BIOINFORMATICS SOLUTIONS INC.

PEAKS® Online 12

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

www.bioinfor.com

PEAKS® Online 12

ASCEND TO NEW HEIGHTS

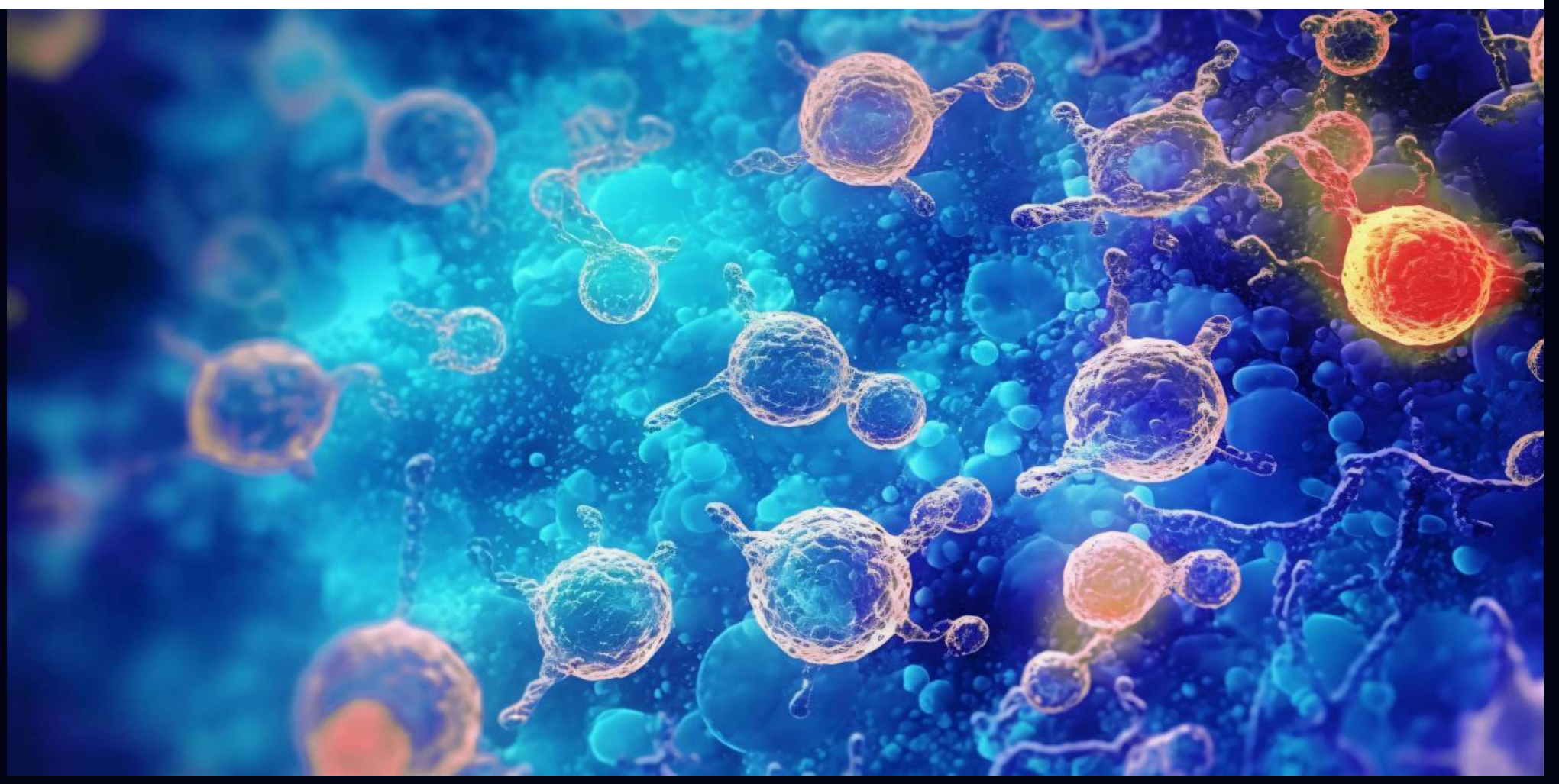
DEEP PROTEOMICS SOLUTIONS

Mass spectrometers and other related analytical techniques have continued to improve at a rapid pace over the past few decades. It is important for researchers to have software that is up to date and can handle the continuously improving data outputs and provide a high-throughput data analysis solution.

Developing innovative software and applying AI-driven technology to proteomics data analysis is vital to advancing research by providing faster, more accurate and sensitive identification and quantification. Together the latest mass spectromery technology and PEAKS® Online will advance the frontier of biological research and facilitate drug discovery.

NEXT GENERATION OF PEAKS®

PEAKS® Online 12 is the next generation of the online platform with optimized workflows to streamline your data analysis. As high-throughput, multi-user solution PEAKS® Online is built for collaboration and handling large scale proteomic studies. From DDA to DIA data support, PEAKS® Online 12 provides a complete solution to bring your research to new heights!



PEAKS® Online is designed to take advantage of powerful and shared computing resources to perform LC-MS/MS peptide and protein 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.

OVERVIEW

03
Deep learning-enabled software
PEAKS® Studio 12 harnesses AI to improve accuracy and sensitivity

04
High performance, multi-user solution

05
Streamlined workflows
PEAKS® Studio 12 provides a complete workflow for DDA and DIA

07
PEAKS® DIA Workflow
DIA pipeline that integrates spectral library search, DB search, and *de novo*

09
PEAKS® *de novo sequencing*
PEAKS® Online 12 harnesses AI to enable higher accuracy with false discovery rate (FDR) estimation

10
PEAKS® *DeepNovo Peptidome*
Identify and validate MHC peptides from coding variants and non-coding regions

11
PEAKS® DB
de novo-assisted databse search integrated with AI for peptide identification

12
Unleash quantitative proteomics
PEAKS® Q add-on supports label free, and labelled quantification

14
Intuitive GUI & automated Quality Control (QC) tool
Visualize and ystematically validate experiment, data quaiity with data analysis results

Extending the

DISCOVERY PROTEOMICS DEEP LEARNING REVOLUTION

PEAKS® Online 12 harnesses deep learning technology to improve identification, accuracy, and sensitivity by using algorithms to predict retention time, fragment ion intensity, mass-to-charge ratios, and ion mobility.

Deep learning is utilized in DDA workflows for PEAKS® *DeepNovo* peptidome for immunopeptidomics and deep learning-boost in PEAKS® DB search results. While in DIA workflows, PEAKS® database search examines the dataset with an *in silico* spectral library generated from the protein sequence database. *In silico* peptide details including the fragment ion pattern, indexed retention time, and ion mobility are predicted using deep learning. New to PEAKS® 12, users can also perform *DeepNovo*-DIA Peptidome for DIA immunopeptidomics studies.

FACILITATE DRUG DISCOVERY AND ADVANCE THE FRONTIER OF BIOLOGICAL RESEARCH THROUGH AI-DRIVEN SOFTWARE SOLUTIONS

In PEAKS® Online 12, we have integrated AI to advance the following workflows:

Peptide *de novo* sequencing: In this release, users will experience the the first *de novo* sequencing with FDR estimation and the most accurate and fastest peptide *de novo* sequencing available in any PEAKS® product. PEAKS® has been known as the gold standard for automated peptide *de novo* sequencing for many years but with the new GPU-enabled "*DeepNovo*", *de novo* sequencing in PEAKS® is the best it has ever been.

DIA Workflow: Deep learning has been integrated into our DIA workflow for spectral library search, database search, and *de novo* sequencing. In this version, users can expect PEAKS® DIA workflow to be faster, increased sensitivity and accuracy, and better support for short gradients. New to PEAKS® 12, the software is integrated with parallel-reaction monitoring (PRM) for discover and validation.

