,000 | - |
| DIA-NN 1.8 LB | 117,277 | 35,615 |



## Deep learning advances accuracy and sensitivity of data analysis

PEAKS Online uses advancements in deep learning-based spectrum prediction models to perform both DDA and DIA database search by predicting the retention time and spectra in silico for each plausible peptide. In addition, PEAKS furthers the use of deep learning to perform *de novo* sequencing for both DDA and DIA data, which could help identify polypeptides from out-of-frame ORFs for example.

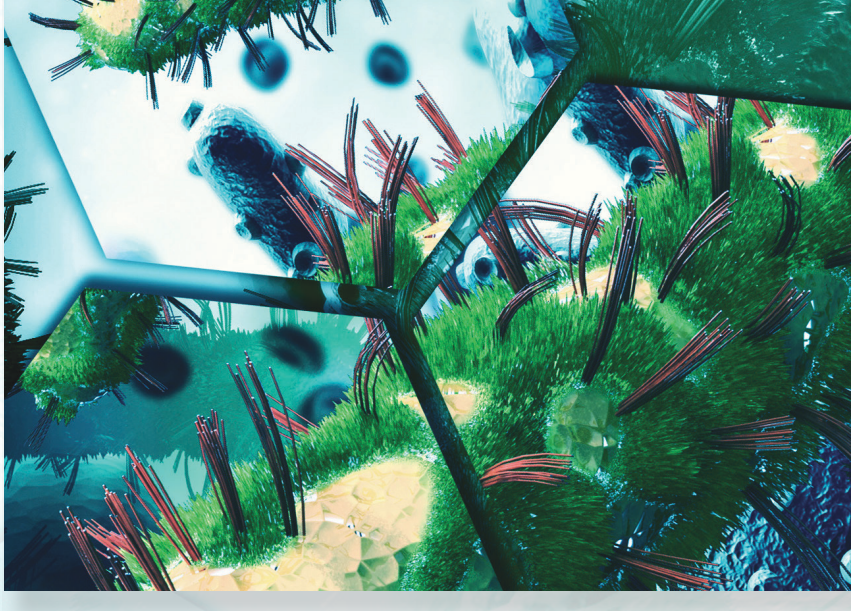
By integrating spectral library search, database search and *de novo* sequencing into a single workflow, PEAKS offers accurate and comprehensive analysis on all types of data. Benchmarking data shows the exceeding peptide identification rates.




2019 ABRF Multi-Laboratory Data-Independent Acquisition Study data used for benchmarking. Sample set consisted of tryptic HeLa digest spiked with four exogenous proteins. Identification results were filtered at a 1% FDR to compare the number of (i) identified precursor peptides using a spectral library search, (ii) identified peptide sequences using a spectral library search, (iii) identified precursor peptides using a database search,and (iv) identified peptide sequences using a database search.

## To overcome the boundaries of discovery, PEAKS introduces the first *de novo* sequencing approach for DIA to provide truly non-biased results

# PEAKS' *de novo* peptide sequencing is well-recognized worldwide and the base of all PEAKS analyses



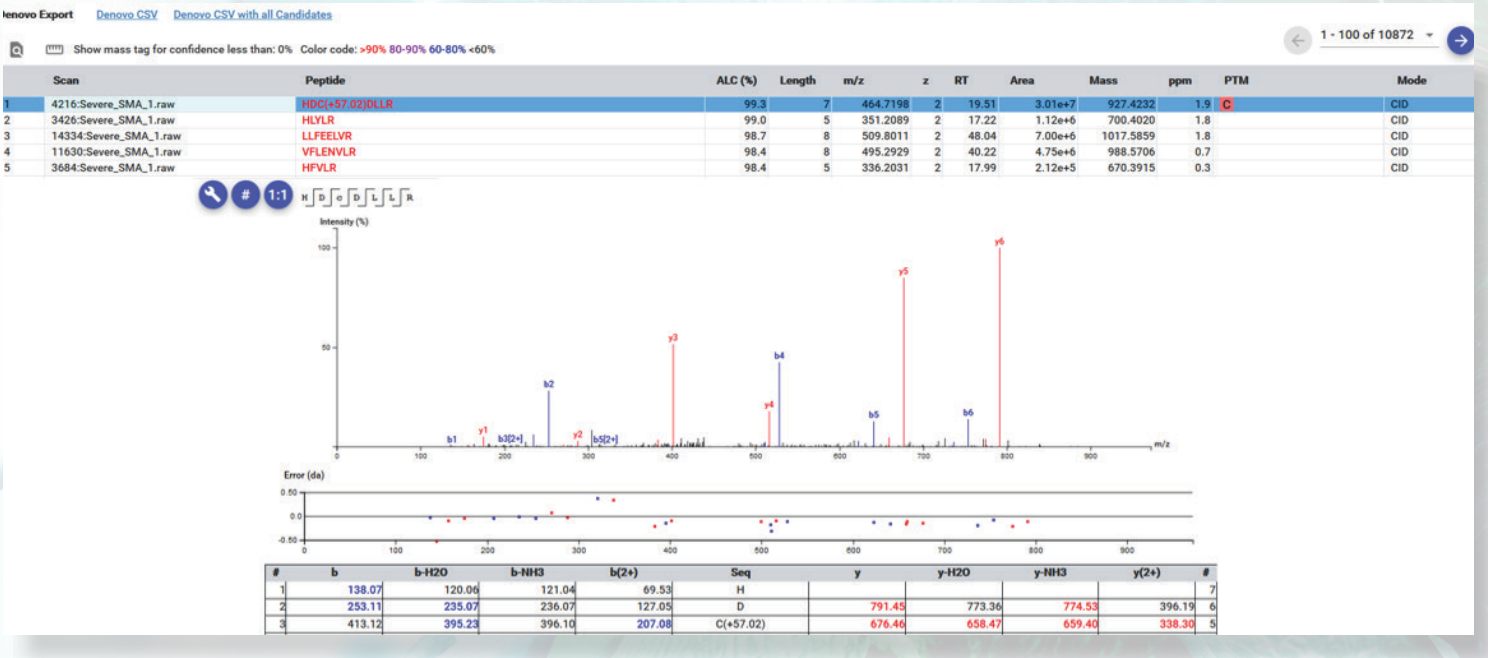
The innovative PEAKS' *de novo* sequencing algorithm accurately constructs a peptide sequence without the use of a database. These sequences are then used to enhance the PEAKS DB, PEAKS PTM, and SPIDER analyses.

