# Triqler for Data Independent Aquisition Data

Patrick Truong Matthew The Lukas Käll

October 19, 2021

**Abstract**

Within mass spectrometry-based proteomics, protein summarization and quantification is recognized as a complex problem. The detection and quantification of each proteoform's protolytic peptides is an error-prone process, and there is a need for computational methods to assess errors and determine which measurments that can be trusted or not. We have previously designed a integrative model, Triqler, that combines identification and quantification errors and summarize results into protein quantities. Here we show that Triqler, is well compatible with data-independent acquisition data, despite being designed for data-dependent acquisition data. Furthermore, we find that it has better performance than other protein summarization tools, when evaluating a relatively large set of different DIA processing methods.

## Introduction

Label-free quantification (LFQ) using Mass spectrometry (MS)-based proteomics enables efficent differential concentration determination of proteins in complex mixtures. The processing of data from such experiments requires multiple different steps, all subjects to errors. We have previously designed a hierarchical Bayesian model, Triqler, able to control for errors from both the identification and quantification process in such experiments [26]. By integrating the errors probabilities from identification and quantification one can obtain better accuracy in calling differentially abundant proteins [26].

Triqler was designed for handling LFQ data from Data-dependent acquisition (DDA). However, many labs prefer Data-independent acquisition (DIA) mass spectrometry [29] as they find that it gives more reproducible peptide detection, and allow for a broader dynamical range in quantification [4,32], compared to DDA. Here, we set out to investigate if Triqler is a suitable tool for protein summarization in DIA data.

For the analysis of DIA data, just as any other type of analysis of complex data, the individual computational methods for all the steps in the processing affect the final results. It is therefore quite hard to evaluate the influence of the individual tools, which should not stop the field from trying to establish what the features of the different processing steps [9, 11, 17]. Unbiased comparisons of software tools is challenging for several reasons [9]. Methods can be assessed by scientist lacking relevant expertise, the tested methods may be lacking sufficient documentation and the interpretation of test results may be subjective [30] [15] [18] [2]. By using the same data set we can assure that the data set is processed consistently and further the analysis by extending it to protein summarization procedures.

Here, we found a previous data set most useful for the comparison of methods, the LFQBench from Navarro et al. [17] constructed a high-quality data set for five widely used tools for analysis of SWATH-MS data (four peptide-centric query tools: OpenSWATH [21], SWATH 2.0, Skyline [16], Spectronaut [5], and one data-centric approach DIA-Umpire [16]), which is readily amendable for a benchmark of other methods.

In the original study, Navarro et al. included a benchmark of different protein summarization strategies, and found that the so-called TOP3 method generally resulted in lower variance and better quantification accuracy than the built-in methods from the tested methods [17]. However, there are reasons to believe that more sophisticated methods would yield better protein quantification than TOP3-method. Simple summarization methods based on mean and median peptide intensity has been shown to produce unreliable protein abundance estimates [12], and more advanced summarization strategies for LFQ data has been

proposed in literature [7,23], and summarization techniques such as PQPQ [10], msStats [6], Diffacto [31], MSqRobSum [24] and Triqler [26] has all been shown to outperform TOP3 and there currently exists no reason why said methods would not theoretically be able to perform well for DIA-data. Hence, we set out to benchmark Triqler as well as a set of other protein summarisation techniques on the LFQBench dataset.

# Materials and methods

## Data description

### Mass spectrometry data

We downloaded the LFQBench dataset [17] from PRIDE identifier PXD002952. Here we used the TripleTOF6600 (ABSciex) section of the study, which was harvested with a setup of 32 fixed windows ms2-windows. We also restricted ourselves to the low ratio difference samples, reffered to as the HYE124 hybrid proteome samples in the original study. That consists of triplicates of Sample A composed of tryptically digested proteins from 65% w/w HeLa, 30% w/w yeast, and 5% w/w E. coli cells, and triplicates of Sample B, composed of 65% w/w, 15% w/w yeast, and 20% w/w E. coli proteins. Samples from HYE110 and the TripleTOF5600 section of HYE124 was omitted in this study. Further details about mass spectrometric instrumentation and data acquisition are available in Navarro et al. [17]. The `.wiff` files were converted to `.mzML` files in a centroided format using msconvert (using windows OS msconvert version 3.0) with peakPicking filter msLevel=1-).