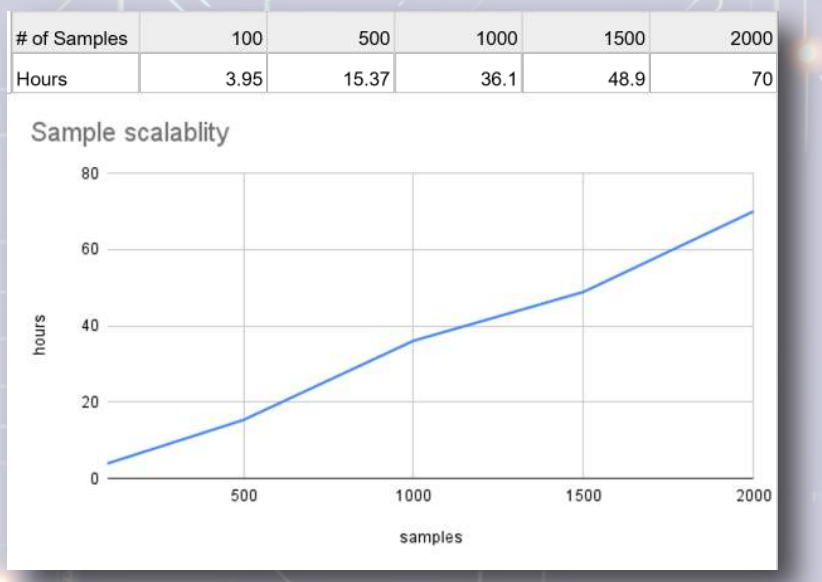
***DeepNovo* Peptidome:** This workflow supports both DDA and DIA data analysis and is a specialized workflow for immunopeptidomics. In this version, we have integrated the latest *DeepNovo* algorithm to improve processing speed and identify peptides with a variety of modifications, sequence variants, and/or splice sites. New to PEAKS® 12, users can link the peptides to genes and bridge the gap between proteomics and genomics to help identify and validate peptides potentially originating from rare sequence variants and non-coding regions.

PEAKS® DB: PEAKS® provides a unique *de novo*-assisted database search called PEAKS® DB to improve both sensitivity and accuracy. The users have the option to include the deep learning boost for MS2 rescoring and improving the number of identifications by ~10%.

ACCELERATE YOUR DISCOVERY PROTEOMICS

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.

# of Samples	100	500	1000	1500	2000
Hours	3.95	15.37	36.1	48.9	70



Align your team's efforts with administrative controls to standardise workflows, databases, PTMs, quantification methods, project sharing and Quality Control (QC) analysis. PEAKS® Online 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.

DID YOU KNOW:

Both new DIA workflow and PEAKS® *DeepNovo* Peptidome enable GPU acceleration

ADVANCED SYSTEM ARCHITECTURE

Built on top of the 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.

Scalable: Vertically and horizontally, add new worker or database node(s) .

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

Dual interfaces: The CLI offers the abilty to automate data analysis workflows and result exporting, the web interface provides a GUI to visually configure workflows and easily assess results in detail.

True automation with PEAKS® daemon: Automated Instrument Link seamlessly connects your acquisition to your analysis in one easy to use workflow.

Streamline your LC-MS/MS data analysis with

NEW WORKFLOWS

Both DDA and DIA technologies are rapidly advancing, and researchers need an analysis method that harmonizes the benefits of both acquisition methods. In recent years, DIA has become increasingly popular due to its parallel nature of acquiring all fragment ions for all precursors within a selected m/z range. This overcomes the limitations of sequential MS/MS acquisition in DDA.

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

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.

DIRECT DATABASE SEARCH:
Given a spectrum and protein sequence database, construct an *in-silico* spectral library using spectrum and iRT prediction models to find a peptide in the spectral library that has the best match with the spectrum.

As a vendor neutral proteomics software developer, we strive to provide a comprehensive solution to facilitate proteomics research and support efficiency in mass spectrometry labs. Avoid the need to use different software for various analytical and acquisition methods.

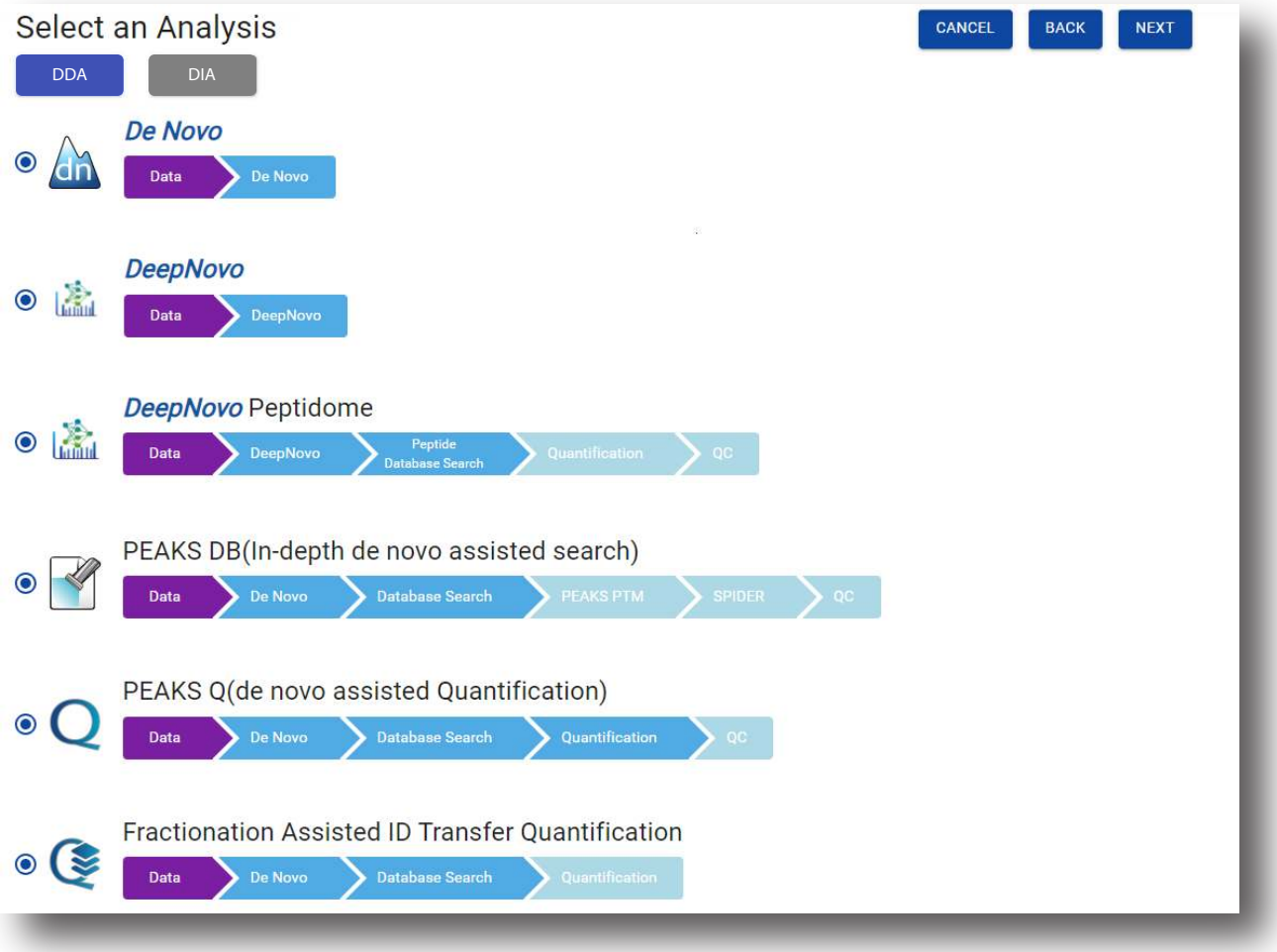
1. When the peptides are believed to be in a protein sequence database, then a database search approach is preferred.
2. When studying a particular proteome, a peptide spectral library for the targeted biological system being studied can be used to focus your analysis.
3. However, when such a sequence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

PEAKS® provides a complete solution for all LC-MS/MS search methods to enhance laboratory efficiency, and reproducibility, and reporting. Seamlessly progress from raw file to report for all your bottm-up mass spectrometry based proteomics.

DDA WORKFLOWS

From *de novo* sequencing and database search to quantification, take advantage of the proven PEAKS® search algorithms efficiently as established workflows.

