

quantms, ms²rescore and multiple search engines
enables deep proteome coverage across protein
quantification, immunopeptidomics, and
phosphoproteomics experiments

Chengxin Dai^{1,2}, Ralf Gabriels^{3,4}, Robbin Bouwmeester^{3,4}, Jonas Scheid^{5,6,7,8}, Lennart Martens^{3,4,9,10}, Oliver Kohlbacher¹², Mingze Bai¹¹, Timo Sachsenberg¹², and Yasset Perez-Riverol^{*13}

¹State Key Laboratory of Medical Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences (Beijing), Beijing Institute of Lifeomics, 102206, Beijing, China

²International Academy of Phronesis Medicine (Guangdong), 510320, Guangdong, China

³CompOmics, VIB Center for Medical Biotechnology, VIB, Ghent, 9052, Belgium

⁴Department of Biomolecular Medicine, Faculty of Medicine and Health Sciences, Ghent University, Ghent, 9052, Belgium

⁵Department of Peptide-based Immunotherapy, Institute of Immunology, University and University Hospital Tübingen, Tübingen, Germany

⁶Cluster of Excellence iFIT (EXC2180) "Image-Guided and Functionally Instructed Tumor Therapies", University of Tübingen, Tübingen, Germany

⁷Quantitative Biology Center (QBiC), University of Tübingen, Tübingen, Germany

⁸Institute for Bioinformatics and Medical Informatics (IBMI), University of Tübingen, Tübingen, Germany

⁹BioOrganic Mass Spectrometry Laboratory (LSMBO), IPHC UMR 7178, University of Strasbourg, CNRS, Strasbourg, 67000, France

¹⁰Infrastructure Nationale de Proteomique ProFI - FR2048, Strasbourg, 67087, France

¹¹Chongqing Key Laboratory of Big Data for Bio Intelligence, Chongqing University of Posts and Telecommunications, Chongqing, China

¹²Department of Computer Science, Applied Bioinformatics, University of Tübingen, Tübingen, Germany

¹³European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Cambridge, United Kingdom

Abstract

The exponential growth of public proteomics datasets has surpassed the analytical capacity of traditional desktop tools, particularly for large-scale automated reanalysis. To address this challenge, we present an integrated workflow that combines quantms, a cloud-native pipeline, with MS²Rescore, a machine learning-based rescoring tool. This workflow enables deep and scalable reanalysis of massive proteomics datasets. Powered by the Nextflow engine for parallel computing, the pipeline incorporates fragment ion intensity predictions from MS²PIP and retention time predictions from DeepLC, improving peptide-spectrum match reliability through Percolator. We applied this approach to four representative datasets covering label-free quantification, TMT labeling, immunopeptidomics, and phosphoproteomics. Compared to traditional methods, our workflow achieved a 16–22.8% increase in identified spectra, along with the quantification of hundreds of additional proteins and phosphosites. These improvements demonstrate that integrating multiple search engines with machine learning-derived features not only enhances identification sensitivity but also deepens quantitative insights for downstream biological interpretation. Overall, this workflow offers a reproducible and scalable solution for the reanalysis of public proteomics data, advancing FAIR principles by promoting scientific transparency, accessibility, and data reuse.

Keywords: Proteomics, Reanalysis, Workflow, Machine learning.

*Corresponding author: yperez@ebi.ac.uk

1 Introduction

In recent years, the field of proteomics has experienced rapid growth in the availability of publicly accessible datasets, accompanied by a shift toward studies analyzing larger sample cohorts. As of June 2025, over 40,000 datasets have been submitted to ProteomeXchange (PX) repositories, including a substantial increase in large-scale submissions comprising more than 100 instrument files [1]. However, conventional desktop tools such as MaxQuant [2], pFind [3], PeptideShaker [4], and Proteome Discoverer are limited in their capacity to perform automated, large-scale quantitative analyses in cloud or distributed environments, hindering the reanalysis of extensive experiments on standard workstations.

We recently developed quantms, an open-source, cloud-based pipeline designed for massively parallel reanalysis of quantitative proteomics datasets [5]. The pipeline is highly modular and flexible, accommodating a wide range of quantitative proteomics approaches. quantms automatically distributes computations using the Nextflow workflow engine across one or more computing nodes, depending on the number of instrument files and samples [6]. To ensure traceability and reproducibility, the pipeline is built entirely on standardized open file formats [7] [8] and reproducible execution environments such as Docker and Singularity, adhering strictly to the FAIR (Findability, Accessibility, Interoperability, and Reusability) principles [9].

With the adoption of machine learning (ML) in the field of proteomics, various models have been developed to accurately predict peptide behavior in LC-MS, such as MS²PIP [10] and DeepLC [11] for fragment ion intensities and retention time prediction, respectively. Early approaches employed decision trees and single-layer neural networks, while more recent deep learning models such as Prosit [12] achieve significantly improved accuracy for predicting fragment ion intensities and retention times. These highly accurate predictions enable superior matching of experimental data to theoretical expectations and have reinvigorated rescore strategies in proteomics. MS²Rescore is a modular Python package that generates multiple features assessing the similarity between observed and predicted peptide behavior,

such as fragment ion intensities, retention time, and ion mobility [13].

Previously, quantms did not leverage measurable peptide properties such as fragment ion intensities and retention times. To overcome this limitation, we integrated MS²Rescore into quantms and incorporated customized features following Nextflow and nf-core best practices. We demonstrate that the enhanced pipeline supports in-depth analysis of large-scale public proteomics datasets across diverse experimental designs, including label-free quantification (LFQ), tandem mass tag (TMT)-based quantification, immunopeptidomics, and phosphoproteomics studies.

2 Methods

2.1 MS/MS Data and quantms settings

To develop and evaluate the performance of the quantms-integrated MS²Rescore workflow, we selected four publicly available benchmark datasets. Three were obtained from the PRIDE Archive under the identifiers PXD001819, PXD019643, and PXD026824, and one from the CPTAC data portal under PDC000127. The PXD001819 dataset contains 48 Sigma UPS1 proteins spiked into a background of yeast cell lysate at nine different concentrations: 0.05, 0.125, 0.25, 0.5, 2.5, 5, 12.5, 25, and 50 fmol/ μ L to evaluate quantification performance. We evaluated five different quantms workflow settings: (1) Comet, (2) Comet and MSGF+, (3) Comet with MS²Rescore, (4) Comet and MSGF+ with MS²Rescore, and (5) Comet, MSGF+, and SAGE with MS²Rescore to explore multiple search engines and their integration with MS²Rescore for improved identification and quantification results. All search parameters were the same as described in the previous publication. An FDR filter at 1% was applied at peptide spectrum match (PSM) and protein level at the dataset level. The search results from quantms and MaxQuant at the PSM and protein group levels are provided in Supplemental File S1.

