

Beyond Data:

Bruker, Biognosys, and Mass Dynamics: Pioneering a Collaborative Era in Accelerated LC-MS Proteomics Discoveries



Introduction

The rapid expansion in LC-MS proteomics data production necessitates advanced, standardized workflows to efficiently handle and analyze large-scale experiments. As cutting-edge instrumentations, such as Bruker's LC-MS timsTOF platforms with dia-PASEF®, emerge, they empower scientists to produce proteomics data at unprecedented throughputs. These deep proteomics experiments, capable of covering hundreds of samples, can now be completed in just a few days. This progress highlights the growing need for innovative solutions to store, process, and retrieve vast quantities of data effectively.

To harness the full potential of their LC-MS platform, Bruker offers software that enables data processing at scales previously thought unattainable. It aims to deliver precise protein identification and quantitation through streamlined workflows - both Bruker ProteoScape™, Bruker's comprehensive solution emphasizing CCS-centric data analytics, and Spectronaut®, the market-leading DIA analysis software from Biognosys.



Bruker ProteoScope, equipped with GPU-powered parallel edge-computing capabilities, offers a robust platform for real-time data analysis in dda-PASEF®, dia-PASEF, PTM searches, and *de novo* sequencing. Beyond its primary proteomics analysis tools, it integrates seamlessly with Bruker TwinScape™, ensuring optimal system performance and maximized uptime, while also introducing intelligent acquisition through instrument feedback loops.

Complementing this, Spectronaut introduces a refined solution for library-free DIA analysis with its directDIA+™ feature. This algorithm enhances the spectrum-centric search in directDIA™ by utilizing an advanced in-silico predicted search space, leading to a notable increase in proteome coverage without compromising quantification quality.

When dia-PASEF or dda-PASEF data is searched with Spectronaut independently or using Spectronaut through Bruker ProteoScope, the data produced is of exceptional quality. This is where that Mass Dynamics intervenes, transforming data into actionable insights.



Through the Mass Dynamics platform, users can transition from library-based or directDIA+ based dia-PASEF data to interactive and intuitive visualizations. This process empowers scientists to efficiently extract knowledge from potentially intimidating proteomics data, setting the stage for addressing complex biological questions.

Additionally, features such as real-time collaboration and the ability to cross-reference both external and internal knowledge bases enhance the overall user experience, promoting a collaborative and informed research environment.

Transforming data into insights

Proteomics and mass spectrometry (MS) have become pivotal tools for uncovering the molecular intricacies distinguishing disease from healthy samples in the landscape of both disease research and drug discovery. This technique is at the heart of contemporary proteomics, revealing the dynamic shifts within the proteome.

The proteomic workflow begins with a robust experimental design, followed by the meticulous preparation phase that includes protein purification, enzymatic digestion, and sample cleanup. The core of this procedure is the data acquisition through a liquid chromatography MS (LC-MS) system. Once processed, this data provides an extensive list of identified proteins, highlighting the quantitative differences in protein abundance across varied conditions.

Yet, simply identifying proteins is just the starting point. The real task lies in deciphering the biological relevance of these proteins. By leveraging sophisticated statistical techniques, such as differential expression analysis, we can start to unravel the biological pathways and processes impacted by these proteomic alterations.

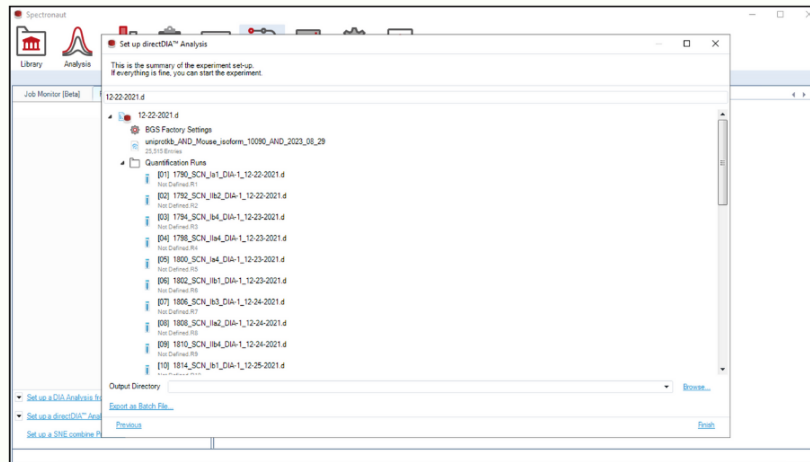
The pinnacle of this workflow—extracting actionable insights—is a multifaceted journey, relying on the acumen of the MS specialist, the biologist's insights, and the selected bioinformatics strategies. This dynamic process demands the synergy of multidisciplinary teams, each bringing their unique expertise to drive novel discoveries and shape future experiments.