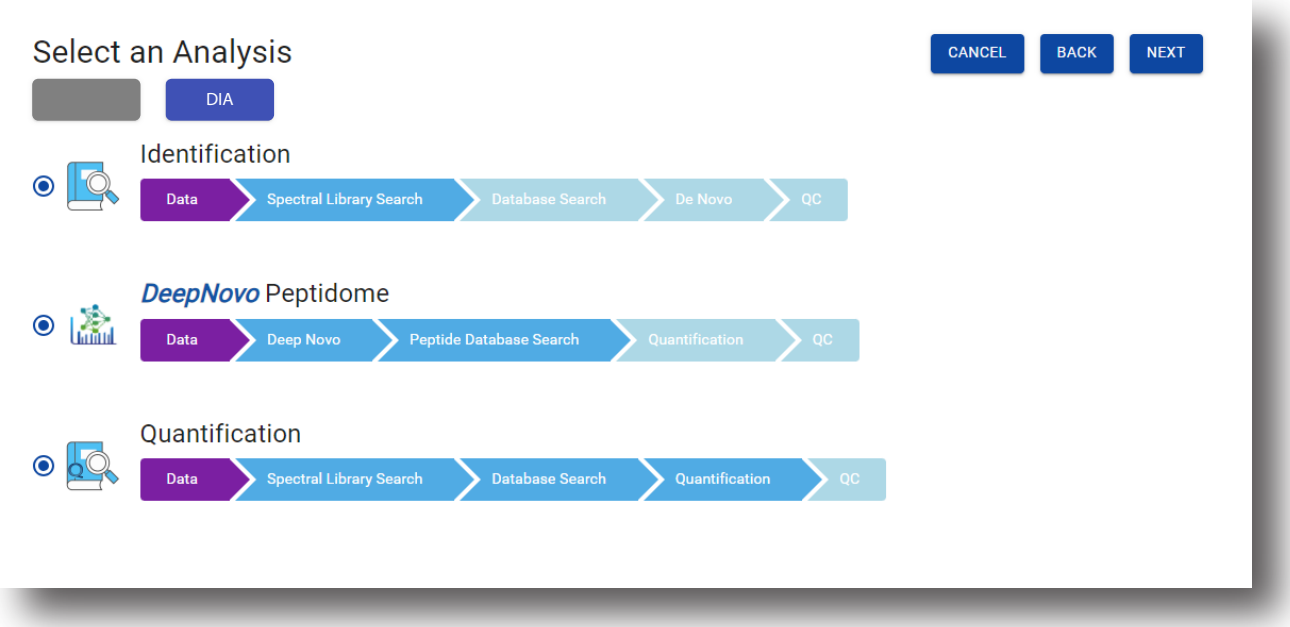
New to PEAKS® Online 12:
The PEAKS® 12 DeepNovo workflow now uses the new GraphNovo algorithm. For the first time, FDR estimation is available for *de novo* sequencing results.



DIA WORKFLOWS

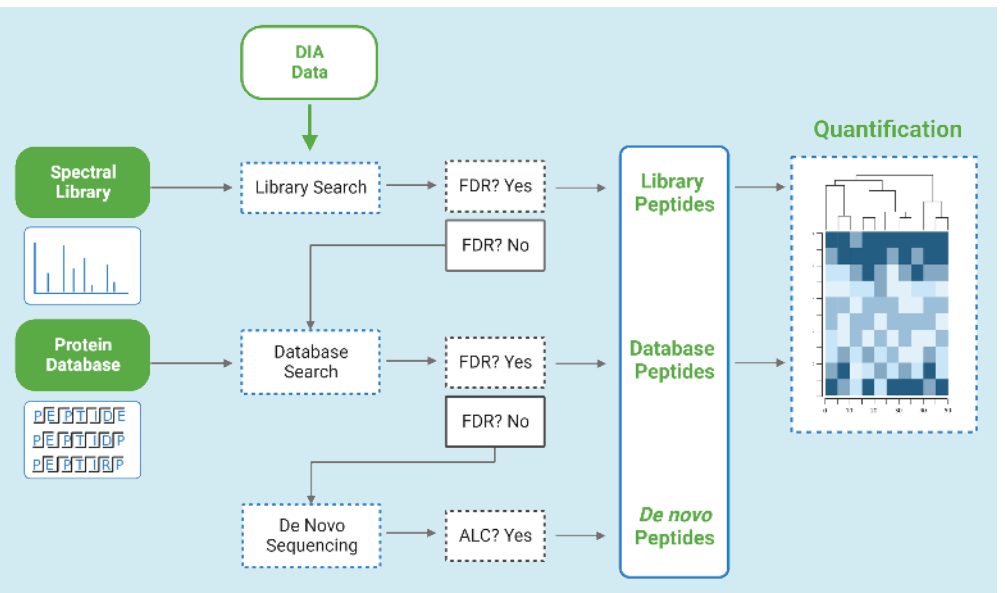
PEAKS® Online features an advanced DIA workflow that incorporates three methods of peptide identification: spectral library search, direct database search, and *de novo* sequencing. For quantification, the label-free method is available as an add-on module.

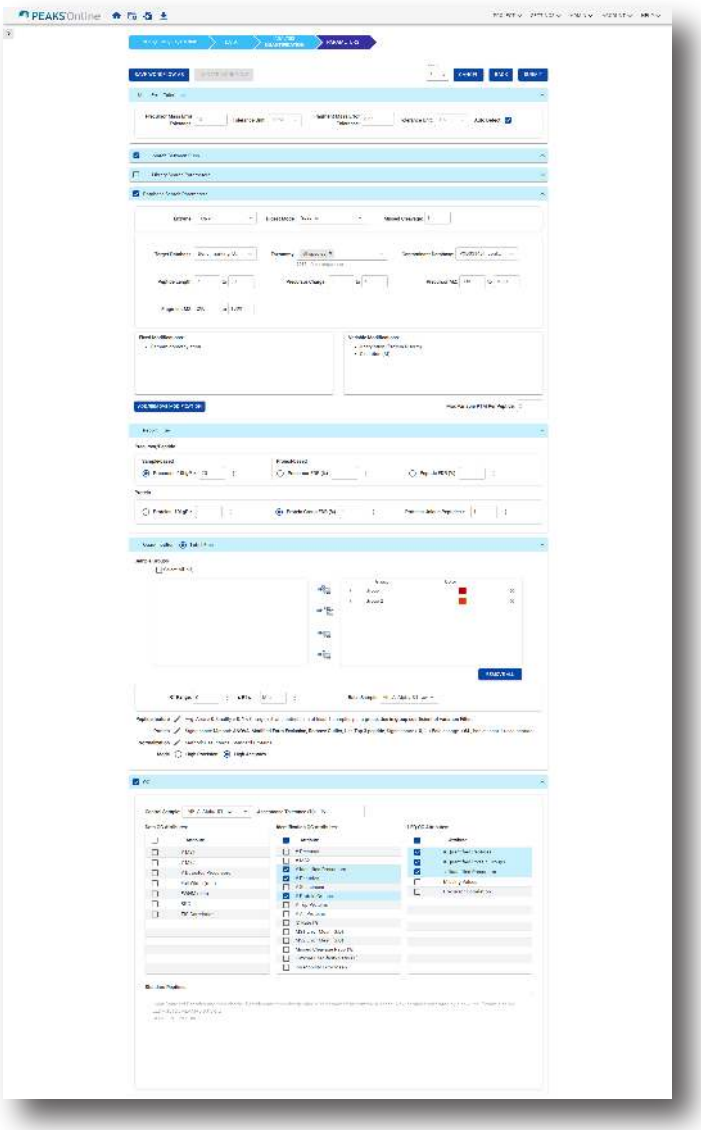
New to PEAKS® Online 12:
Match Between Runs is now available in the identification workflow.



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

PEAKS® offers a robust solution for DIA data analysis. It incorporates three methods of peptide identification: spectral library search, direct database search, and *de novo* sequencing. The search is performed using an expanding search space. First, a library search is performed against a library of previously identified spectra. By estimating the false discovery rate, peptides that pass the filter are saved. MS/MS spectra that don't match a peptide within the false discovery rate threshold are brought forward to a direct database search. Confident database matches are added to the result. Then, using the same FDR approach, unmatched spectra from the database search are analysed using de novo sequencing.





Advantages of a DIA Workflow

- Decrease bias by including all peptides in analysis
- Reproducibility of peptide detection and quantification across MS runs
- Quantify proteins in complex mixtures over a dynamic range
- Eliminate under sampling
- Increased sensitivity and depth of proteome coverage
- Increased precision and reproducibility when compared to DDA
- Eliminate the cost and time with label free quantification

Deep learning advances accuracy and sensitivity of data analysis

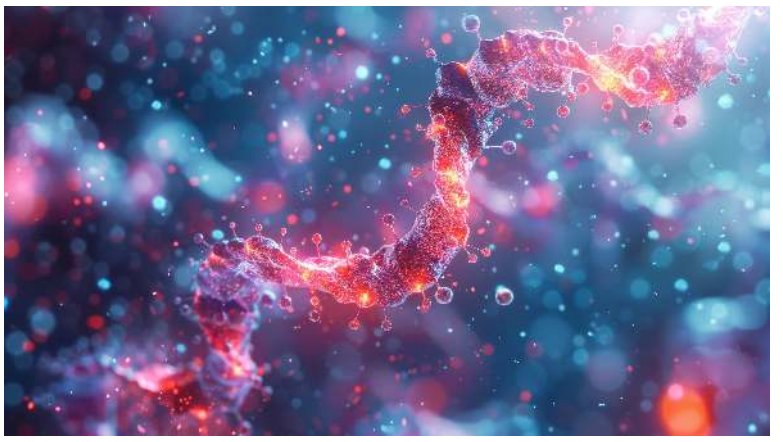
PEAKS® Online uses advancements in deep learning-based spectrum prediction models to perform both DDA and DIA database search by predicting the retention time and spectra *in silico* for each plausible peptide. In addition, PEAKS® furthers the use of deep learning to perform *de novo* sequencing for 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.