### Sequence database

Uniprot FASTA files with one protein sequence per gene was downloaded for each species (UP000005640, UP000000625, and UP000002311, acquired on 2021-06-16). The unfiltered FASTA files contained 20 590 human proteins, 6 046 yeast proteins, and 4 373 E. coli proteins. To reduce for the effect of the different protein inference strategies for the tested protein summarization tools, a modified .fasta file, without shared peptides, was used for database search. The filter randomly removed protein sequences with shared peptides, so that the final database did not contain any tryptic peptides with length >7 amino acids mapping to peptides shared with other proteins. After filtering the FASTA file contained 20 302 proteins (288 human proteins less proteins than unfiltered database), 5 848 yeast proteins (198 yeast proteins less proteins than unfiltered database), and 4 306 E. Coli proteins (67 E. Coli proteins than unfiltered database). Replacing the I/L amino acids with to handle mass-equivalence did not result in any considerable differences (See Supplementary Table S1). We also added pseudo-reverse sequences to the database as decoys for target-decoy analysis using OpenSwathDecoyGenerator.

## General workflow

Just as in LFQBench, we used two separate strategies to generate peptide abundances from the DIA-runs, first we used a spectral library consisting of selected spectra from separate DDA runs, we refer to this workflow as SL hereon, and second we searched pseudo-spectra generated directly from the DIA data, we will refer to this workflow as PS from hereon. The workflows are shown in figure 1. We will describe the parameter chioces of both theses two methods below.

### Spectral library matching (SL)

For the Spectral Library (SL) workflow we searched the LFQBench provided DDA runs with MSFragger [13] with a precursor mass tolerance of [-20, 20] ppms, fragment tolerance of 20 ppms, and constructed a spectral library with EasyPQP [1] to from the DDA-search results using a PSM-level FDR treshold of 0.01, peptide-level FDR treshold of 0.01 and protein-level FDR treshold of 0.01 (default settings). Decoys for the spectral library was generated with OpenSwathDecoyGenerator with pseudo-reverse method. The spectral library was subsequently matched to the DIA data with the OpenSwath Workflow and PyProphet [25] was used for

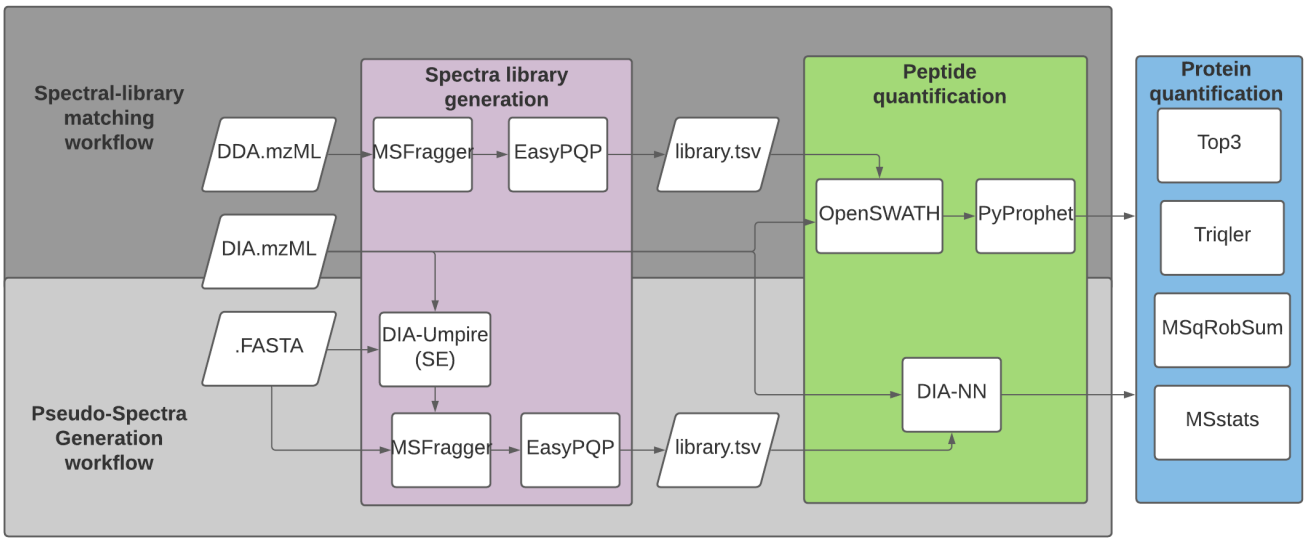


Figure 1: **The spectrum-library matching and pseudo spectrum workflow pipelines.**

computing false discovery rate (m_score) for the peptide quantification. This resulted in a set of detected peptides together with their assesed peptide identification accuracies and abundance estimates.