This Application Note showcases Mass Dynamics' innovative stance in this field. By seamlessly integrating LC-MS data processing with the intuitive generation and interactive exploration of biological insights, Mass Dynamics offers a highly dynamic perspective of the proteome, anchored by interactive plots, figures, and seamless linking to external knowledge.

Seamless data to knowledge

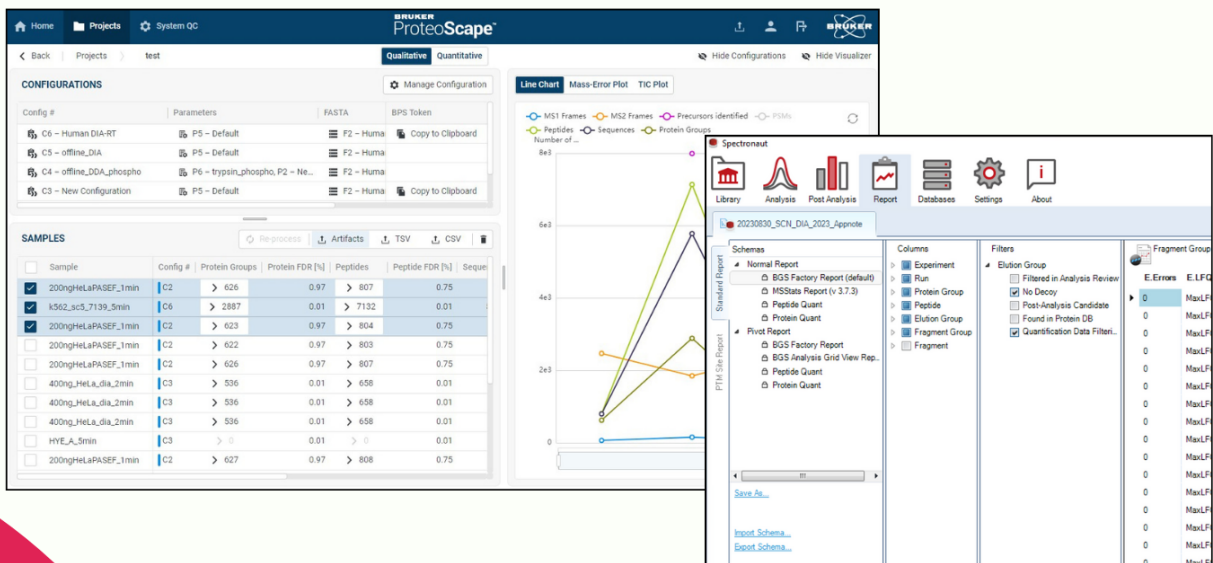
Processing data using Spectronaut using DirectDIA+

Raw dia-PASEF data can be processed following the directDIA analysis pipeline, using a protein FASTA database. After selecting your run files, a wizard guides you through the process, including a prompt to assign the relevant FASTA files, GO annotation file, and analysis settings. A summary of the analysis is shown prior to commencing.



How do you export data processed by ProteoScope or Spectronaut, to then upload to Mass Dynamics?

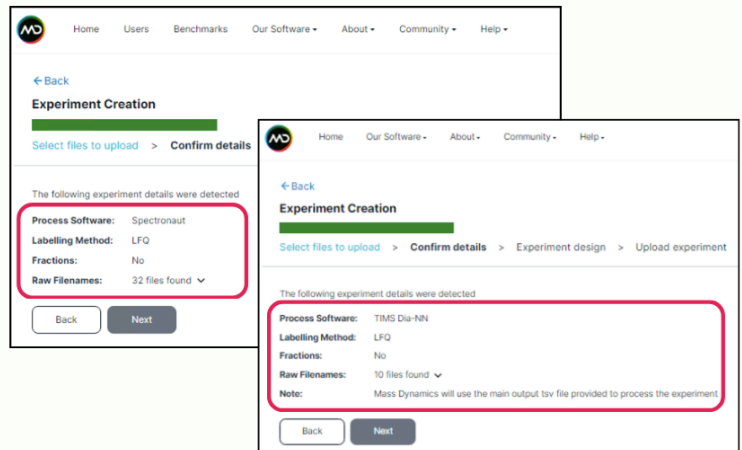
Bruker ProteoScope and Spectronaut allow straightforward export options, allowing you to design and customize results according to your needs. Seamless uploading of ProteoScope dia-PASEF and Spectronaut dia-PASEF results into Mass Dynamics requires the export of the default report.tsv and log.txt (Bruker ProteoScope) and BGS Normal report (Spectronaut).