2.2 Rescoring Features and Postprocessing

The quantms-rescoring Python package integrates MS²Rescore and computes a broad range of features based on DeepLC-predicted versus observed retention times, as well as MS²PIP-predicted versus observed MS2 fragment ion intensities. Signal-to-noise ratio (SNR) features are also computed for specific scenarios. To enhance compatibility and performance, we implement new strategies for model selection, feature export, and model training. For retention time (RT)-related features, retraining of the DeepLC model is supported, allowing the selection of the model with the lowest mean absolute error (MAE) on calibration PSMs. These calibration PSMs are selected from each run at a specified ratio. For MS2 intensity-related features, we perform overfitting tests using the top-scoring 60% of non-decoy PSMs in the calibration set. quantms-rescoring automatically selects the most suitable MS²PIP model for a given batch of PSMs by comparing the correlation scores of predicted and observed intensities across candidate models. Furthermore, it verifies whether the MS²PIP predictions meet predefined correlation and scoring thresholds. This is achieved by filtering out decoy PSMs, ranking the results by PSM score, and selecting a calibration subset. The method ensures that at least 80% of calibration PSMs exceed the correlation threshold. If none of the models satisfy this criterion, MS²PIP rescoring is skipped unless explicitly forced by the user. This mechanism filters out pre-trained models that are incompatible with the data. Additionally, users can specify a list of features for export. All steps are fully parallelized at the run level, significantly reducing runtime for large-scale datasets.

Then, the quantms workflow calculates a posterior error probability for each PSM using Percolator. This is performed under three different feature configurations (1) a baseline model using only search engine-derived features, (2) the baseline model plus MS²Rescore-derived features, and (3) the above combined with the newly added SNR features. To merge results from multiple search engines, ConsensusID aggregates PSMs into unified scores. Final PSM-level q-values are then obtained either directly from Percolator or calculated using OpenMS’s target-decoy approach based on the predicted probabilities. For phosphoprote-

teomics datasets, LuciPHOr2 is employed to assign site-level localization scores and estimate the associated false localization rate using tools from the OpenMS toolkit.

3 Results

3.1 quantms with MS²Rescore Enhances Label-Free Identification and Quantification

To systematically evaluate the performance of the quantms-integrated MS²Rescore workflow at both identification and quantification levels, we first analyzed the public benchmark dataset PXD001819. For benchmarking the identifications, five different workflows configurations were designed and compared, including (1) Comet with Percolator, (2) two search engines with Percolator, (3) Comet with MS²Rescore features and Percolator, (4) two search engines with MS²Rescore features and Percolator, and (5) three search engines with MS²Rescore features and Percolator to determine whether features from fragment intensity-based and retention time-based predictors enhanced the identification and quantification process.

Significant PSMs were filtered based on q-values, and the FDR was used as a key metric to compare workflows. As shown in Figure 1, using consensus scores from two search engines significantly increased identification rates at a fixed FDR threshold. Specifically, combining Comet and MSGF+ improved identified spectra by 17% over Comet alone, and incorporating MS²PIP and DeepLC features through MS²Rescore led to an additional 16% increase. quantms achieved a 28% improvement in the number of PSM identifications compared to MaxQuant. At the quantification level, including MS²Rescore-derived features allowed more low-abundance UPS1 proteins to be quantified (Figure 2B), highlighting the contributions of the integrated workflow. Although MaxQaunt reported more UPS proteins in 2500amol, the proteins that were only quantified in MaxQuant were all quantified by match between runs (MBR) rather than from direct MS/MS identifications. The main reason for this is

that quantms used different MBR quality control.

To better understand the contribution of individual features, we extracted the top 20 feature weights from Percolator (Supplemental Figure 1). Over half of the top-weighted features were derived from MS²Rescore. Notably, SpecPearsonNorm had a strong positive weight, indicating that a better correlation between predicted and observed intensities improves confidence. Conversely, RtDiffBest had a negative weight, suggesting that large deviations in retention time reduce match quality. Interestingly, peptide length also emerged as a significant positive feature after MS²Rescore integration, likely because longer peptides generate more fragment ions and thus are more reliably identified.

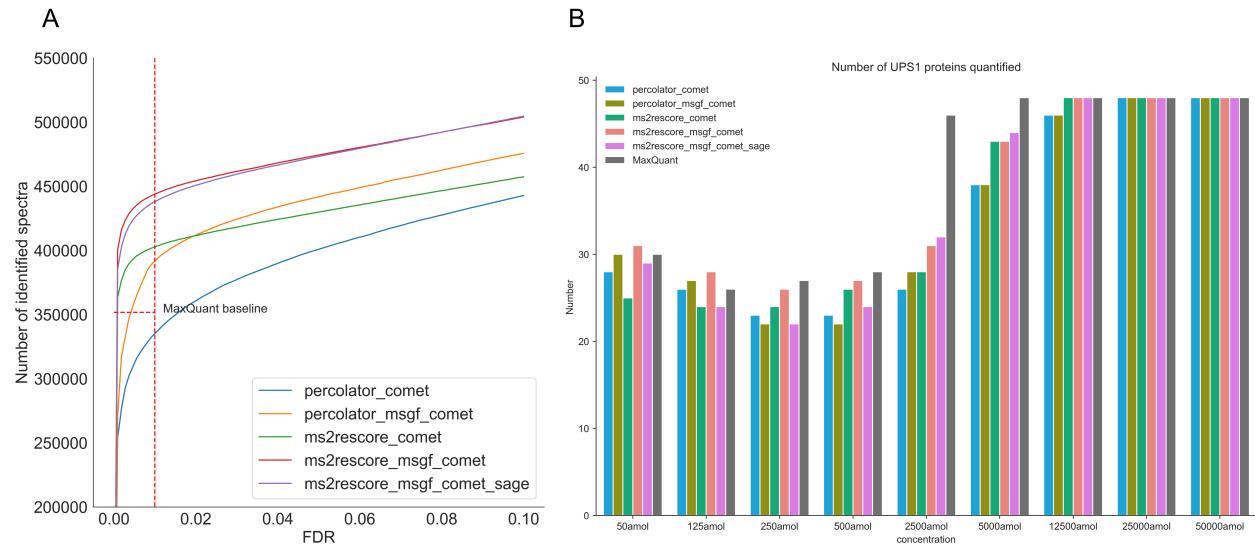


Figure 1: The number of identified spectra as a function of differing FDR levels for different workflow settings. (A) Results for the PXD001819 dataset. (B) Results for the PXD001819 dataset showing protein quantification. Five different workflow settings are shown: the combination of Comet search engine with Percolator rescore, the combination of Comet with MSGF+, the combination of two search engines with MS²Rescore features, the combination of two search engines with MS²Rescore and SNR features, and the combination of three search engines with MS²Rescore and SNR features. Multiple search engines consensus identification, additional intensity-based and retention time-based features, and SNR features lead to both higher identification rates and high-quality identification.