### Pseudo Spectrum generation (PS)

For the PS workflow we also used the fragpipe software, employing DIA-Umpire to extract pseudo-spectra from the DIA data. The pseudo-spectra were subsequently searched using using MSFragger with a precursor mass tolerance of -20, 20] ppms, fragment tolerance of 20 ppms, and allowing for [M,[^] variable modifications (M - oxidation on methionine and ^ is a terminus modifier for protein N-terminal). A spectral library was build from the resulting PSMs at a PSM-level treshold of 0.01 using easyPQP. DIA-NN was used for peptide quantifications, a method that uses a built-in custom implementation of the mProphet algorithm to compute $q$ values [8, 20].

This is confusing. Spell out what it means. Do you look for N-term methionines, or is it something different?

This is a very libral treshold. Is this a default level?

## Protein summarization

The peptides was summarized to proteins using Top3, MSstats, MSqRobSum and Triqler for both spectral library and pseudo-spectra pipelines. The multiple test correction is performed with $q$ value for Triqler and Top3, while Benjamini-Hochberger [3] is used for MsStats and MSqRobSum. The $q$ value approach aims to give a unbiased estimator of FDR, while the Benjamini-Hochberger approach guarantees that the reported FDR is less than the true FDR. The consequence of this is that even though the methods are interchangeable, the Benjamini-Hochberger approach has slightly less statistical power and be slightly more conservative [14].

# Results

| OSW | | | | | | |
|---|---|---|---|---|---|---|
| Condition | | 1 | | | 2 | |
| Run | 002-Pedro | 004-Pedro | 006-Pedro | 003-Pedro | 005-Pedro | 007-Pedro |
| Peptides | 20 747 | 21 445 | 21 016 | 20 494 | 22 792 | 22 787 |
| Proteins | 2 836 | 2 857 | 2 869 | 2 818 | 2 909 | 2 918 |

Table 1: **Number of identified peptides and proteins for the samples using spectral library workflow.**

| DIANN | | | | | | |
|---|---|---|---|---|---|---|
| Condition | | 1 | | | 2 | |
| Run | 002-Pedro | 004-Pedro | 006-Pedro | 003-Pedro | 005-Pedro | 007-Pedro |
| Peptides | 21 036 | 20 783 | 21 040 | 21 243 | 21 325 | 21 248 |
| Proteins | 3 377 | 3 350 | 3 368 | 3 358 | 3 381 | 3 356 |

Table 2: **Number of identified peptides and proteins for the samples using the pseudo-spectra workflow.**

## Test of assumptions

We strived to validate the Triqler method for a DIA setting. We hence investigatred some properties of DIA data to see if they agree with the assumption we mnade when designing Triqler for DDA data. DIA data is known to encompass a larger dynamic range than DDA data . This could affect one of Triqles assumptions, that the noise structure is mainly multiplicative, i.e. that the standard deviation within a sample group is proportional to its mean. When investigating all the peptide abundance measurements at a 1% identification FDR from the TripleTOF6600 section of the LFQBench dataset, we found a relatively linear relation between standard deviation and mean (Figure 2A). Further, Triqler assumes that the missing peptide abundance values follow a censored normal distribution, which is also roughly fullfilled by DIA data (See Figure 2B).