Seamless knowledge interrogation

Uploading Spectronaut data into Mass Dynamics

The enhanced integration between Bruker ProteoScape, Spectronaut, and Mass Dynamics now facilitates effortless uploading of the default report.tsv and log.txt (Bruker ProteoScape) and BGS Normal report (Spectronaut). This revamped upload feature automatically recognizes the data format, ensuring a seamless upload process. Once uploaded, users can easily interrogate the data and compare it against both internal and external knowledge bases.



How does Mass Dynamics help you interrogate your data?

- [Enrichment Analysis](#): Dive into gene sets for profound biological insights.
- [Effortless Integration with Spectronaut](#): Directly upload BGS outputs for instant analysis.
- [Intuitive Visualizations](#): Tools like volcano plots offer comprehensive data views.
- [Knowledge Access](#): Direct links to gene set/pathway databases for enriched understanding.
- [Precision Data Processing](#): Harnessing algorithms like CAMERA [1] for meticulous data scrutiny.
- [Collaborative Insights](#): Real-time sharing and collaboration tools.
- [Swift Insight Generation](#): Accelerated visualization and hypothesis creation.
- [Contextual Exploration](#): Uncover pivotal relationships in biological pathways.

What knowledge bases are currently accessible

Mass Dynamics seamlessly integrates a blend of external and internal knowledge bases for a comprehensive view of your data.

External Knowledge:

- **Reactome** [2]: Interactive biological pathways.
- **Gene Ontology (GO)** [3]: Standardized gene annotations.
- **UniProt** [4]: Essential protein sequence and function data.
- **String** [5]: Database for protein-protein interactions.

Internal Knowledge:

- **Your Experiments**: Use your own data as a reference point. Cross-compare with other datasets, whether your own or publicly available, for deeper insights.

This knowledge fusion ensures a robust and efficient analysis of your proteomic data.

Leveraging Mass Dynamics to unravel Age and Sex Differences in Mice: A Proteomics Case Study

Deep proteome profiling reveals signatures of age and sex differences in paw skin and sciatic nerve of naïve mice

Feng Xian, Julia Regina Sondermann, David Gomez Varela, Manuela Schmidt

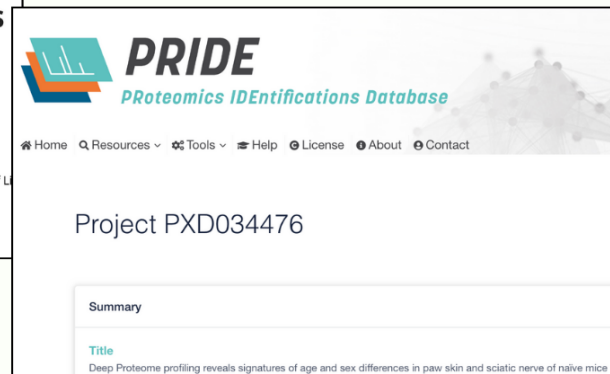
Systems Biology of Pain, Division of Pharmacology & Toxicology, Department of Pharmaceutical Sciences, Faculty of Life Sciences, University of Vienna, Austria

Nov 30, 2022 · <https://doi.org/10.7554/eLife.81431>

In biomedical research, the age and sex of mice significantly influence study results. Yet, many studies overlook these factors, leading to potential inaccuracies. This research used proteomics to analyze mouse paw skin (PS) and sciatic nerves (SCN), revealing notable differences between male and female mice of different ages [6]. The study identified key proteins and pathways affected by age and sex, offering valuable insights into skin and nerve biology. This work emphasizes the importance of considering age and sex in preclinical mouse studies for more accurate and comprehensive results.

Elevating Proteomic Analysis Efficiency with Mass Dynamics

To demonstrate the potential of modern data analysis tools, a case study was conducted, re-analyzing a dataset focused on age and sex differences in mouse proteomics using Mass Dynamics (MD).



Dataset Integration

The initial step involved raw DIA data processing using Spectronaut, followed by seamless upload of the pre-processed data (i.e., BGS default output) into MD.