3.2 Improved Identification and Quantification in TMT Experiments via MS²Rescore Integration

We next applied the workflow to a large-scale TMT-labeled dataset from CPTAC (PDC000127). The integration of multiple search engines and MS²Rescore features led to substantial improvements in both identification and quantification (Figure 2). The PSM identification rate increased by 3.6%, and 921 proteins were newly quantified compared to the workflow without MS²Rescore (Figure 2A, 2B). Among these, 59 proteins had abundance levels within the top 10% (Figure 2C). To assess the biological impact of rescoreing, we conducted differential expression analysis using the MS²Rescore-enhanced workflow. As shown in Figure 2D, 27 newly quantified proteins were significantly differentially expressed. Notably, FOXG1—one of these proteins—has been associated with prognosis in Clear Cell Renal Cell Carcinoma, as reported by Yang et al. [14]. These findings illustrate that quantms incorporating MS²PIP and DeepLC-derived features enhances not only identification sensitivity but also enables deeper biological insights.

We further examined the rescoreing contribution of individual features by analyzing the top 20 SVM weights from Percolator (Supplemental Figure S2). For Comet-based results, SpecPearsonNorm, DotProdIonYNorm, and RtDiffBest were highly weighted. The interpretations of these weights were consistent with those from the label-free dataset. Importantly, peptide length again emerged as a key feature. Similar weight distributions were observed for MSGF+ and SAGE results, confirming the generalizability of these feature effects across different search engines.

3.3 quantms with MS²Rescore improves HLA Class I immunopeptidome identification

The HLA Ligand Atlas (cite) (PXD019643) is a comprehensive immunopeptidomics repository of benign primary tissue. To further assess the performance of the quantms workflow

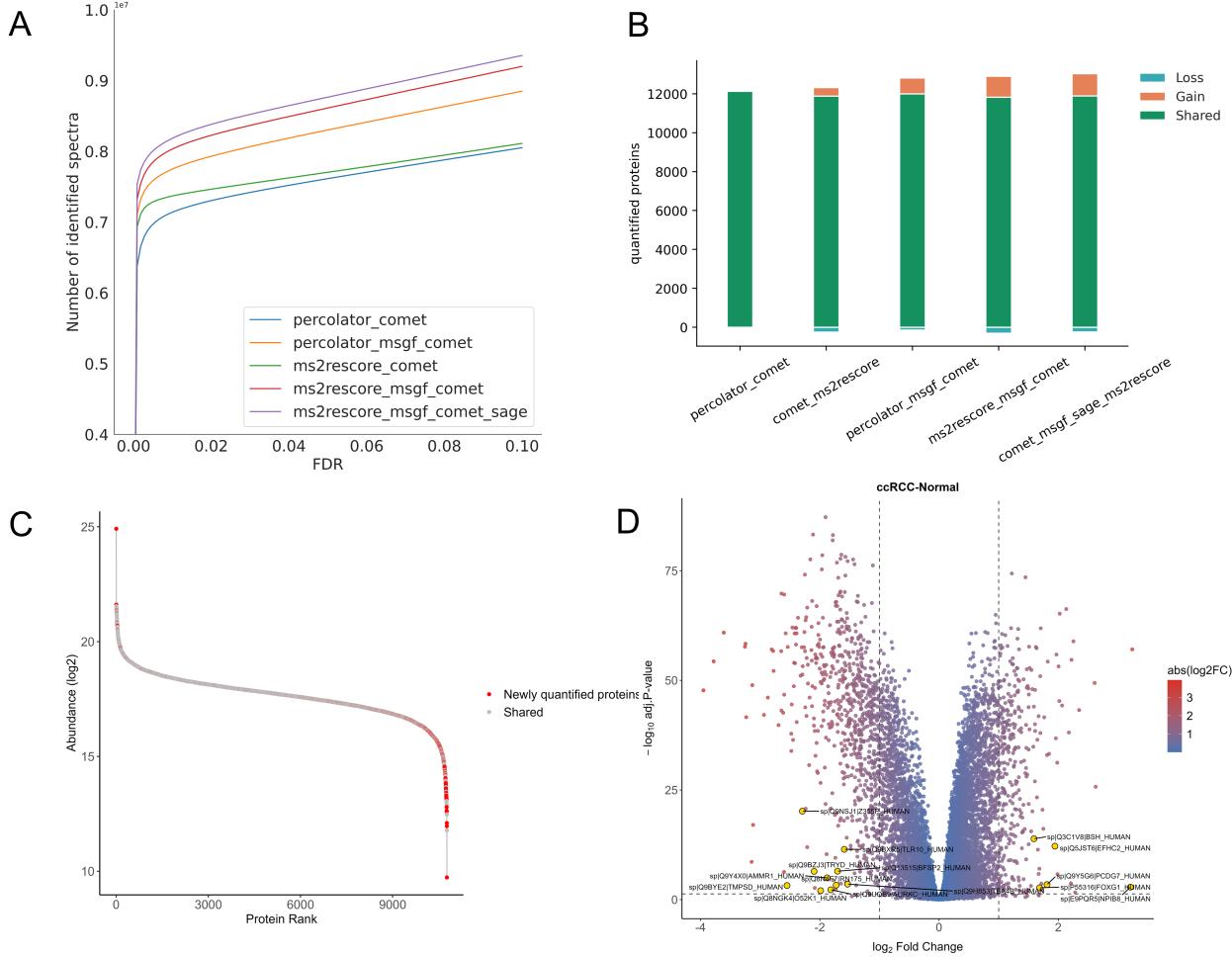


Figure 2: Comparison of identification and quantification results for different workflow settings on PDC000127. (A) The number of identified spectra as a function of differing FDR levels for different workflow settings. (B) The number of quantified proteins. The green part indicates the intersection between a workflow and Percolator-Comet results. (C) The rank of protein abundance from Comet and MSGF+ with MS²Rescore. The red dots represent proteins quantified only in the Comet and MSGF+ with MS²Rescore workflow compared to the Comet and MSGF+ workflow without MS²Rescore. (D) The volcano plot for differential expression analysis from MSstatsTMT. The yellow dots represent proteins quantified only in the Comet and MSGF+ with MS²Rescore workflow compared to the Comet and MSGF+ workflow without MS²Rescore.

using MS²Rescore, we analyzed the HLA Class I data of the HLA Ligand Atlas. Five workflow configurations were evaluated: (1) Comet + Percolator, (2) Comet + MS²Rescore, (3) Comet + MSGF+ + Percolator, (4) Comet + MSGF+ + MS²Rescore features, and (5) Comet + MSGF+ + MS²Rescore + SNR features. These configurations were compared

based on both the total number of peptide-spectrum matches (PSMs) and the number of unique peptide sequences identified. Across all comparisons, the use of multiple search engines combined with MS²Rescore features significantly enhanced identification performance. At both 1% and 0.1% FDR thresholds, integrating MSGF+ with Comet increased the number of identified spectra by 11.7% compared to using Comet alone. Adding MS²Rescore features on top of the multi-engine workflow yielded a further 22.8% improvement, as shown in Figure 3.