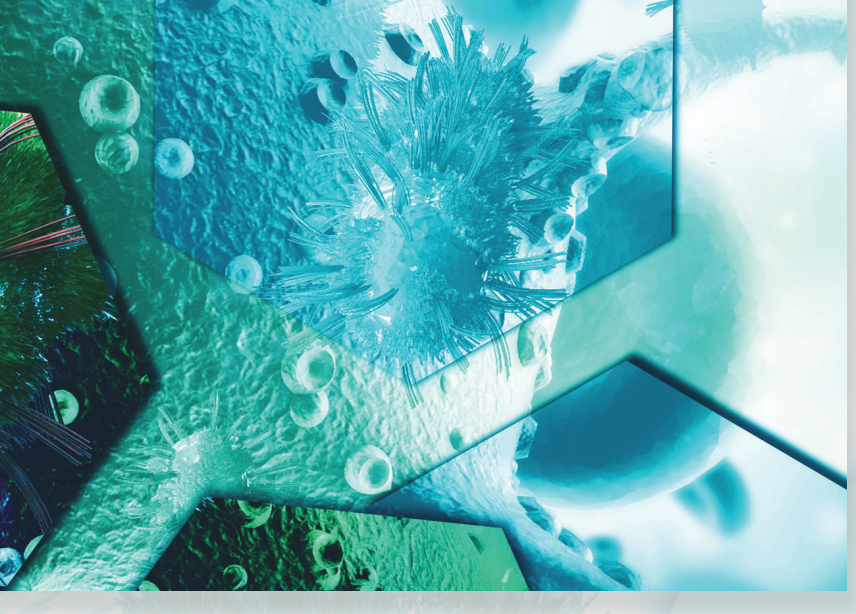
## Examine accuracy at the amino acid level

Local confidence scores are assigned for each amino acid. You can easily separate confident amino acid assignments from false positives.

## Find novel peptides not recorded in protein database

*De novo* results from scans missed in protein databases are summarized in 'de novo only' results. Partial protein matches or "*de novo* tags" are also given and can be viewed directly.