Swift and Comprehensive Analysis

Leveraging MD's advanced algorithms, what would traditionally take weeks of analysis was accomplished in a significantly reduced timeframe. (hours). This speed did not compromise the depth or quality of the analysis.

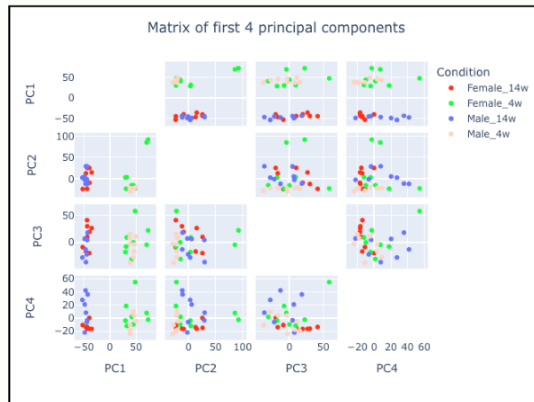
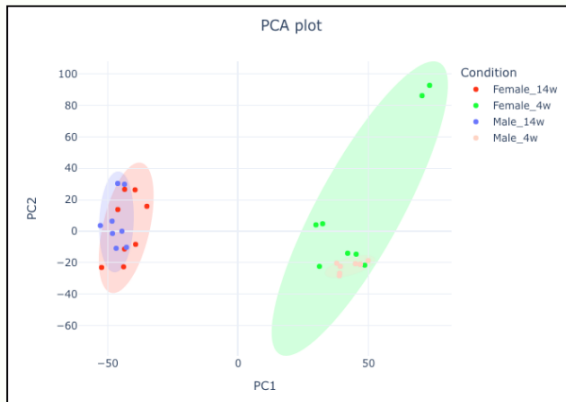
Deep Biological Insights

Beyond validating the initial findings from the dataset, MD enabled the exploration of nuanced variations in protein expressions, offering a more granular understanding of the data.

Contextual Understanding

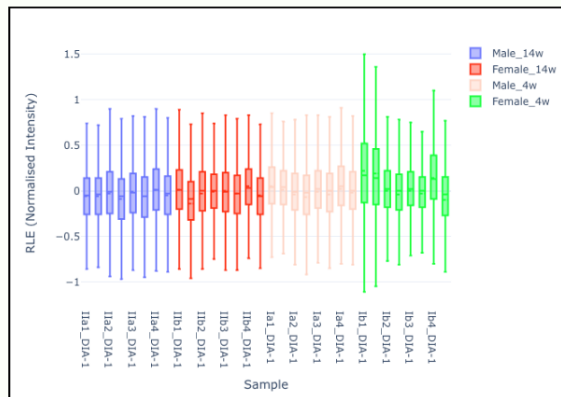
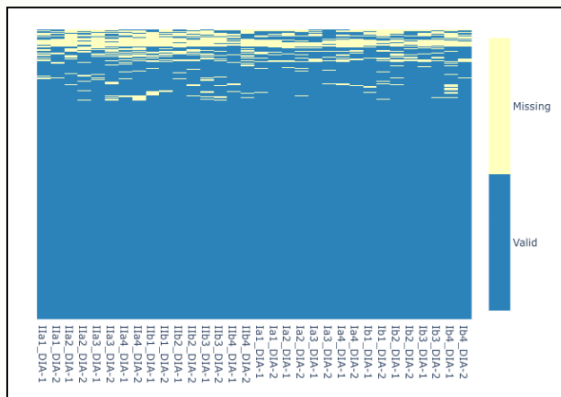
The platform's ability to correlate identified proteins with established biological pathways meant that the data wasn't just numbers — it told a story. This context provided richer insights into the implications of the findings.