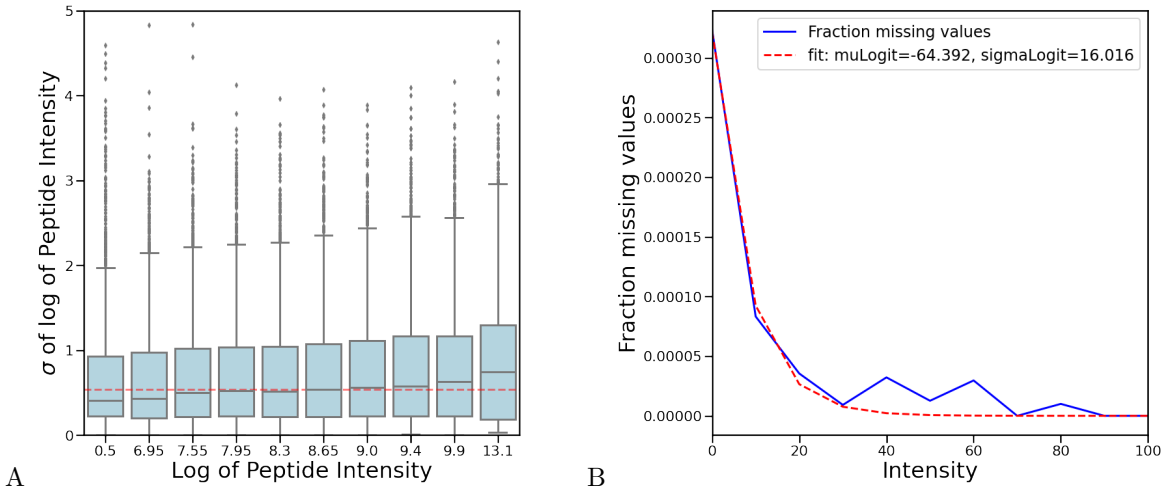


Figure 2: **Triqler's assumptions of peptide abundance values are valid also for the petide intensities in DIA data.** The red dotted line marks the median of the standard deviation. (A) We grouped the peptide intensities to equally sized bins and plotted the gaussian kernal density estimate for each bin. The bins are approximately the same shape, demonstrating that $\log(\sigma) \approx \log(\mu) + \log(k)$ and hence $\sigma \approx \mu k$.(B) We plotted the fraction of peptides observing one missing value for each triplicate in the set. We obseve that this empirical distribution roughly follows the assumed function $1 - (\frac{1}{2} + \frac{1}{2}tanh(\frac{x-\mu_m}{\sigma_m}))$.(See Supplementary Figure S2 and Supplementary Figure S3).

## Test of performance

We wanted to compare the performance of Triqler against other protein summarization methods. An inherent problem when comparing different protein summarization softwares is that the performance is affected by which protein inference structure is used, in a data set dependent manner [22]. For example, when reporting number of differentially abundant proteins in protein mixtures, a protein inference scheme that infer any protein contain a detected peptide will report more differentially abundant proteins than more restrictive scheme that just report a parsimonous set of proteins when benchmarking with mixtures of cell lysate mixtures data sets. For such data sets there is no mechanisms restrictive mechanism detecting situations where non-present proteoforms are reported as long as they are part of and reported with protein abundance rates compatible to the right proteome. To elaviate, or at least minimize, this problem from our comparison, we restricted the searched FASTA files by removing proteins with shared peptides. This should give a fair comparison of protein summarization regardless of the protein inference method.

find the prefered term for "cell lysate mixture"

Also, we selected to compare the efficency of Triqler to the ones of the protein summarization strategies from msStats, msSqRobSum and Top-3 using two different processing strategies, spectral libraries using DDA data, and when generating pseudo-spectra with DIA-Umpire [28]. Here we selected to run Triqler with a lower bound estimate as described in The&Käll [27], which ended up as 0.76 for the spectral library data, and 0.52 for the pseudo-spectra enabled data.

### Fold change distributions

To get an overview of the results of the methods we first made histograms of the reported protein level fold-changes as reported by the compared methods in Figure 3. An FDR filter is not applied on the histograms. We observe that Triqler and Top3 has less unbias than MSstats and MSqRobSum by seeing that the apex of the distributions are centered more closely to the true values. MSqRobSum report a very large number of proteins at zero log2 fold-change. These peaks are insignificant peaks.

FIXME: Remove this, Figure **??**.

### Differential abundance

We also report the amount of differentially abundant proteins in Figure 4 to show the performance of of the compared methods. We compare two-sided differential abundance since Triqler computes two-sided fold change. A fold-change treshold of 0.4 has been applied to Triqler, but not Top3, MSstats or MSqRobSum. This should give an disadvantage to Triqler. We observe that Triqler and MSqRobSum are the best performing protein quantification methods for *E. Coli* and Yeast in both spectral library matching and pseudo-spectrum workflows. We observe that MSqRobSum and MSstats report far more differentially abundant human proteins than Triqler and Top3. These are false hits.