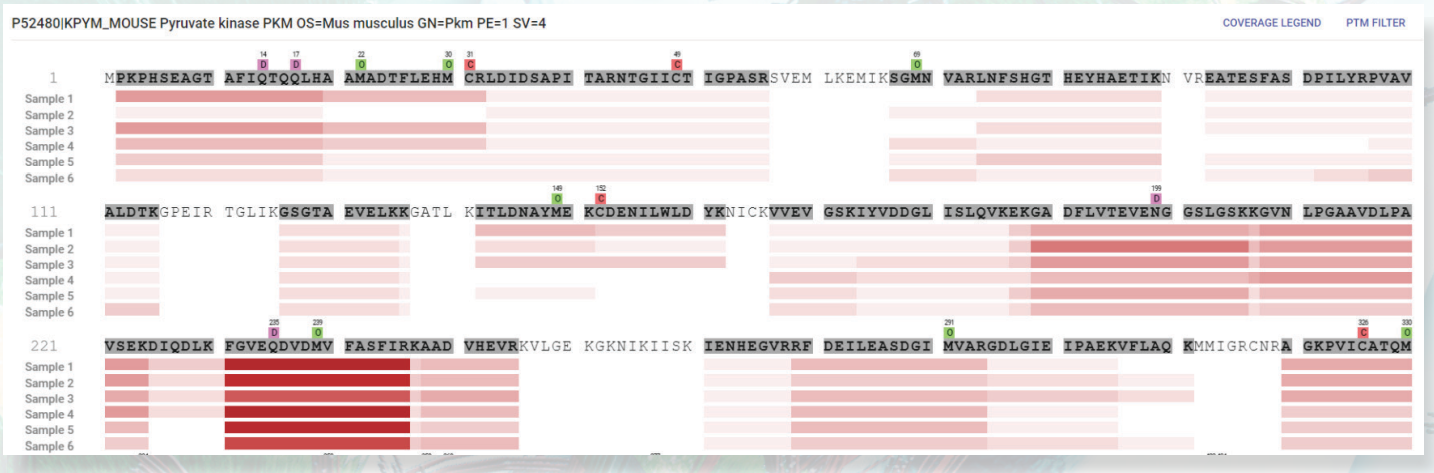




**PEAKS enhances the separation of true/false hits by integrating *de novo* sequencing into a database search workflow. This unique approach identifies more peptides and proteins with increased confidence.**

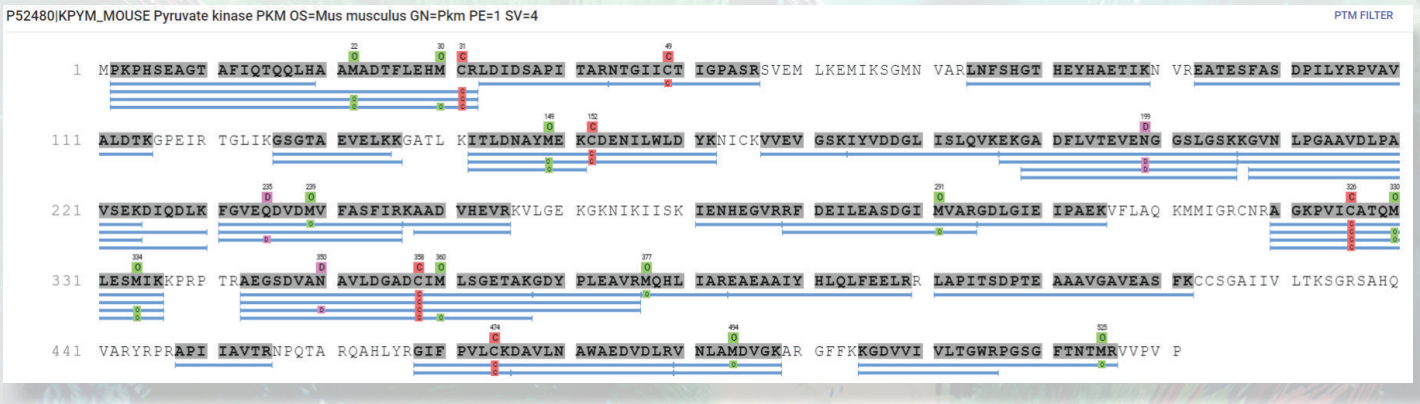
## Protein coverage heatmap for quick and easy comparison across multiple samples

Easily compare between multiple samples in a project using the PEAKS Online protein coverage heatmap. The increasing colour intensity indicates a higher abundance of supporting spectra within the corresponding sample.



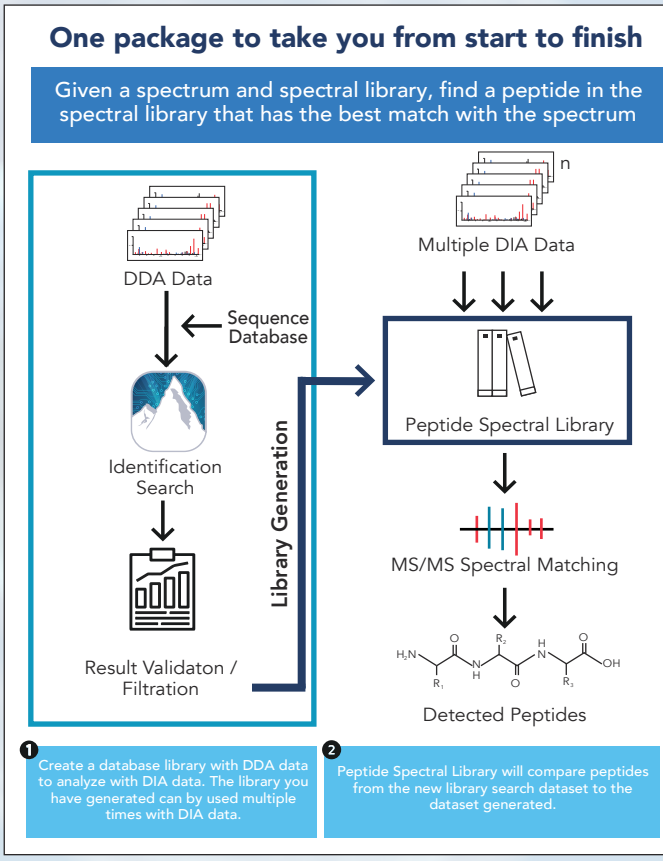

## No need to sacrifice details when analyzing large datasets

Interested in the protein coverage in a particular sample? PEAKS Online allows users to select an individual sample to view the detailed coverage information, just as in PEAKS Studio.