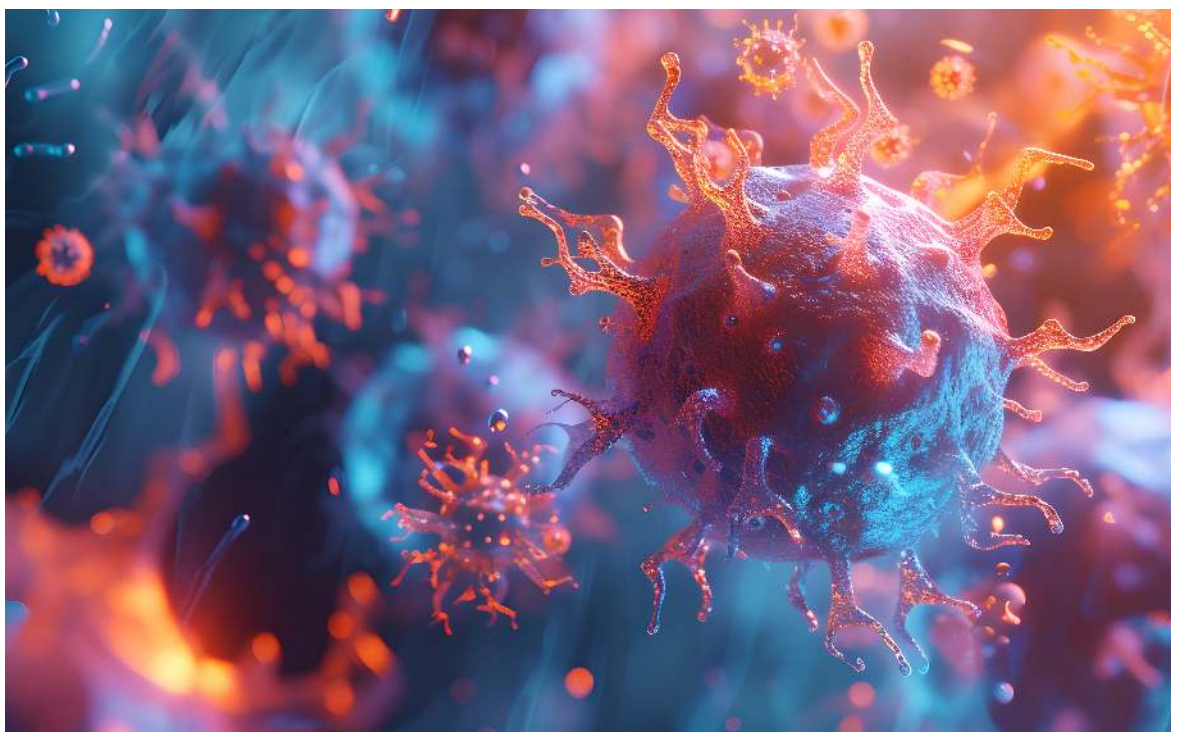
To overcome the boundaries of discovery, PEAKS® introduces the first *de novo* sequencing approach for DIA to provide truly unbiased results

Next generation PEAKS® *DeepNovo*

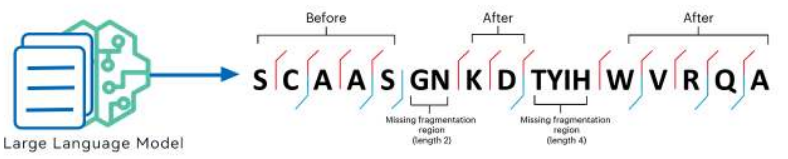
deep learning-enabled *de novo* sequencing with FDR control



De novo sequencing is a fundamental method to discover novel protein and peptide with tandem MS. PEAKS *de novo* has been well known as the gold standard for peptide and protein *de novo* sequencing. The new PEAKS *de novo* engine, *DeepNovo* provides higher accuracy and faster speed, also enables FDR estimation.

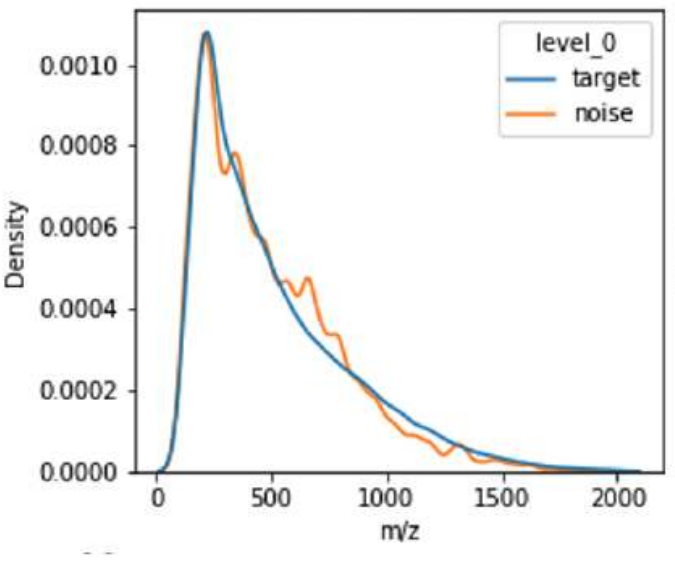


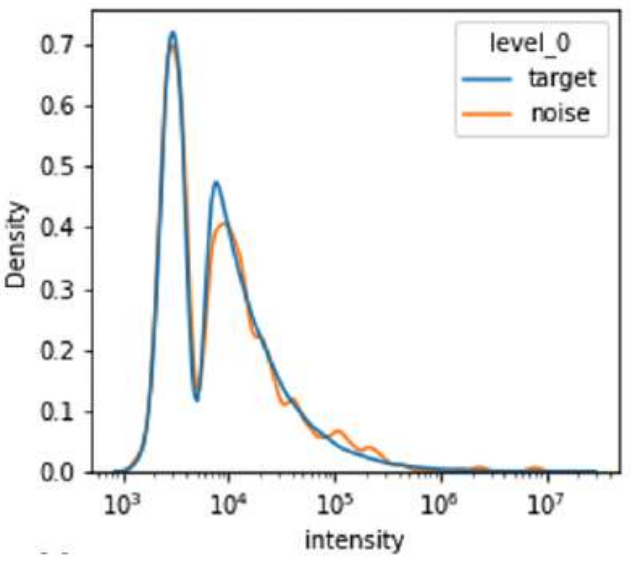
Empowered by the latest *GraphNovo* algorithm, the new *DeepNovo* resolved the issue of accumulated prediction errors due to missing fragment ions by finding the optimal path in the first stage to guide sequence prediction in the second stage.

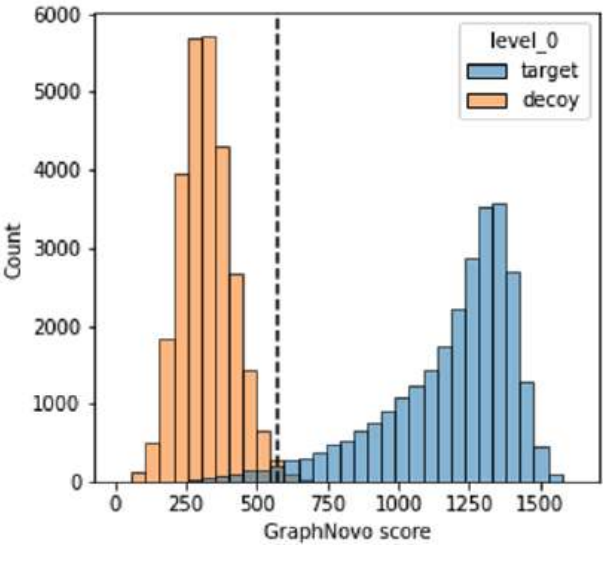


As a key technology for finding novel peptides from MS data, *DeepNovo* provides an advanced solution and will push research like antibody sequencing and neoantigen discovery further.

In PEAKS® 12, decoy spectra are used in *DeepNovo* to estimate an FDR for *de novo* peptide sequencing. The decoy generation process ensures decoy spectra has the same fragment ion distribution as target spectra. Furthermore, *DeepNovo* scoring function efficiently separates target and decoy. An FDR curve of *de novo* sequencing is now available in *DeepNovo*.