To better understand the impact of MS²Rescore-derived features on rescoring, we visualized the distributions of decoy PSMs, rejected target PSMs, and accepted target PSMs using the Pearson correlation coefficient (PCC) and retention time error (Figure 3C, D). The accepted target PSMs were clearly separated from both decoy and rejected targets based solely on PCC values, and exhibited consistently low retention time deviations. These patterns were observed across both individual search engine support and multi-engine consensus, further confirming that the integration of intensity and RT-based features enhances PSM discrimination. Overall, these results demonstrate the robustness of the quantms workflow in immunopeptidomics, highlighting its ability to leverage machine learning-derived features to maximize identification accuracy through multi-engine consensus rescoring.

3.4 Enhanced Phosphopeptide Identification and Localization Using MS²Rescore integration

In addition, we investigated the performance of our quantms workflow with MS²Rescore on post-translational modification experiments (PXD026824). For phosphoproteomics analyses, different workflow settings were evaluated: (1) Comet + Percolator, (2) Comet + MS²Rescore, (3) Comet + MSGF+ + Percolator, and (4) Comet + MSGF+ + MS²Rescore features. As shown in Figure 4, combining search engine consensus results with MS²Rescore features substantially improves Percolator’s ability to discriminate between true and false PSMs. This is evident from the leftward shift in score distributions compared to the Comet-

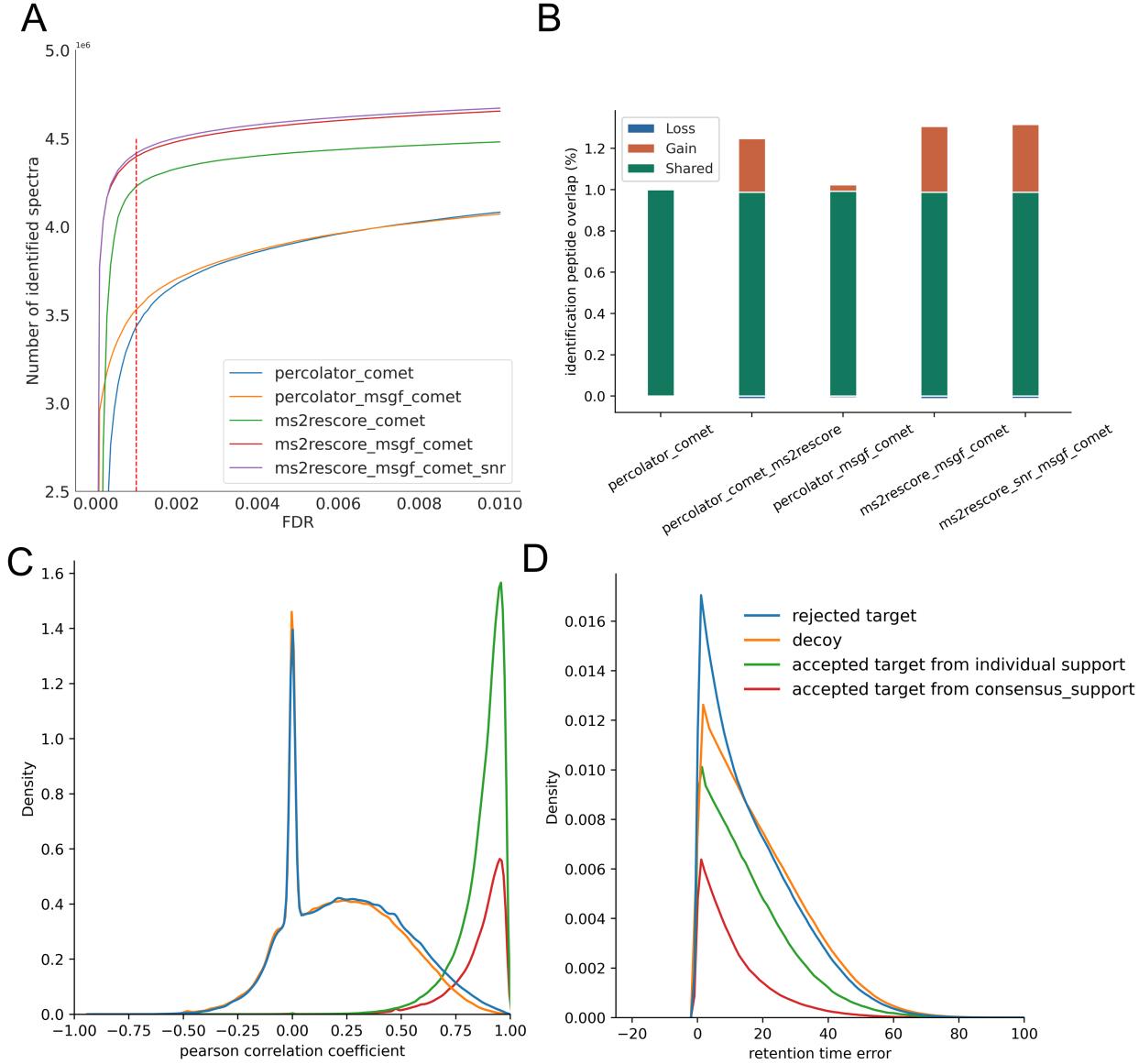


Figure 3: Comparison of immunopeptides identification results for different workflow settings on PXD019643. (A) Line plot showing the number of spectra identified for four workflow settings at different PSM FDR levels. (B) Percentage of unique identified peptides using different workflow settings. Density plots showing the distribution of the smallest retention time error between observed and predicted retention time (C) and the Pearson correlation between observed and predicted peak intensities (D) for each PSM split into decoys (red), rejected targets with q-value >0.01 (blue), and accepted targets from individual support and consensus support with q-value <0.01 (green). Note that the rejected target distribution coincides with the decoy distribution. The pronounced spike at Pearson correlation = 0 reflects spectrum comparisons in which all predicted or observed peaks have zero intensity, causing a division-by-zero situation that leaves the Pearson correlation undefined.

only workflow, indicating improved stringency. Additionally, an upward shift in the plot demonstrates a notable gain in the identification rate. Specifically, incorporating MS²Rescore features into the two-search-engine consensus increased the number of identified spectra by 19% compared to consensus identification alone (Figure 4A). The top 20 feature weights from Percolator, displayed in Supplemental Figure S4, highlight the consistent importance of features such as DotProdIonYNorm, similar to previous datasets.