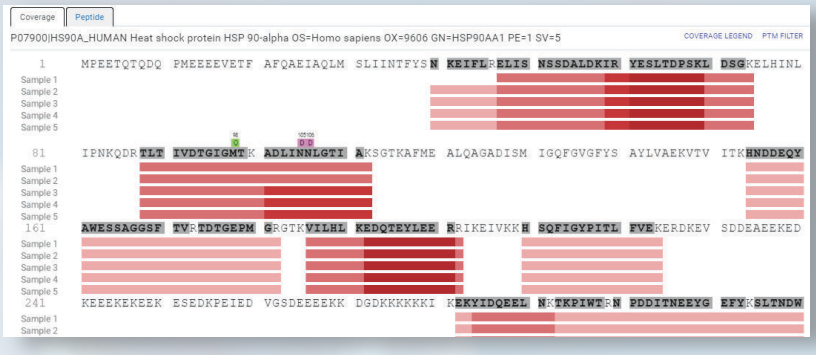


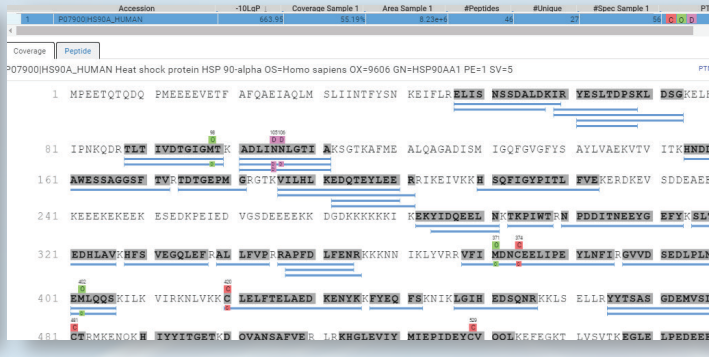
# PEAKS spectral library search uses RT, CCS, & AI-based spectral prediction to improve the separation of true and false hits

Use PEAKS Online for high-throughput analysis of large sample cohort proteomics studies to overcome the increased search space and long processing times. With the addition of the spectral library workflow researchers can focus on the purpose of the study by gathering targeted information and screen samples for known peptides and proteins.



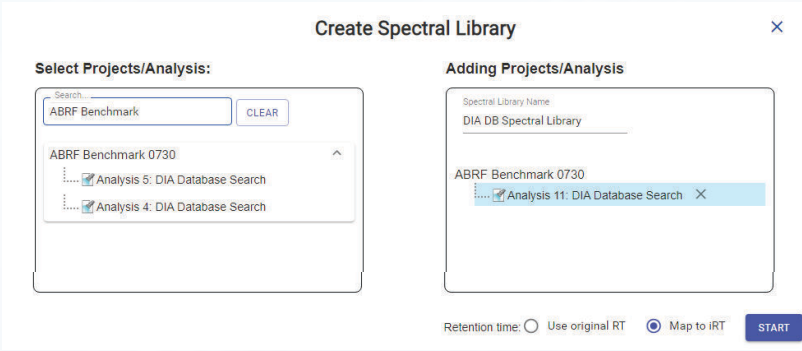

Addition of a sequence database with your spectral library allows protein inference. Easily determine the source of your peptide using the protein coverage view and compaare how they vary from one sample to the next.