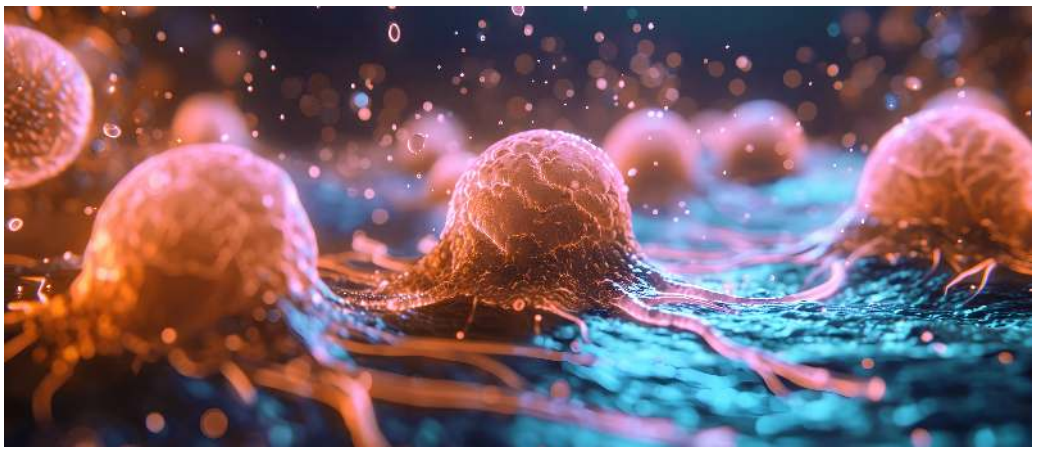




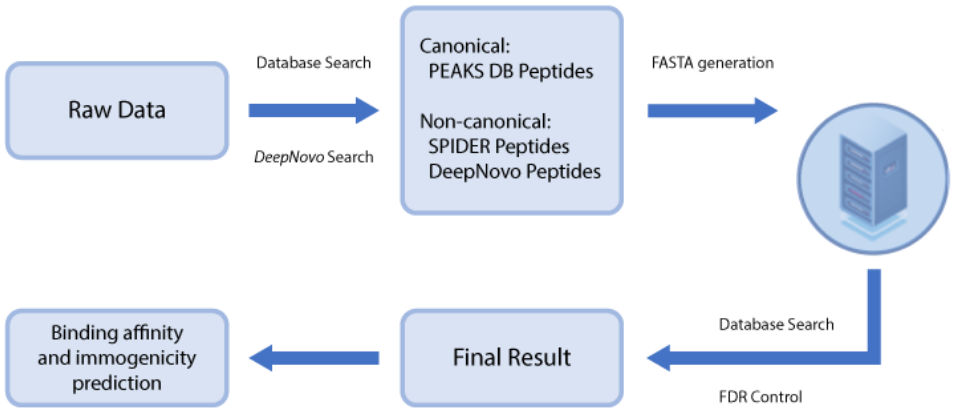


As a deep learning based solution, *DeepNovo* harnessed the computation power of NVDIA GPUs for speed. NVIDIA A100 enabled *de novo* sequencing 1200 + spectrum/s.

PEAKS® *DeepNovo* Peptidome
WORKFLOW

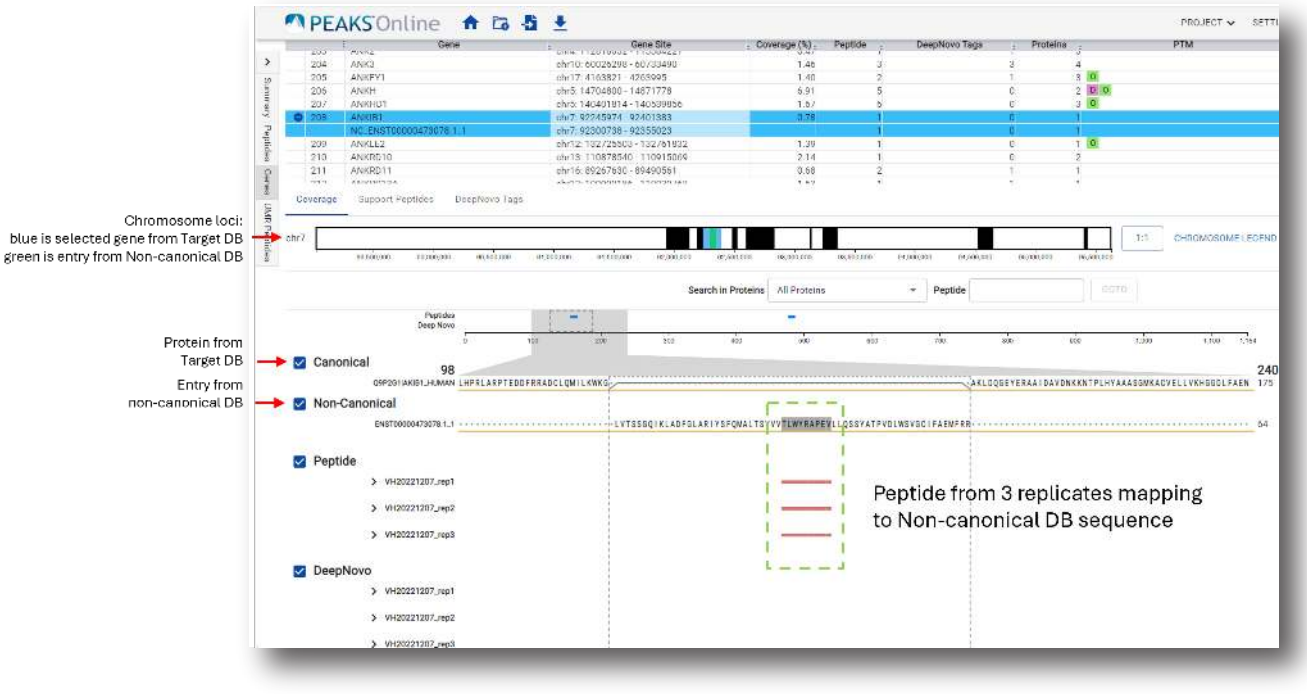


This new solution is a specialized workflow for peptidomics data that combines database searching, *de novo* sequencing, and identification of mutated peptides. By training *DeepNovo* deep learning model using peptidomics datasets the sensitivity and accuracy of peptide identification can significantly be improved. Furthermore, *de novo* peptides (non canonical) are combined with database peptides (canonical) for more accurate estimation of false discovery rate. The final output of peptides are categorized as Database, *DeepNovo* or Homologs (mutated peptides) and can be directly exported for binding affinity and immunogenicity predictions.



Peptide mapping to chromosome loci

To better understand non-canonical peptide biosynthesis and where in the genome each peptide is expressed, a Gene tab is integrated in the *DeepNovo* Peptidome workflow to display this information. The workflow accepts both canonical database files from predicted coding regions (Target database), as well as custom database files from next-generation sequencing (NGS). Target databases such as Uniprot are used to identify canonical peptides mapped to open reading frames of proteins and their isoforms. Custom NGS databases can be used for mapping peptides to non-coding regions and identifying non-canonical peptides. Lastly, the Gene tab also displays *DeepNovo* peptides, where only part of the peptide sequence matches to either canonical or non-canonical databases. This information is helpful in identifying post-translationally spliced peptides.



DID YOU KNOW:

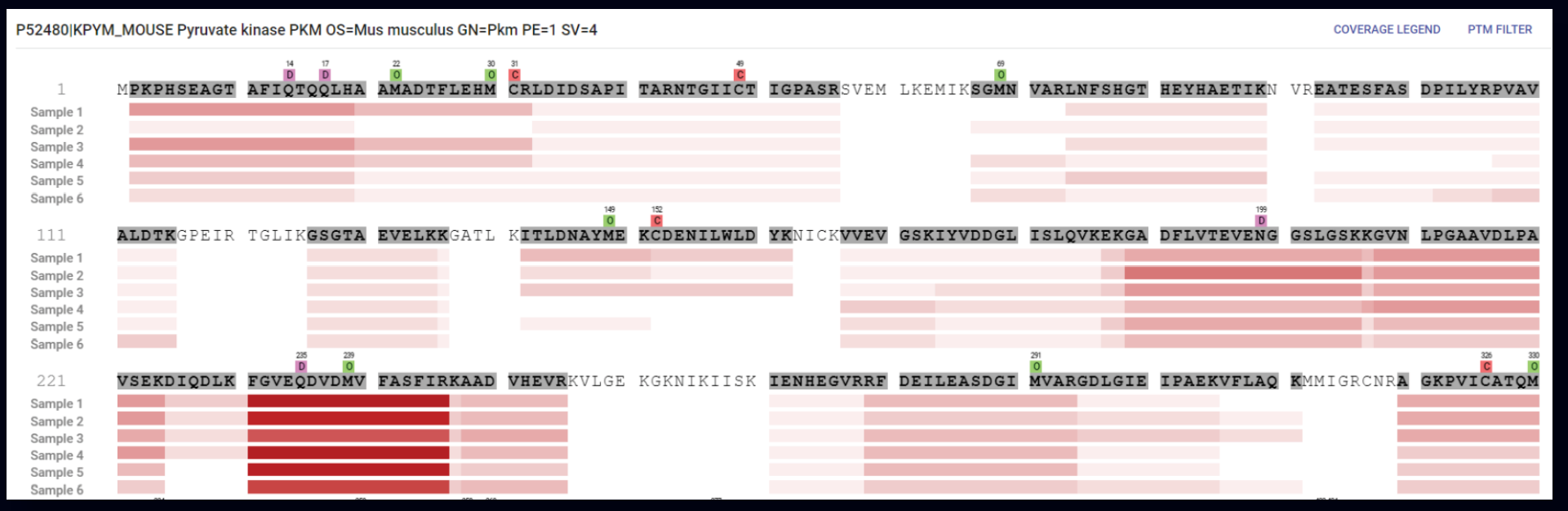
As a deep learning based solution, *DeepNovo* harnesses the computation power of NVDIA GPUs for speed. NVIDIA A100 enabled *de novo* sequencing

1200 + spectrum/second

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.