Given the importance of false localization rate (FLR) control in phosphoproteomics, we also assessed the impact of different workflows on phosphopeptide identification at varying FLR thresholds (Figure 4B, C). At a 0.01 local FLR, the consensus workflow with MS²Rescore features identified 17% more phosphorylated peptides than its counterpart without MS²Rescore. Notably, 1,345 phosphorylated peptides were uniquely identified by the MS²Rescore-enabled workflow. At the phosphosite level, this workflow also uncovered 350 novel protein phosphorylation sites not reported in the other settings (Figure 4D). Collectively, these results underscore the value of MS²Rescore in enhancing both identification sensitivity and site-level localization in phosphoproteomics, establishing quantms as a robust solution for large-scale PTM data analysis.

4 Discussion

Advancements in deep learning-based tools have significantly improved the sensitivity of peptide-spectrum match (PSM) rescoring, yet their integration into streamlined, reproducible, and quantitative proteomics workflows has remained limited, especially for large-scale public data analysis. In this study, we demonstrate the systematic incorporation of MS²Rescore into the open-source quantms pipeline and evaluate its impact across various experimental settings, including label-free quantification (LFQ), tandem mass tag (TMT)-based quantification, immunopeptidomics, and phosphoproteomics. Our results show that enhanced feature sets derived from MS²PIP and DeepLC not only increase the identifica-

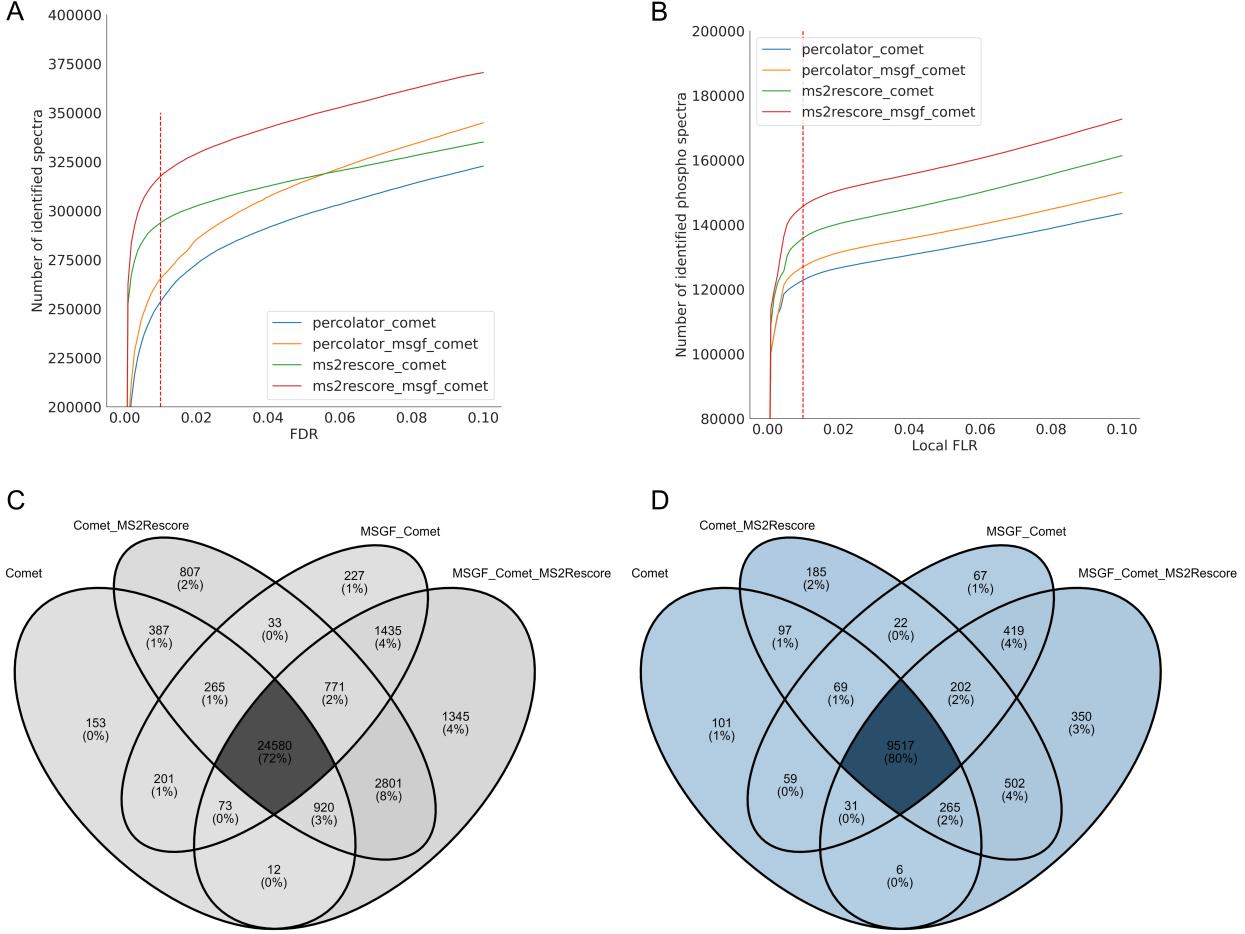


Figure 4: Comparison of phosphorylated peptide identification results for different workflow settings on PXD026824. (A) Line plot showing the number of spectra identified for four workflow settings at different PSM FDR levels. (B) Line plot showing the number of phosphorylated spectra identified for four workflow settings at different local false localization rate levels after an FDR of less than 0.01 at the PSM level. (C) Venn diagram of peptides quantified for four settings. (D) Venn diagram of protein phosphosites at protein FDR 0.01 and FLR 0.01 for four settings.

tion rates but also improve the quantification depth. Importantly, we provide evidence that the improved identification process via rescoring propagates downstream into quantification, leading to increased numbers of proteins with reliable abundance estimates, as well as enhanced detection of differentially expressed proteins, and enhanced localization of PTMs.

To demonstrate the applicability and advantages of quantms integrated with MS²Rescore in quantitative proteomics, we performed a series of reanalyses using publicly available large-scale datasets with well-established benchmarks. In label-free experiments, MS²Rescore-

enhanced workflows achieved significant increases in PSM and peptide identification rates while maintaining or improving quantification reproducibility across replicates. For TMT experiments, the integration of MS²Rescore into quantms increased the number of quantifiable proteins. In more challenging applications, such as immunopeptidomics and phosphoproteomics, the deep learning-based rescore pipeline consistently yielded higher identification rates and improved sensitivity. These improvements will contribute to meaningful biological insights, such as phosphosite discovery and the identification of differentially expressed proteins with potential clinical relevance. In addition, we counted the run times with different configurations and the added sensitivity from rescore does add 15%-50% of run time, as shown as Supplemental Table S1. This can be alleviated by increasing the number of quantms-rescoring parallel processes.