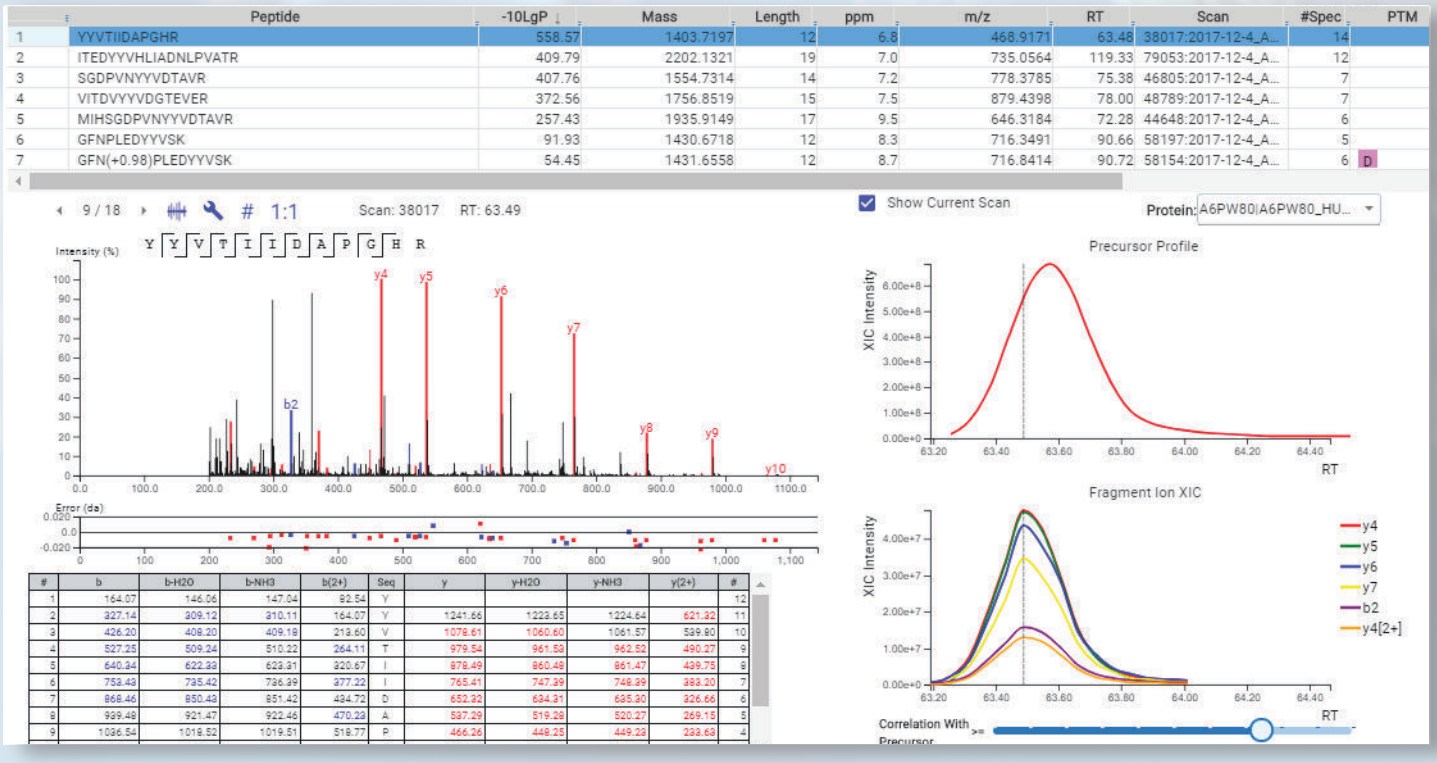



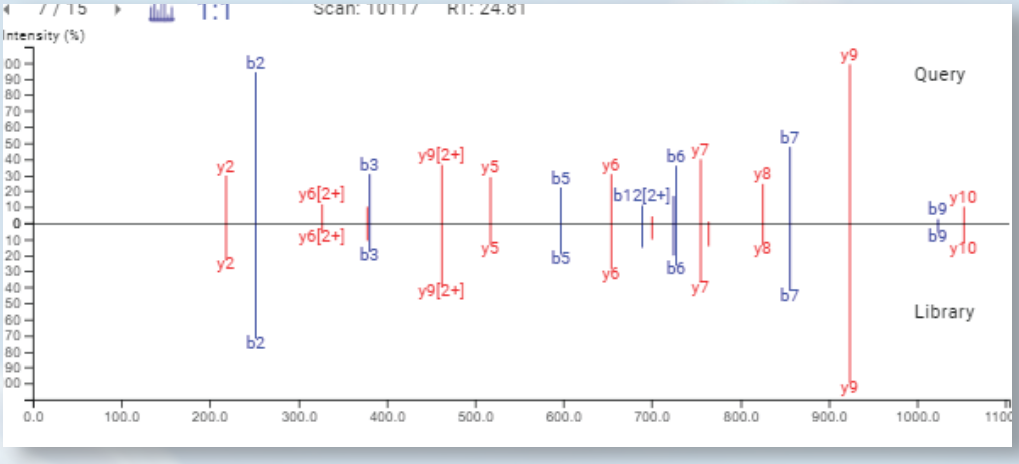

## Creating a library is as simple as one click

Easily generate a spectral library within PEAKS Online by simply selecting the desired PEAKS identification results from your list of completed analyses. Alternatively, PEAKS Online supports library generation from external sources.



Increased complexity of DIA data is a result of its information-rich acquisition method. DIA data provides highly accurate information on MS1 and MS2 levels with few missing values. Precursor Profile and Fragment Ion XIC views have been added to help with quality control and validation.