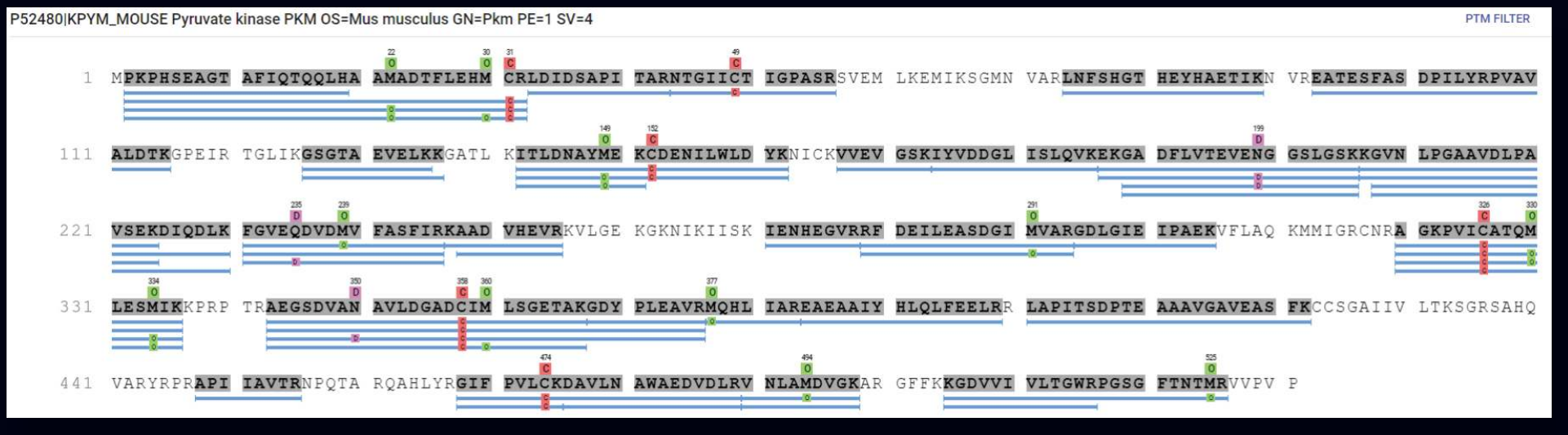
Protein coverage heatmap for quick and easy comparison across multiple samples

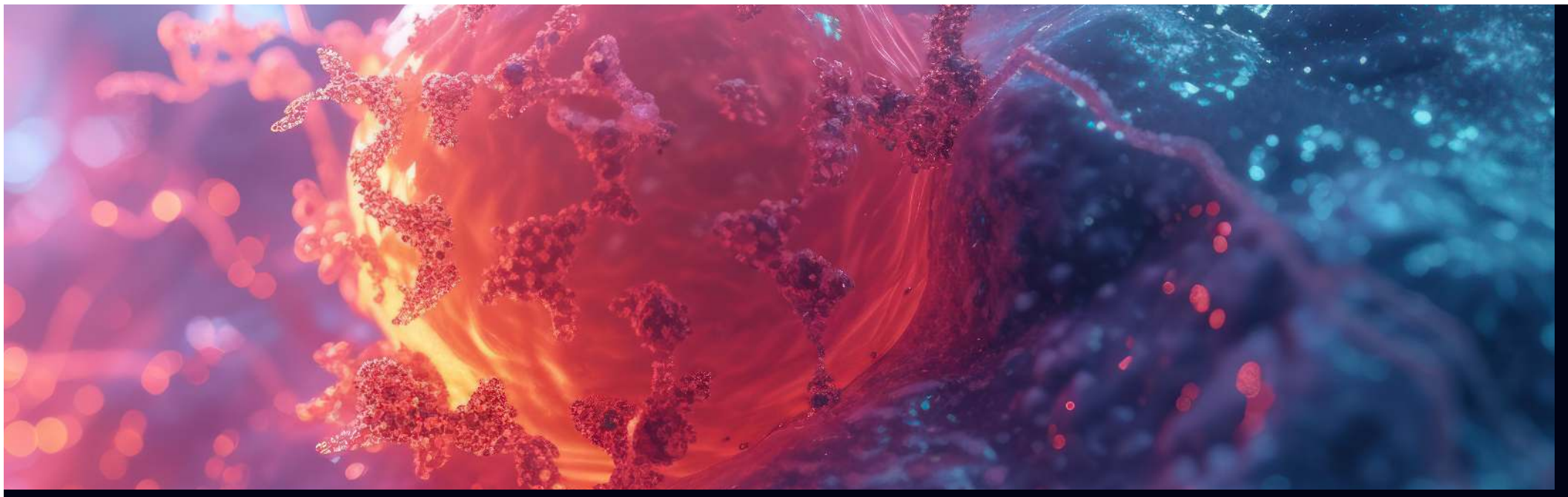
Easily compare between multiple samples in a project using the PEAKS® Online protein coverage heatmap. The increasing colour intensity indicates a higher abundance of supporting spectra within the corresponding sample.



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.

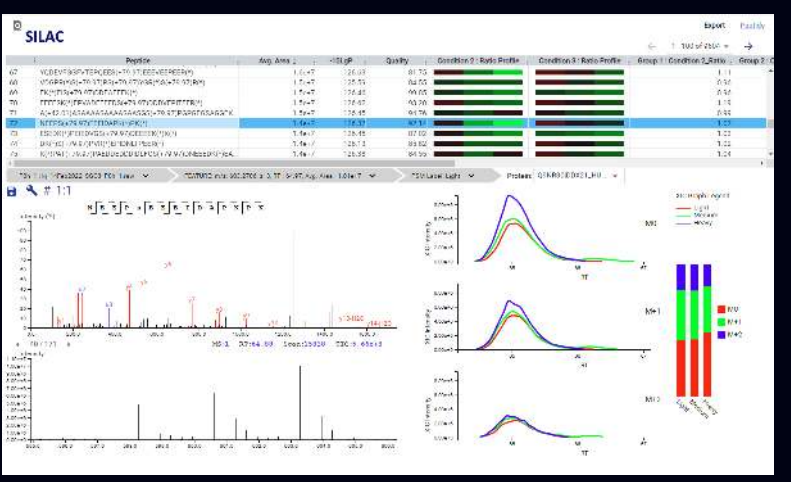




UNLEASHING
QUANTITATIVE PROTEOMICS WITH THE POWER OF PEAKS® Q

Quantification provides greater insight into proteomic mysteries. Researchers need a software tool to support them as they press further in to the understanding of life sciences.

PEAKS® is equipped with not only a powerful identification algorithm, but also embraces paralleled quantification capabilities to perform:



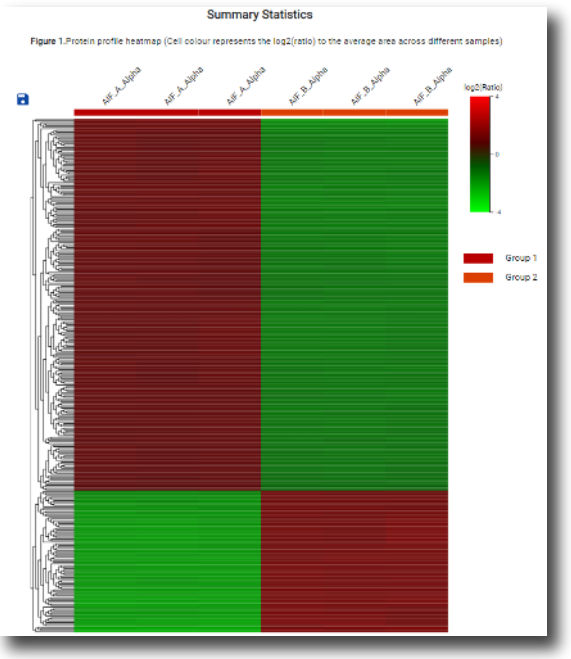
Reporter ion quantification:
Isobaric tags (ex. TMT/iTRAQ) have identical masses and chemical properties that allow heavy and light isotopologues to co-elute. The tags are then cleaved from the peptide by collision-induced dissociation during MS/MS, which is used for quantification. For large-scale, protein quantification studies, researchers can use PEAKS® Q to expand the sample size with reference channels to enhance the accuracy of quantification.