Accelerated confidence in your data



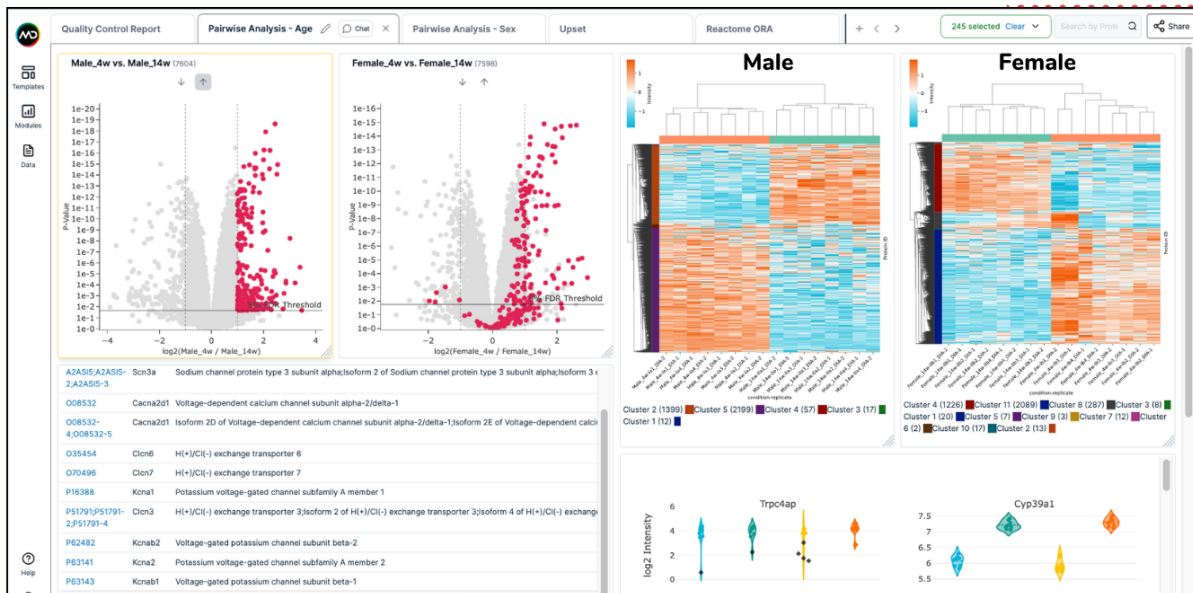
Instant Principle Component Analysis (PCA)

MD offers swift insights by automating the creation of experimental quality control figures (see above). In this instance, the left plot displays the four distinct groups on a PCA, autogenerated as part of the quality control report template. The PCA plots for mouse PS (not shown) and SCN (shown above) highlighted that the primary differentiation observed is based on age rather than sex. Meanwhile, the charts on the right showcase the data as a matrix of the top four principal components, allowing for an instant visualization of which components correlate with age or gender.



Instantly Visualise Data Coverage and Expression Variability

In the presented figures (above), the left panel showcases a "missingness" heatmap that illustrates the coverage achieved throughout the experiment. This heatmap reveals minimal data gaps, with only minor disparities between the sample groups. On the other hand, the right panel features a relative log expression box plot. This plot quickly highlights a marginally increased variation in the final group, suggesting a potential need for a more detailed examination of the experimental data for that specific group.