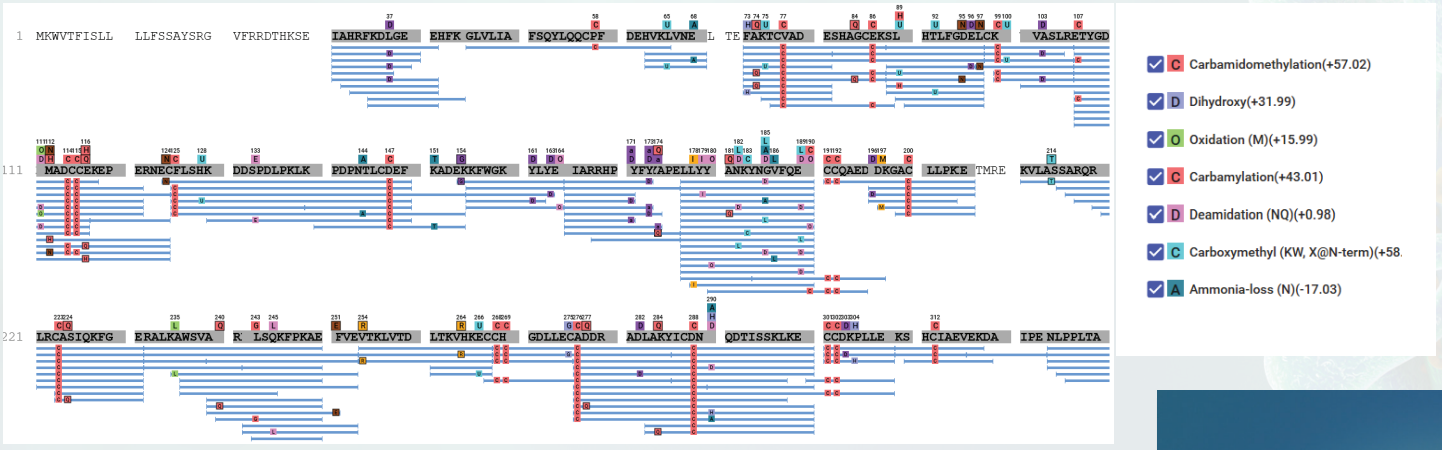



Click the icon to view the Query vs. Library Mirror Plot. Using a spectral library, you can match the observed spectra to the identified spectra.

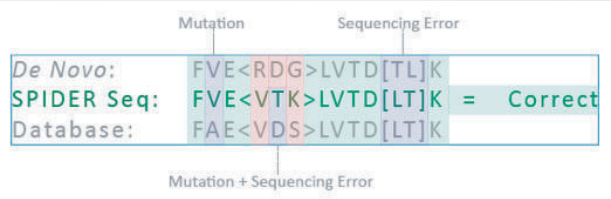
## Designed to discover hidden modifications

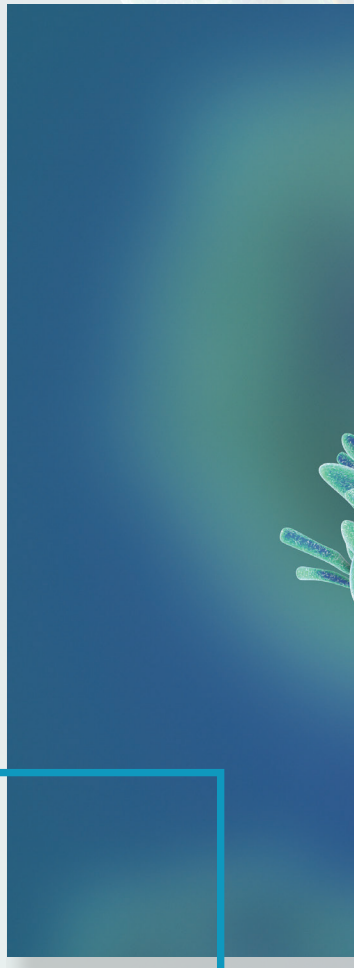
In PEAKS PTM, the unassigned spectra with highly confident de novo scores are reanalyzed to assess any unknown or unexpected modifications. Use a set list of your PTMs of interest, or turn on all 313 naturally occurring, biological modifications from the Unimod database. This multiple-round search approach can help you maximize the identification and sensitivity of your PTM analysis.




## Powered by de novo sequencing to reconstruct the true sequence

SPIDER tries to match the de novo sequence tags with the database proteins. By minimizing the sum of the de novo errors between the reference sequence and the *de novo* sequence, SPIDER, reconstructs a "real" sequence to find peptides with single amino acid variants.



**The characterization of PTMs and sequence variants is crucial to the understanding of biological pathways.**

## Site localization confidence and result validation

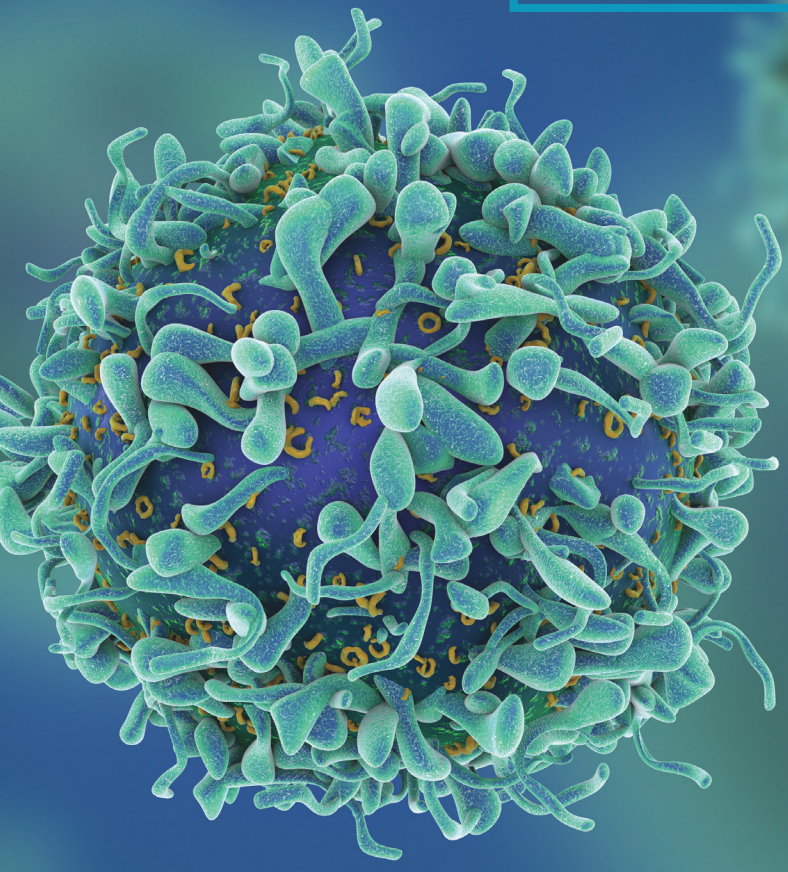
Use PEAKS to measure the probability of any given local site modifications by assessing it's A-Score and/or ion intensity. Allow PEAKS to help you confidently report identified PTMs and sequence variants.