Precursor ion quantification:
Stable Isotope Labelling by Amino Acids in Cell Culture (SILAC) is a powerful and popular approach for mass spectrometry (MS)-based quantitative proteomics. PEAKS® Q's SILAC quantification enables unsurpassed sensitivity of peptide feature detection through a novel peptide feature detection algorithm to find peptide feature pairs. Researchers can take advantage of the intuitive interface showing paired features at first glance and minimize the biases from missing values.

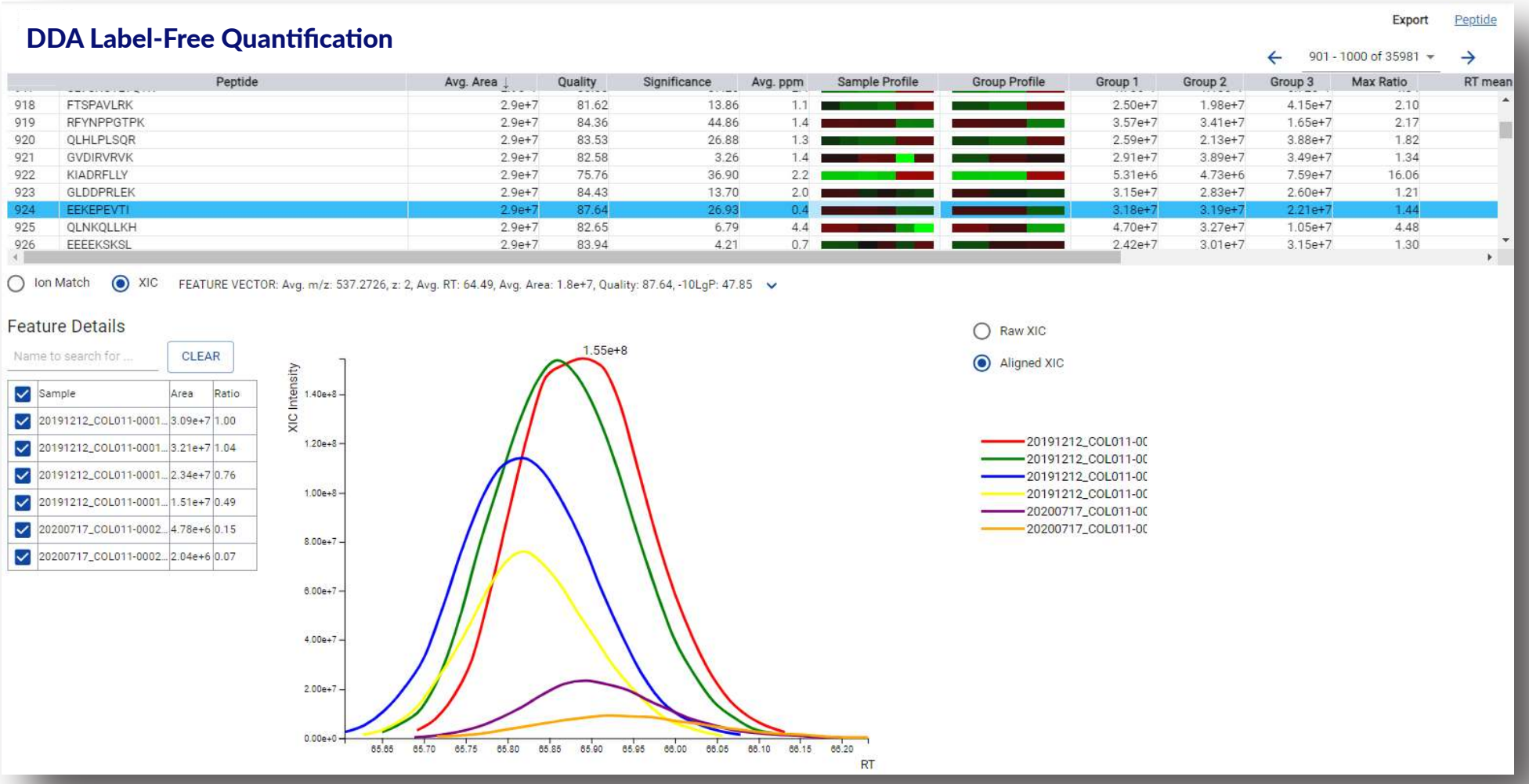
Label-free quantification:

Label-free quantification (LFQ) for discovery proteomics allows for the unbiased and comprehensive analysis of protein abundance across multiple samples without the need for costly and time-consuming labeling steps.

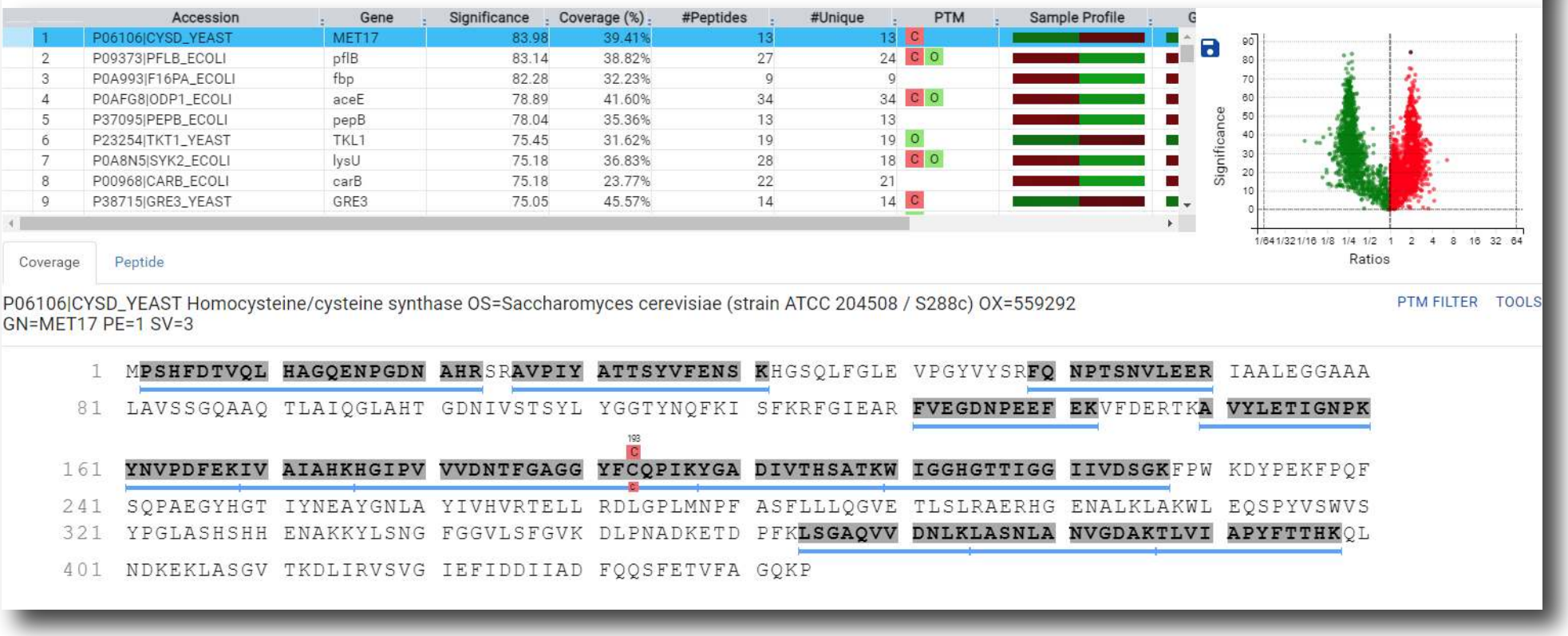
PEAKS® Q enables labs to uncover subtle changes in protein expression and to achieve high sensitivity and specificity in their experiments, thereby advancing our understanding of biological pathways, identify biomarkers for diseases, and understand cellular responses to various stimuli. From the PEAKS® GUI, users can then thoroughly investigate differences in peptide/protein abundance between samples with confident and accurate results.