Overall, the results highlight the synergistic benefits of integrating machine learning-based features within end-to-end, cloud-based proteomics workflows. By embedding MS²Rescore into the quantms pipeline, we provide a practical and accessible solution for the community to leverage cutting-edge spectral prediction and retention time modeling without the need for complex manual configuration. This integrated approach enables reproducible reanalysis of large public datasets, increases identification confidence, and strengthens quantitative conclusions across diverse experimental designs. Future work will focus on expanding this approach to other types of post-translational modifications and exploring the integration of additional machine learning tools to further enhance proteomics data analysis capabilities.

Acknowledgments

We thank all those who supported this research, including funding bodies and the proteomics community for making proteomics data sets publicly available. Yasset Perez-Riverol would like to acknowledge funding from EMBL core funding, Wellcome grants (208391/Z/17/Z, 223745/Z/21/Z) and the BBSRC grant ‘DIA-Exchange’ [BB/X001911/1]. R.G., L.M., and

R.B. acknowledge funding from [12B7123N,G010023N,G028821N,12A6L24N]. L.M. acknowledge funding from the Horizon Europe Projects BAXERNA 2.0 [101080544] and COMBINE [101191739], and from the Ghent University Concerted Research Action [BOF21/GOA/033]. L.M. is further supported by the CHIST-ERA project ODEEP-EU [G0GDV23N].

References

- [1] Yasset Perez-Riverol, Chakradhar Bandla, Deepti J Kundu, Selvakumar Kamatchi-nathan, Jingwen Bai, Suresh Hewapathirana, Nithu Sara John, Ananth Prakash, Mathias Walzer, Shengbo Wang, and Juan Antonio Vizcaíno. The PRIDE database at 20 years: 2025 update. *Nucleic Acids Research*, 53(D1):D543–D553, January 2025.
- [2] Jürgen Cox and Matthias Mann. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, 26(12):1367–1372, December 2008.
- [3] Le-heng Wang, De-Quan Li, Yan Fu, Hai-Peng Wang, Jing-Fen Zhang, Zuo-Fei Yuan, Rui-Xiang Sun, Rong Zeng, Si-Min He, and Wen Gao. pFind 2.0: a software package for peptide and protein identification via tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, 21(18):2985–2991, September 2007.
- [4] Marc Vaudel, Julia M Burkhart, René P Zahedi, Eystein Overland, Frode S Berven, Albert Sickmann, Lennart Martens, and Harald Barsnes. Peptideshaker enables reanalysis of ms-derived proteomics data sets. *Nature biotechnology*, 33(1):22–24, 2015.
- [5] Chengxin Dai, Julianus Pfeuffer, Hong Wang, Ping Zheng, Lukas Käll, Timo Sachsenberg, Vadim Demichev, Mingze Bai, Oliver Kohlbacher, and Yasset Perez-Riverol. quantms: a cloud-based pipeline for quantitative proteomics enables the reanalysis of public proteomics data. *Nature Methods*, July 2024.

- [6] Paolo Di Tommaso, Maria Chatzou, Evan W Floden, Pablo Prieto Barja, Emilio Palumbo, and Cedric Notredame. Nextflow enables reproducible computational workflows. *Nature Biotechnology*, 35(4):316–319, April 2017.
- [7] Chengxin Dai, Anja Füllgrabe, Julianus Pfeuffer, Elizaveta M. Solovyeva, Jingwen Deng, Pablo Moreno, Selvakumar Kamatchinathan, Deepti Jaiswal Kundu, Nancy George, Silvie Fexova, Björn Grüning, Melanie Christine Föll, Johannes Griss, Marc Vaudel, Enrique Audain, Marie Locard-Paulet, Michael Turewicz, Martin Eisenacher, Julian Uszkoreit, Tim Van Den Bossche, Veit Schwämmle, Henry Webel, Stefan Schulze, David Bouyssié, Savita Jayaram, Vinay Kumar Duggineni, Patroklos Samaras, Mathias Wilhelm, Meena Choi, Mingxun Wang, Oliver Kohlbacher, Alvis Brazma, Irene Papathodorou, Nuno Bandeira, Eric W. Deutsch, Juan Antonio Vizcaíno, Mingze Bai, Timo Sachsenberg, Lev I. Levitsky, and Yasset Perez-Riverol. A proteomics sample metadata representation for multiomics integration and big data analysis. *Nature Communications*, 12(1):5854, October 2021.
- [8] Lennart Martens, Matthew Chambers, Marc Sturm, Darren Kessner, Fredrik Levan- der, Jim Shoffstahl, Wilfred H. Tang, Andreas Römpf, Steffen Neumann, Angel D. Pizarro, Luisa Montecchi-Palazzi, Natalie Tasman, Mike Coleman, Florian Reisinger, Puneet Souda, Henning Hermjakob, Pierre-Alain Binz, and Eric W. Deutsch. mzML—a Community Standard for Mass Spectrometry Data. *Molecular & Cellular Proteomics*, 10(1):R110.000133, January 2011.
- [9] Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino Da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J.G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A.C. 'T Hoen, Rob Hooft, Tobias Kuhn, Ruben

Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene Van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan Van Der Lei, Erik Van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1):160018, March 2016.

- [10] Sven Degroeve and Lennart Martens. MS2PIP: a tool for MS/MS peak intensity prediction. *Bioinformatics (Oxford, England)*, 29(24):3199–3203, December 2013.
- [11] Robbin Bouwmeester, Ralf Gabriels, Niels Hulstaert, Lennart Martens, and Sven Degroeve. DeepLC can predict retention times for peptides that carry as-yet unseen modifications. *Nature Methods*, 18(11):1363–1369, November 2021.
- [12] Siegfried Gessulat, Tobias Schmidt, Daniel Paul Zolg, Patroklos Samaras, Karsten Schnatbaum, Johannes Zerweck, Tobias Knaute, Julia Rechenberger, Bernard Delanghe, Andreas Huhmer, Ulf Reimer, Hans-Christian Ehrlich, Stephan Aiche, Bernhard Kuster, and Mathias Wilhelm. Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning. *Nature Methods*, 16(6):509–518, June 2019.
- [13] Louise M. Buur, Arthur Declercq, Marina Strobl, Robbin Bouwmeester, Sven Degroeve, Lennart Martens, Viktoria Dorfer, and Ralf Gabriels. MS² Rescore 3.0 Is a Modular, Flexible, and User-Friendly Platform to Boost Peptide Identifications, as Showcased with MS Amanda 3.0. *Journal of Proteome Research*, 23(8):3200–3207, August 2024.
- [14] Wenjie Yang, Hualin Chen, Lin Ma, Jie Dong, Mengchao Wei, Xiaoqiang Xue, Yingjie Li, Zhaoheng Jin, Weifeng Xu, and Zhigang Ji. A comprehensive analysis of the FOX family for predicting kidney renal clear cell carcinoma prognosis and the oncogenic role of FOXG1. *Aging*, 14(24):10107–10124, December 2022.