### Comparison of ability to differentiate differentially abundant proteins

The LFQBench set contain varying concentrations of *E. Coli* and Yeast concentrations in a background of HeLa-cells. As a first test of performance we compared the methods'ability to infer differentially abundant *E. Coli* and Yeast protein as a function of the number of false positives from the HeLa-background, see Figure 5. Overall it seems like Triqler reports more true differential abundat protein for every false differential abundant protein for both spectral library matching and pseudo-spectrum methods. Surprisingly, Top3 performs has more true differentially abundant proteins for every false protein than MSstat and MSqRobSum.

### Comparison of statistical calibration

We subsequently set out to test the statistical calibration of the different summarization methods. We hence investigated the relation between the fraction of wrongly reported differential abundant proteins (i.e. the number of human proteins), and the estimated false discovery rate (See Figure 6). We observe that Triqler and Top3 shows a fraction of human protein that is close to the $q$ value, while MSstats and MSqRobSum seem to have a larger amount of wrongly identified differentially abundant protein than the reported FDR.

## Discussion

In bottom-up proteomics summarizing peptides to protein has been a long-established practice. There are several reasons protein summarization is beneficial. It gives lower variance among technical and biological replicates than peptide-level analysis, it reduces the amount of hypotheses tested and reduced the amount of missing values, which can have major impact on the quality of the analysis [19].

Here we have shown that Triqler operates well for DIA data, despite originally intended for DDA data. We also find that Triqler outperforms other protein summarization methods on our engineered benchmark set, both in terms of sensitivity and accuracy in its error estimates.

We do see some differences in how DIA and DDA ppetide-level abundance data appear. For instance there are more missing values in DDA than DIA data. However, we find that qualitatively the types of data are similar, at least in the sense that Triqler's underlying assumption of missing values seem to be as valid for DDA and DIA data.

Key notes for the discussion:

- Summarize our findings from above.
- Discuss the protein inference problem.
- Tell how we have performed our protein search on a truncated database for fair comparison.
- Highlight fair comparison.
- Discuss future improvements.
- Future improvement: Protein inference strategies which can account for multiple proteoforms.

- Discuss why dataset such as LFQ-bench is good (because it makes it easy to perform benchmarking and comparison).
- Add some more general notes about the field in general.

## Acknowledgements

## Funding

This work was supported by grants from the Swedish Research Council (grant 2017-04030)

## Supporting information

## References

[1] EasyPQP:simple library generation for openswath. `https://github.com/grosenberger/easypqp`. Accessed: 2021-09-03.

[2] The difficulty of a fair comparison. *Nature Methods*, 12(273), 2015.

[3] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300, 1995.

[4] Marshall Bern, Gregory Finney, Michael R Hoopmann, Gennifer Merrihew, Michael J Toth, and Michael J MacCoss. Deconvolution of mixture spectra from ion-trap data-independent-acquisition tandem mass spectrometry. *Analytical chemistry*, 82(3):833–841, 2010.

[5] Roland Bruderer, Oliver M Bernhardt, Tejas Gandhi, Saša M Miladinović, Lin-Yang Cheng, Simon Messner, Tobias Ehrenberger, Vito Zanotelli, Yulia Butscheid, Claudia Escher, et al. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues*[s]. *Molecular & Cellular Proteomics*, 14(5):1400–1410, 2015.

[6] Meena Choi, Ching-Yun Chang, Timothy Clough, Daniel Broudy, Trevor Killeen, Brendan MacLean, and Olga Vitek. Msstats: an r package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics*, 30(17):2524–2526, 2014.

[7] Jürgen Cox, Marco Y Hein, Christian A Luber, Igor Paron, Nagarjuna Nagaraj, and Matthias Mann. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed maxlfq. *Molecular & cellular proteomics*, 13(9):2513–2526, 2014.

[8] Vadim Demichev, Christoph B Messner, Spyros I Vernardis, Kathryn S Lilley, and Markus Ralser. Diann: neural networks and interference correction enable deep proteome coverage in high throughput. *Nature methods*, 17(1):41–44, 2020.

[9] Craig Dufresne, David Hawke, Alexander R Ivanov, Antonius Koller, Brendan MacLean, Brett Phinney, Kristie Rose, Paul Rudnick, Brian Searle, Scott Shaffer, et al. Abrf research group development and characterization of a proteomics normalization standard consisting of 1,000 stable isotope labeled peptides. *Journal of Biomolecular Techniques: JBT*, 25(Suppl):S1, 2014.