## Cross-species homology search with SPIDER

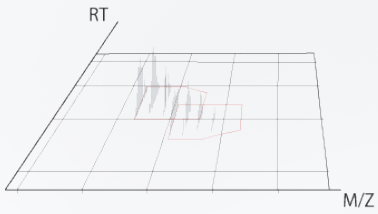
*De novo* tag homology search tolerates common de novo sequencing errors such as (AT/TA) and (N/GG). Find confident hits that are different from the database entry with our de novo tailored homology search.

**SPIDER provides a specialized approach for:**

- Resolving database errors
- Antibody sequence confirmation
- Potential biomarker discovery
- Mutated peptide identification



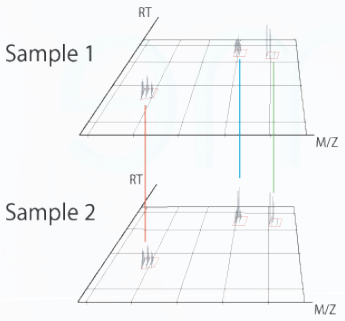
# Add the PEAKS Q module to your PEAKS Online data analysis workflow for robust label & label-free quantification




Ion Detection

## Accurate and sensitive protein- and peptide-level quantification in all dimensions

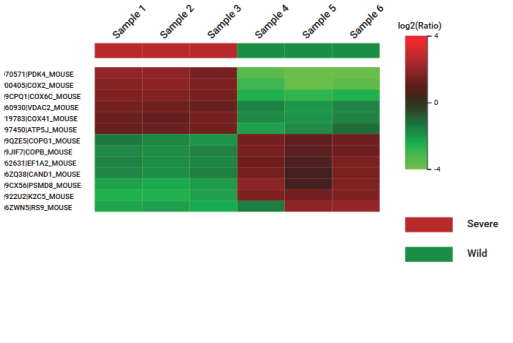
To achieve accurate and sensitive quantification for the large amount of data collected from LFQ experiments, use PEAKS Q to automate ion peak alignment and comparison. Users with access to IMS-MS technology for the extra 4th- ion mobility dimension, enable the PEAKS IMS module to further enhance feature alignment for more accurate quantification results.

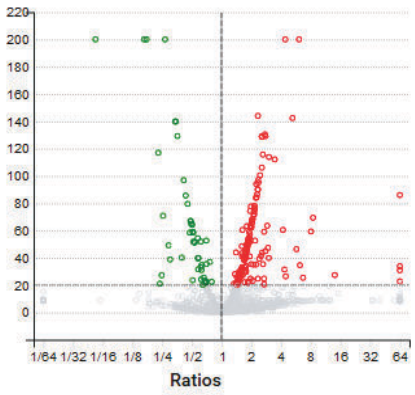


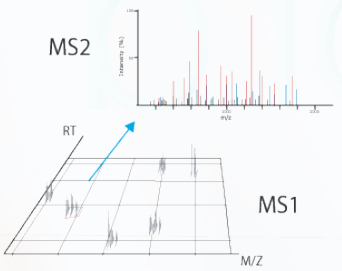

Alignment

## Easily interpret protein and peptide change in abundance between samples and groups

PEAKS Q presents its quantitative results in various graphs and charts to quickly analyze the change in proteins and peptides with just a glance at the results. Export the results in text format for precise, detailed information.




Ion Identity

| | Accession | Significance | Coverage | #Peptides | #Unique | PTM | Sample Profile | Group Profile | Avg. Mass | Description |
|---|---|---|---|---|---|---|---|---|---|---|
| 16 | P19536\|COX5B_MOUSE | 116.96 | 28.13% | 4 | 4 | C O | | | 13813 | Cytochrome c oxidase subunit 5B, mitochondria... |
| 17 | P19324\|SERPH_MOUSE | 115.63 | 40.05% | 13 | 13 | O | | | 46534 | Serpin H1 OS=Mus musculus GN=Serpinh1 PE=... |
| 18 | P17427\|AP2A2_MOUSE | 113.82 | 9.38% | 9 | 5 | | | | 104017 | AP-2 complex subunit alpha-2 OS=Mus muscul... |
| 19 | Q8CIE6\|COPA_MOUSE | 112.05 | 3.35% | 3 | 3 | C | | | 138432 | Coatomer subunit alpha OS=Mus musculus GN... |
| 20 | P16015\|CAH3_MOUSE | 106.03 | 81.92% | 20 | 20 | C | | | 29366 | Carbonic anhydrase 3 OS=Mus musculus GN=C... |
| 21 | Q9JI91\|ACTN2_MOUSE | 100.46 | 47.87% | 59 | 35 | C D O | | | 103834 | Alpha-actinin-2 OS=Mus musculus GN=Actn2 P... |
| 22 | P27659\|RL3_MOUSE | 97.12 | 15.63% | 5 | 5 | O | | | 46110 | 60S ribosomal protein L3 OS=Mus musculus GN... |