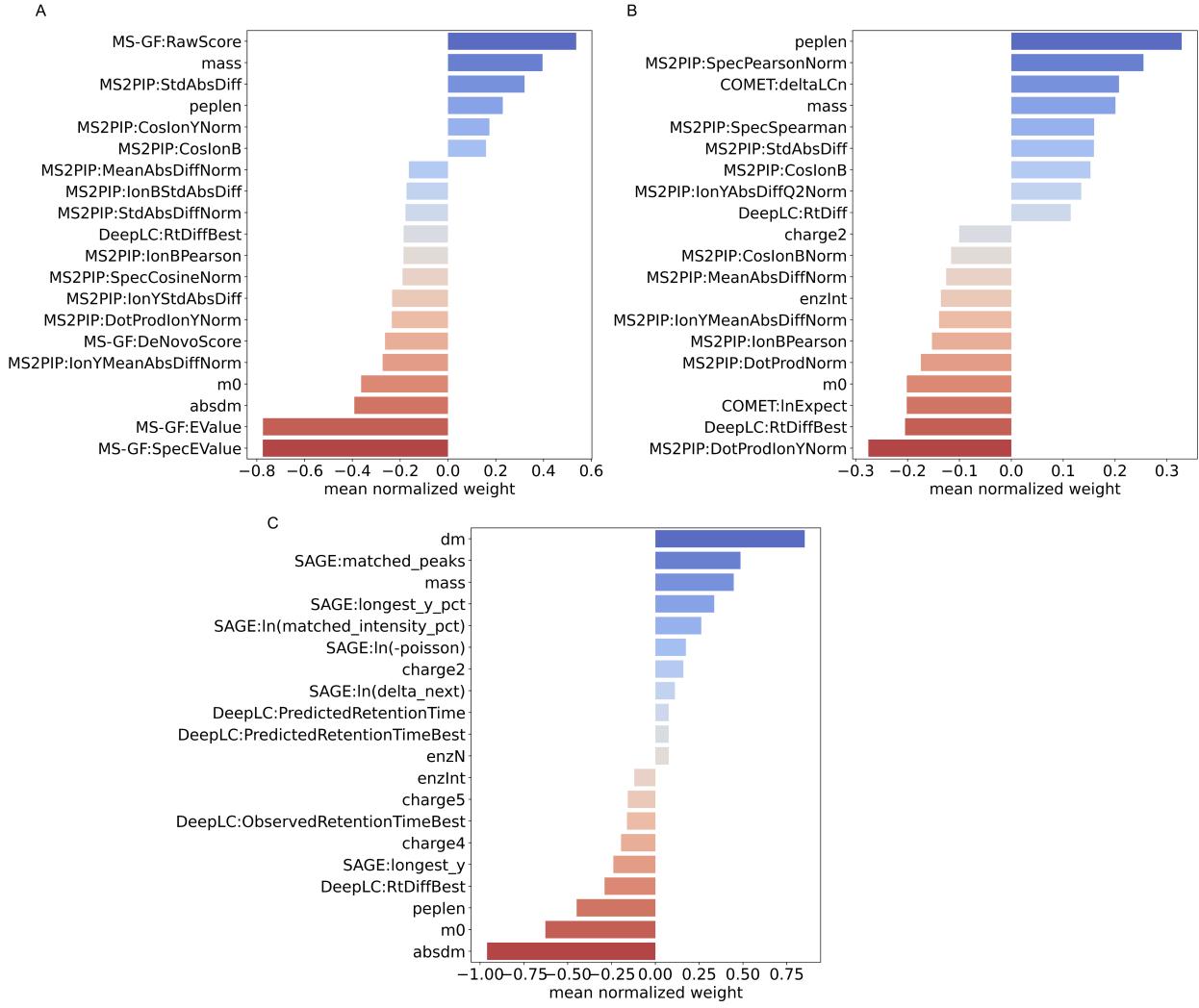


Figure S1: The weights that Percolator assigns to different features of the search engine feature vector in MSGF+ (A), Comet (B), and SAGE (C) combined with Percolator settings from PXD001819. The more different from zero a weight is (both positive and negative), the more important that feature is in the final Percolator classification.