DDA Label-Free Quantification



DIA Label-Free Quantification

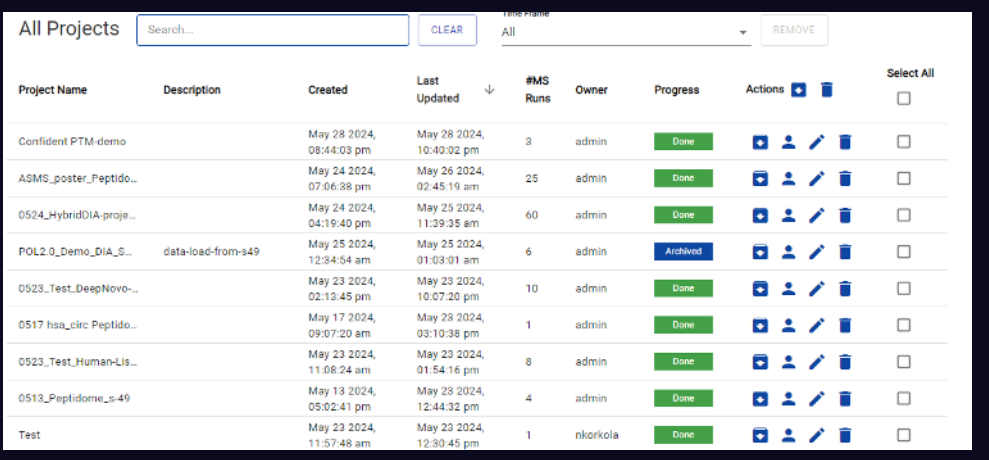


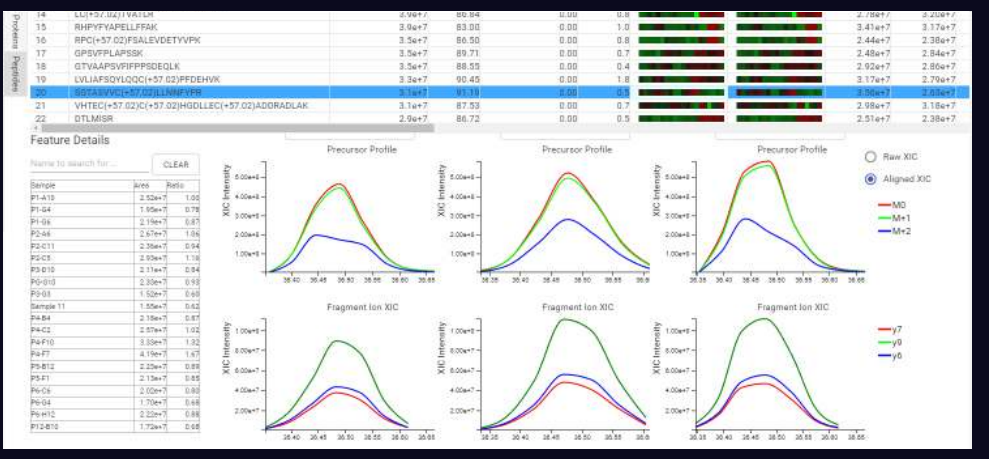
UPDATED SOFTWARE ARCHITECTURE AND GRAPHICAL USER INTERFACE (GUI)

The updated software architecture provides increased speed and stability for identification and quantification. Support all your discovery proteomics research with higher throughput and robust data analysis for both, DDA and DIA support.

Users still get the intuitive data visualisation that PEAKS® is known for, but with a new look and optimised workflows to streamline your data analysis.

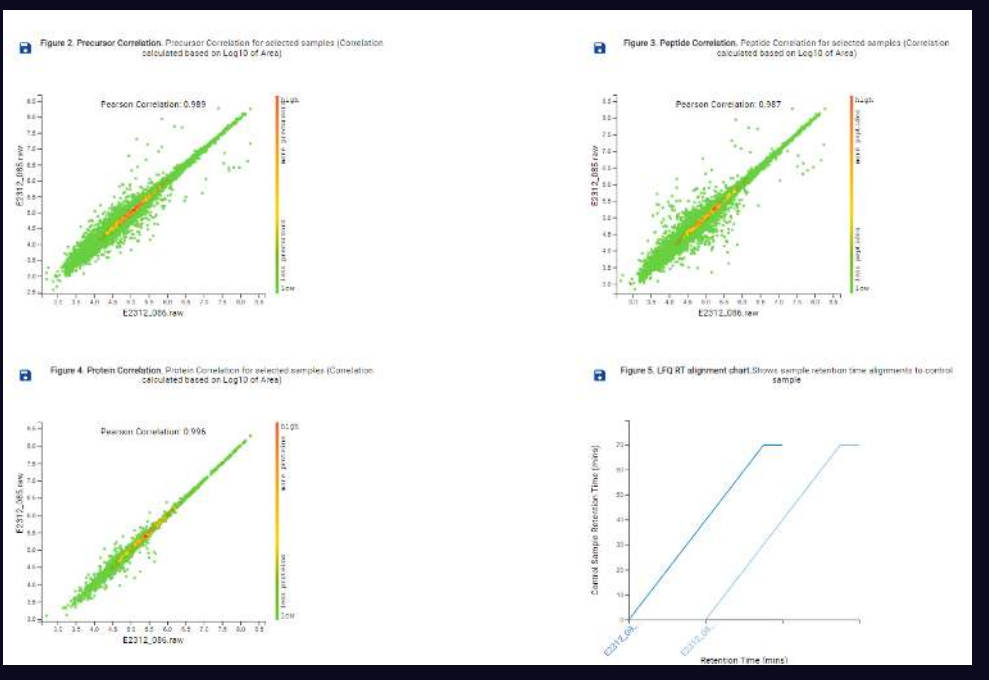
As a vendor neutral software, all DDA and DIA experiments can be performed in PEAKS® . Save your workflows and create customized PTMs, enzymes, labelled quantification methods, etc.

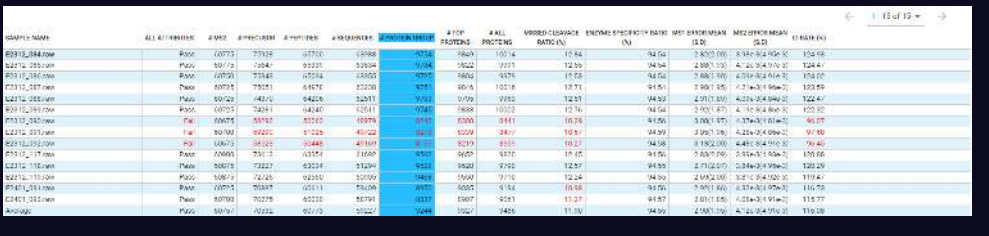




AUTOMATED QUALITY CONTROL (QC) TOOL

PEAKS® Online presents a specialized QC tool designed to improve the accuracy of protein and peptide identification and quantification. This module delivers detailed QC analysis, presenting complex information in an organized and user-friendly manner, which aids in troubleshooting instrumentation and data processing issues. By seamlessly integrating database searching, quantification, and QC analysis into a single workflow, the statistical analysis can be consolidating into a comprehensive report. The entire process is automated, systematic, customizable, and sers can freely export data and figures as needed.





Next generation PEAKS® Online 12

ACCELERATE DISCOVERY PROTEOMICS WITH HIGH THROUGHPUT DATA ANALYSIS

ACCURATE AND SENSITIVE IDENTIFICATION FOR BOTH, DDA AND DIA ANALYSES

ROBUST LABEL-FREE, LABELLED AND PRM QUANTIFICATION SUPPORT

AI TECHNOLOGY INTEGRATED TO ADVANCE *DE NOVO*, DIA-DB, DDA-DB, AND QUANT

AUTOMATED QC TOOL FOR IN-DEPTH ANALYSIS FROM RAW DATA TO RESULTS

Deep learning technology

Deep learning enabled to improve identification accuracy and sensitivity by using algorithms to predict retention time, fragment ion intensity, mass-to-charge ratios, and ion mobility.

High performance Computing for large-scale proteomics

PEAKS® Online uses the latest distributed computing technology to achieve high-throughput performance with established PEAKS® workflows for multiple users on a network. Accelerate your discovery proteomics research.

Assess essential attributes of the raw data & results

The new automated Quality Control (QC) analysis tool is designed for both, DDA and DIA data and will supply the elements to determine the quality of the data and evaluates the setup of the experiment.

Bioinformatics Solutions, Inc.
140 Columbia St. W., Suite 202
Waterloo, Ontario N2L 3K8
Canada
Tel: (519) 885-8288
Fax: (519) 885-9075
sales@bioinfor.com
www.bioinfor.com