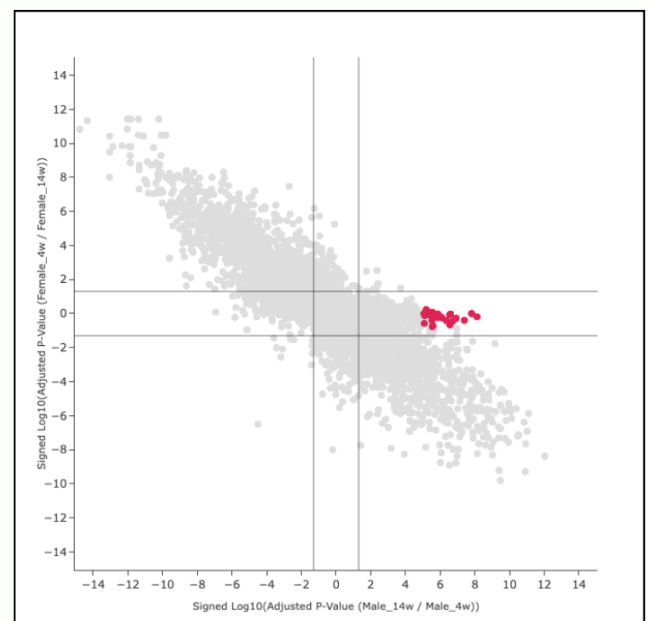


Instant interactive volcano plot and heatmaps

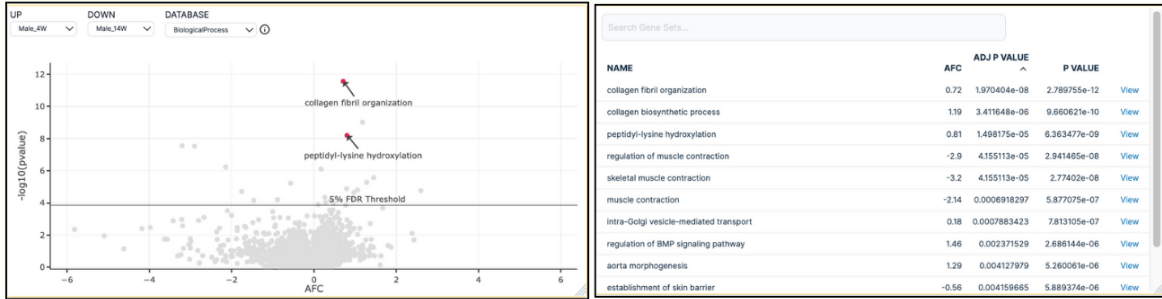
As shown in the figure above, researchers can customize tabs with various modules to delve deeper into their results. For instance, the volcano plots provide a pairwise analysis highlighting differential protein expression. These interactive volcano plots enable users to compare results across different pairs within their experiment, generate lists of proteins, and further examine these proteins using additional modules like the heatmap. These curated protein lists can then be utilized for over-representation analysis, pinpointing relevant pathways for the study. Additionally, for a comprehensive dataset assessment, users can opt for the 'Gene Set Enrichment Analysis (GSEA)' template under Enrichment Analysis.

Instant Log-Log P-Value Plots

Log-log p-value plots are powerful graphical representations that allow for the visualization of data spanning several orders of magnitude, especially when assessing the significance of changes in large datasets. Plotting both axes on a logarithmic scale provides a clearer view of the distribution of p-values, making it easier to discern patterns and clusters within the data. In the provided figure, the log-log plot distinctly reveals a cluster of highly up-regulated proteins in the male mice (14w vs 4w) compared to the female mice (14w vs 4w). This visualization technique, thus, offers an intuitive way to quickly identify and interpret significant trends and differences in complex datasets.



Interrogate knowledge in context



Understand biological changes in the context of your experiment

While age and sex in rodent-based research are known to strongly affect experimental outcomes *in vivo* and *in vitro*, comprehensive interrogation of proteome differences is still limited [6]. Understanding how changes in observed protein abundance from proteomic studies relate to known biological functions is crucial to define better what the changes represent. For example, adopting the GSEA template to interrogate the Gene Ontology Biological Process (GO-BP) showed enriched pathways for collagen fibril organization and peptidyl-lysine hydroxylation (as reported in [6], see Figure above for male PS) when comparing 4-week PS vs. 14-week PS, irrespective of sex.

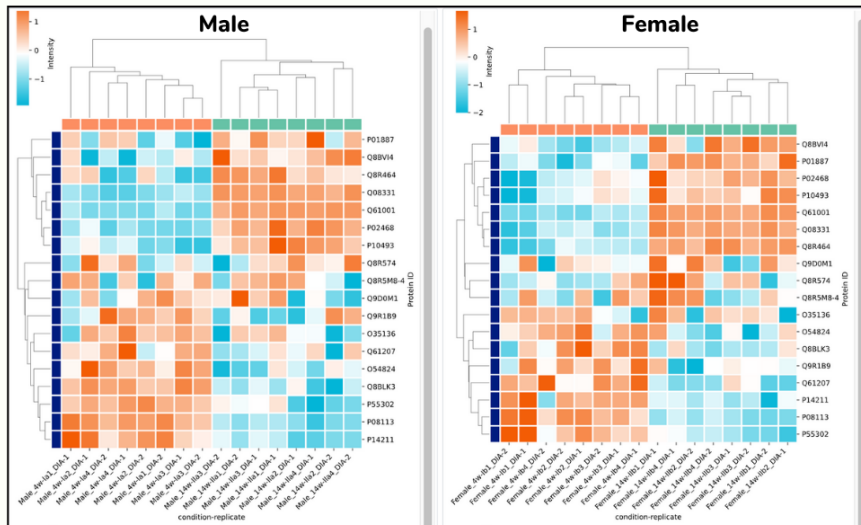
Over-representation analysis using Reactome also showed ‘collagen formation’, ‘collagen biosynthesis and modifying enzymes’, and ‘Assembly of collagen fibrils and other multimeric structures’ for males and females at 4-week PS (data not shown). In contrast, ORA showed distinct pathways according to sex at 14 weeks, with ‘Striated Muscle Contraction’ being over-represented in male PS only and ‘Terminal pathway of complement’ being over-represented in female PS at 14 weeks. The proteins associated with muscle contraction can be overlaid with volcano plots in the service to represent this (refer to the Figure below).