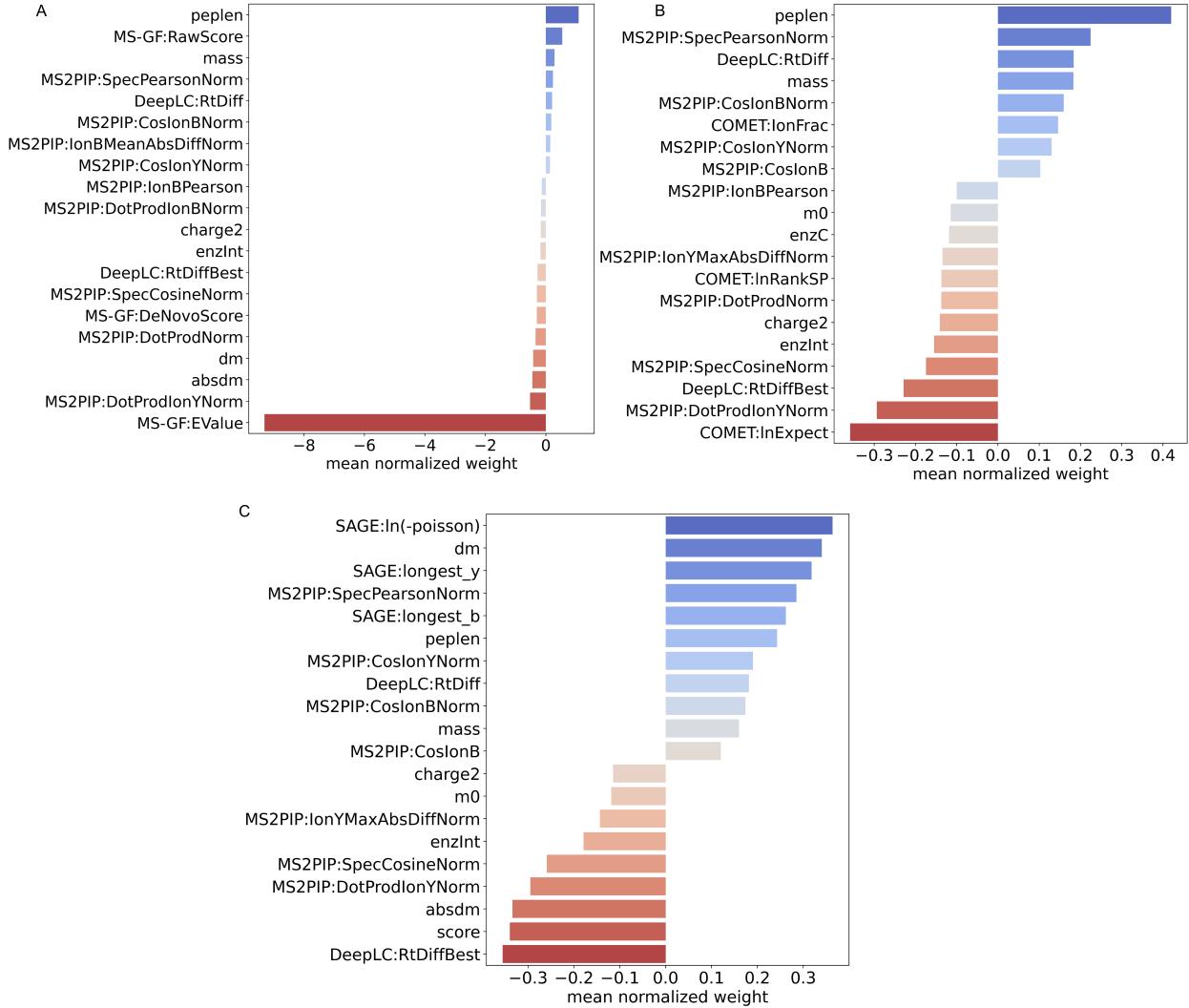


Figure S2: The top 20 normalized weights from Percolator for MSGF+ (A), Comet (B), and SAGE (C) identification results from PDC000125.

Table S1: Runtime, memory, and CPU consumption for different configurations. The memory and CPU values represent the maximum resources required for a single job. The quantms workflow was run in a 64 cluster queue.

Datasets	Configuration	Runtime	Memory	CPU
PDC000127	Comet&MSGF+&MS ² Rescore	9h39m	30G	7
PDC000127	Comet&MSGF+	6h28m	8G	4
PDC000127	Comet	5h10m	8G	4
PXD019643	Comet&MSGF+&MS ² Rescore	15h57m	20G	8
PXD019643	Comet&MSGF+	13h09m	5.5G	4
PXD019643	Comet	2h05m	5.5G	4
PXD026824	Comet&MSGF+&MS ² Rescore	2h05m	66G	8
PXD026824	Comet&MSGF+	1h17m	48.4G	4
PXD026824	Comet	24m	48.4G	4

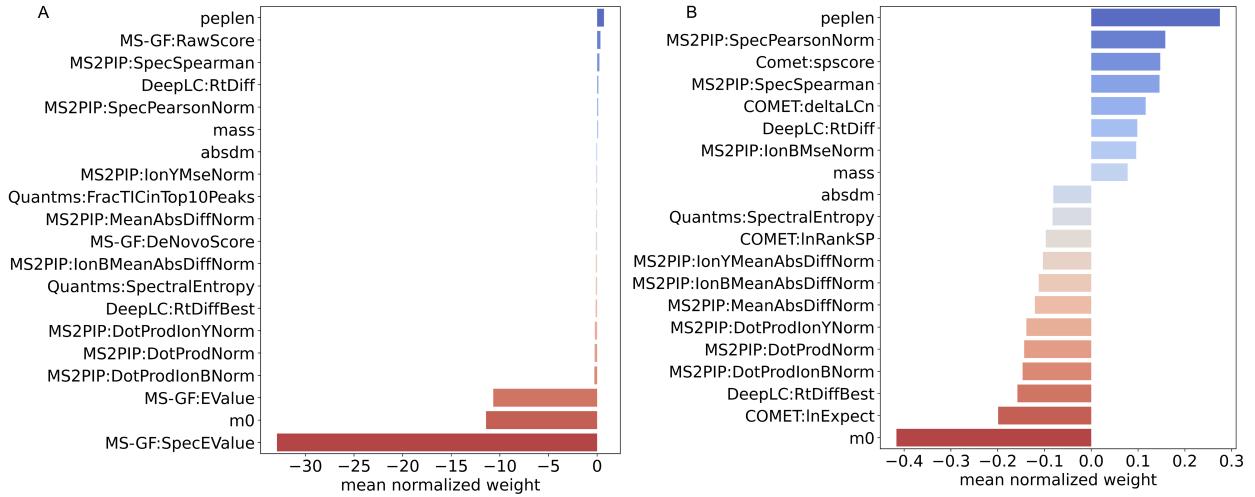


Figure S3: The top 20 normalized weights from Percolator for MSGF+ (A) and Comet (B) identification results in the PXD019643 dataset.

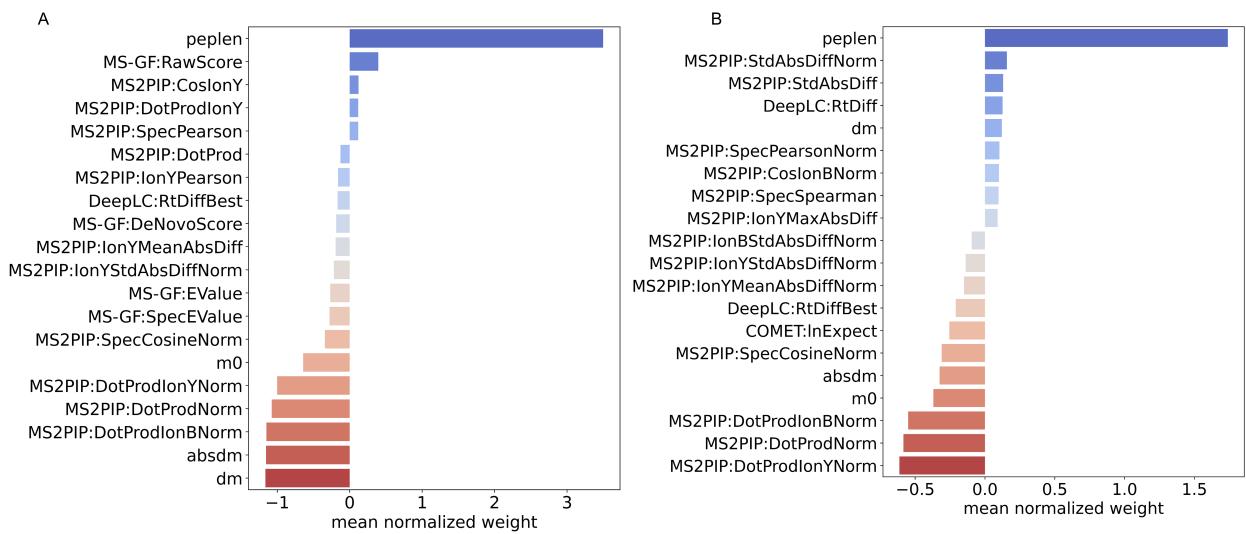


Figure S4: The top 20 normalized weights from Percolator for MSGF+ (A) and Comet (B) identification results in the PXD026824 dataset.