P16015|CAH3_MOUSE Carbonic anhydrase 3 OS=Mus musculus GN=Ca3 PE=1 SV=3

PTM FILTER TOOLS

1 MAKEWGYASH NGPDHWHELY PIAKGDNQSP IELHTKDIKH DPSLQPWSAS YDPGSAKTIL NNGKTCRVVF DDTYDRSMLR GGPLSGPYRL RQFHLHWGSS DDHGSEHTVD

111 GVKYAAELHL VHWNPKYNTF GEALKQPDGI AVVGIFLKIG REKGEFQILL DALDKIKTKG KEAPFTHFDP SCLFPACRDY WTYHGSFTTP PCEECIVWLL LKEPMTVSSD

221 QMAKLRSLFS SAENEPPVPL VGNWRPPQPV KGRVVRASFK

| # | Peptide | Used | Quality | Significance | Avg.ppm | Avg.Area | Sample Profile | Group Profile | Area Severe | Area Wild | Max Ratio | #Vector | Start | End | PTM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | HGSFTTPPC(+57.02)EEC(+57.02)IVW... | | 0.39 | 60.00 | 0.0 | 1.68e+4 | | | 0.00e+0 | 1.01e+5 | 64.00 | 1 | 194 | 212 | C C |
| 2 | VVFDDTYDR | | 3.81 | 60.00 | 0.0 | 1.85e+5 | | | 1.11e+6 | 0.00e+0 | 64.00 | 1 | 68 | 76 | |
| 3 | APFTHFDPSC(+57.02)LFPAC(+57.02)R | | 3.67 | 3.26 | 2.1 | 2.47e+5 | | | 1.83e+4 | 4.82e+5 | 26.40 | 1 | 173 | 188 | C C |
| 4 | DYWTYHGSFTTPPC(+57.02)EEC(+57.... | | 4.09 | 2.66 | 2.7 | 9.06e+4 | | | 1.11e+4 | 1.78e+5 | 16.07 | 1 | 189 | 212 | C C |
| 5 | QFHLHWGSSDDHGSEHTVDGVK | | 6.97 | 2.28 | 1.8 | 1.76e+5 | | | 3.17e+4 | 3.41e+5 | 10.76 | 1 | 92 | 113 | |
| 6 | EWGYASHNGPDHWHELYPIAK | | 8.56 | 1.80 | 3.0 | 4.47e+5 | | | 1.12e+5 | 7.81e+5 | 6.95 | 1 | 4 | 24 | |
| 7 | EKGEFQILLDALDK | | 8.97 | 1.74 | 1.1 | 1.05e+6 | | | 2.78e+5 | 1.82e+6 | 6.57 | 2 | 152 | 165 | |
| 8 | QPDGIAVVGIFLK | | 7.10 | 1.64 | 2.3 | 1.98e+5 | | | 5.66e+4 | 3.39e+5 | 5.99 | 1 | 136 | 148 | |
| 9 | GEFQILLDALDK | | 22.56 | 1.59 | 1.0 | 1.09e+6 | | | 3.75e+5 | 1.81e+6 | 4.82 | 1 | 154 | 165 | |
| 10 | YAAELHLVHWNPK | ✓ | 20.77 | 1.49 | 0.9 | 7.16e+6 | | | 2.65e+6 | 1.17e+7 | 4.41 | 3 | 114 | 126 | |

| Group | Area | Ratio |
|---|---|---|
| Severe | 2.65e+6 | 1.00 |
| Wild | 1.17e+7 | 4.41 |


## Priority processing for multi-user platform

Ensure deadlines are met by assigning priority to the analyses. Default priority can be set per user, and priority can be adjusted once an analysis has begun. Users can also receive an email notification when analysis is complete.

## Easily setup real-time processing of LC-MS data

Setup parsing rules for a daemon instead of adding individual data files into a project to allow real-time processing of data as it comes off the instrument.

## Work as a team and collaborate with ease

In PEAKS Online, users can easily share projects, databases, and standardized workflows between the whole research team.

## Easily monitor performance of PEAKS Online

Administrator(s) of PEAKS Online can check the usage of the master and worker nodes to make sure everything is running at optimal performance.



Bioinformatics Solutions Inc.
470 Weber Street North, Suite 204
Waterloo, Ontario N2L 6J2
Canada
Tel: (519) 885-8288
Fax: (519) 885-9075
sales@bioinfor.com
www.bioinfor.com