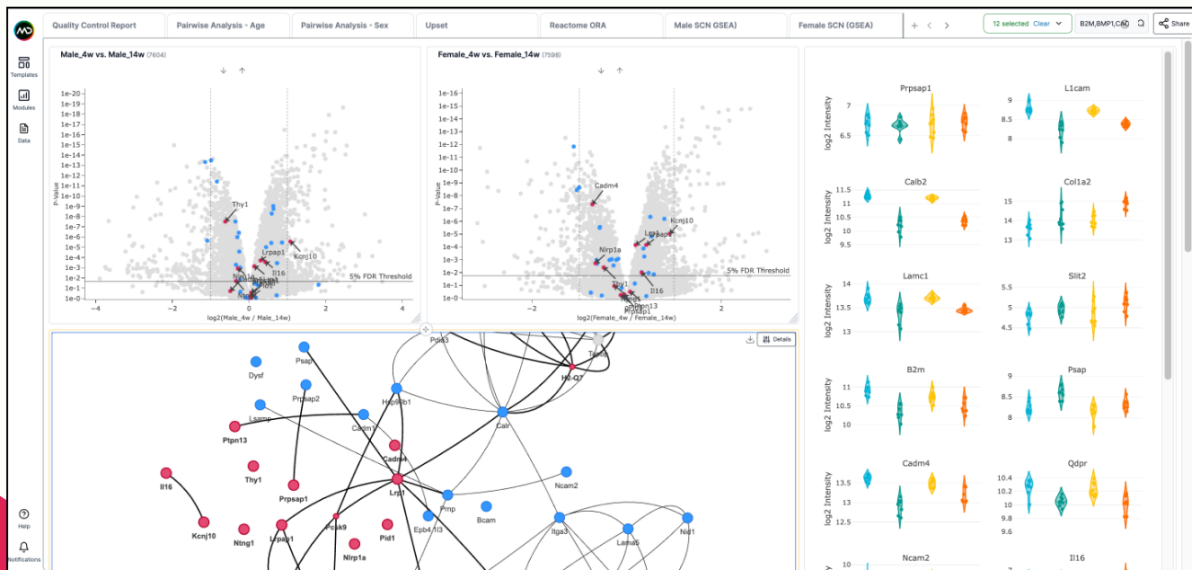
Interrogate knowledge in context

Understand biological changes in the context of your experiment

Known information can be easily assessed on the Mass Dynamics platform, such as with importing protein or gene lists for further analysis. For instance, known ligands of neuron receptors are known for their importance for SCN (patho)physiology [6]. Analysis of observed glial cell ligands showed a portion displaying a strong age dependency, as shown in the heatmaps below.



Integration with STRING can be leveraged to interrogate known interactions and how these relate to observed abundance changes.



Sharing and Collaboration

Share, collaborate and publish, allow independent analysis

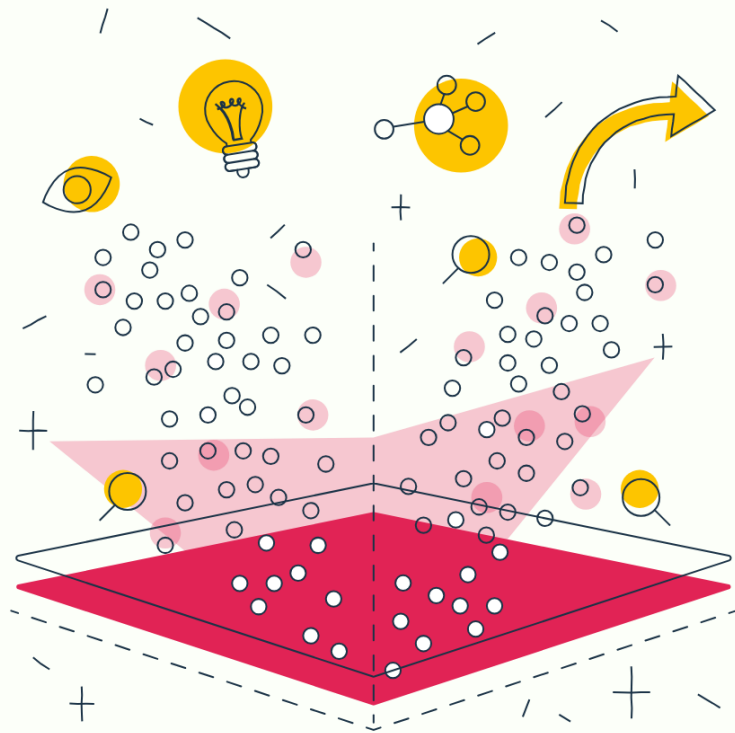
The MD platform is built on a cloud-based infrastructure, not requiring any downloads, installations, licences or management of infrastructure. It has sharing and commenting features, with direct notifications in-app and by email, with the ability to define user access rights. Modules in the service allow note-taking and checklists for improved collaboration. Images and results can be exported as entire reports or specific modules to *.SVG, *.PNG, as required. Finally, the analysis of results can be made public to allow interactive assessment of results by reviewers and the community.

The screenshot displays the MD platform's sharing and collaboration interface. At the top, there is a header bar with a selection indicator '111 selected Clear' and a search box 'Search by Proteins'. A 'Share' button is located on the right. Below the header, a 'Comments for Knowledge Integration' section is visible, featuring three comments from users 'Anna' and 'mark-admin@massdynamics.com'. The interface also includes a vertical sidebar with a scatter plot and a slider for 'R Threshold' set between 8 and 10. At the bottom, there is a text input field for 'Add a comment' and a 'Publish' button.

Conclusion

The synergy between Bruker's data generation (Data), raw data processing by the Bruker ProteoScape and Biognosys Spectronaut software (Information), and Mass Dynamics' analytical capabilities (Knowledge) has revolutionized the proteomics research landscape. The integrated workflows ensure that data is generated and processed with precision and interpreted with depth and clarity. Mass Dynamics, in particular, offers a user-friendly platform that facilitates in-depth analysis of protein data generated by Bruker ProteoScape and Spectronaut. Users can efficiently identify differential protein expressions and generate relevant protein lists through customizable modules like interactive volcano plots. These lists can be further analyzed using tools like heatmaps and the 'Gene Set Enrichment Analysis (GSEA)' template, ensuring a comprehensive understanding of the data and the identification of pertinent pathways. This streamlined approach enhances data interpretation and accelerates the discovery process, making the journey from data generation to insights both seamless and efficient.

See how you can explore Proteomics data with Mass Dynamics



About Bruker Corporation (Nasdaq: BRKR)



Bruker is enabling scientists to make breakthrough discoveries and develop new applications that improve the quality of human life. Bruker's high-performance scientific instruments and high-value analytical and diagnostic solutions enable scientists to explore life and materials at molecular, cellular, and microscopic levels. In close cooperation with our customers, Bruker is enabling innovation, improved productivity, and customer success in life science molecular and cell biology research, in applied and pharma applications, in microscopy and nanoanalysis, as well as in industrial applications. Bruker offers differentiated, high-value life science and diagnostics systems and solutions in preclinical imaging, clinical phenomics research, proteomics and multiomics, spatial and single-cell biology, functional structural and condensate biology, as well as in clinical microbiology and molecular diagnostics. Please visit www.bruker.com

About Biognosys



At Biognosys, we believe that deep proteome insights hold the key to breakthrough discoveries that transform science for better lives. With our versatile portfolio of next-generation proteomics solutions, including the TrueDiscovery®, TrueTarget®, and TrueSignature® research service platforms, our flagship software Spectronaut®, and the PQ500™ kit, we make the proteome actionable to empower research, drug development, and clinical decision-making. Our solutions provide a multi-dimensional view of protein expression, function, and structure in all biological species and sample types. Our unique, patented technologies utilize high-resolution mass spectrometry to quantify thousands of proteins with industry-leading precision, depth, and throughput. Through our strategic partnership with Bruker (Nasdaq: BRKR), we make proteomics globally accessible. For more information, visit biognosys.com.

About Mass Dynamics



Our mission is to free humanity and society from the burden of disease by unlocking the magic of Mass Spectrometry (MS), Proteomics and the power of existing biological knowledge. We do this by delivering a powerful software platform that seamlessly connects multi-disciplinary life scientists to answer biological questions and understand the building blocks of life - better, faster and easier.



www.massdynamics.com



hello@massdynamics.com



github.com/massdynamics



Sign up for a free trial: <https://massdynamics.com/free-trial>

References

1. Wu, D, and Smyth, GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Research* 40, e133 (2012).
2. Fabregat A, et al. Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinformatics*;18(1):142 (2017).
3. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* 47, D330–D338 (2019).
4. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 49, D480–D489 (2021).
5. Szklarczyk D, Kirsch R, et. al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 2023 Jan 6;51(D1):D638-646.
6. Xian F, Sondermann JR, et. al. Deep proteome profiling reveals signatures of age and sex differences in paw skin and sciatic nerve of naïve mice. *Elife.* 2022 Nov 30;11:e81431.