[10] Jenny Forshed, Henrik J Johansson, Maria Pernemalm, Rui MM Branca, AnnSofi Sandberg, and Janne Lehtiö. Enhanced information output from shotgun proteomics data by protein quantification and peptide quality control (pqpq). *Molecular & Cellular Proteomics*, 10(10), 2011.

[11] Laurent Gatto, Kasper D Hansen, Michael R Hoopmann, Henning Hermjakob, Oliver Kohlbacher, and Andreas Beyer. Testing and validation of computational methods for mass spectrometry. *Journal of proteome research*, 15(3):809–814, 2016.

[12] Ludger JE Goeminne, Andrea Argentini, Lennart Martens, and Lieven Clement. Summarization vs peptide-based models in label-free quantitative proteomics: performance, pitfalls, and data analysis guidelines. *Journal of proteome research*, 14(6):2457–2465, 2015.

[13] Andy T Kong, Felipe V Leprevost, Dmitry M Avtonomov, Dattatreya Mellacheruvu, and Alexey I Nesvizhskii. Msfragger: ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics. *Nature methods*, 14(5):513–520, 2017.

[14] Keegan Korthauer, Patrick K Kimes, Claire Duvallet, Alejandro Reyes, Ayshwarya Subramanian, Mingxiang Teng, Chinmay Shukla, Eric J Alm, and Stephanie C Hicks. A practical guide to methods controlling false discoveries in computational biology. *Genome biology*, 20(1):1–21, 2019.

[15] Felipe da Veiga Leprevost, Valmir C Barbosa, Eduardo L Francisco, Yasset Perez-Riverol, and Paulo C Carvalho. On best practices in the development of bioinformatics software. *Frontiers in genetics*, 5:199, 2014.

[16] Brendan MacLean, Daniela M Tomazela, Nicholas Shulman, Matthew Chambers, Gregory L Finney, Barbara Frewen, Randall Kern, David L Tabb, Daniel C Liebler, and Michael J MacCoss. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*, 26(7):966–968, 2010.

[17] Pedro Navarro, Jörg Kuharev, Ludovic C Gillet, Oliver M Bernhardt, Brendan MacLean, Hannes L Röst, Stephen A Tate, Chih-Chiang Tsou, Lukas Reiter, Ute Distler, et al. A multicenter study benchmarks software tools for label-free proteome quantification. *Nature biotechnology*, 34(11):1130–1136, 2016.

[18] Huisong Pak, Frederic Nikitin, Florent Gluck, Frederique Lisacek, Alexander Scherl, and Markus Muller. Clustering and filtering tandem mass spectra acquired in data-independent mode. *Journal of The American Society for Mass Spectrometry*, 24(12):1862–1871, 2013.

[19] Deanna L Plubell, Lukas Käll, Bobbie-Jo Webb-Robertson, Lisa Bramer, Ashley Ives, Neil L Kelleher, Lloyd M Smith, Thomas J Montine, Christine C Wu, and Michael J MacCoss. Can we put humpty dumpty back together again? what does protein quantification mean in bottom-up proteomics? *bioRxiv*, 2021.

[20] Lukas Reiter, Oliver Rinner, Paola Picotti, Ruth Hüttenhain, Martin Beck, Mi-Youn Brusniak, Michael O Hengartner, and Ruedi Aebersold. mprophet: automated data processing and statistical validation for large-scale srm experiments. *Nature methods*, 8(5):430–435, 2011.

[21] Hannes L Röst, George Rosenberger, Pedro Navarro, Ludovic Gillet, Saša M Miladinović, Olga T Schubert, Witold Wolski, Ben C Collins, Johan Malmström, Lars Malmström, et al. Openswath enables automated, targeted analysis of data-independent acquisition ms data. *Nature biotechnology*, 32(3):219–223, 2014.

[22] Oliver Serang, Luminita Moruz, Michael R Hoopmann, and Lukas Käll. Recognizing uncertainty increases robustness and reproducibility of mass spectrometry-based protein inferences. *Journal of proteome research*, 11(12):5586–5591, 2012.

[23] Jeffrey C Silva, Marc V Gorenstein, Guo-Zhong Li, Johannes PC Vissers, and Scott J Geromanos. Absolute quantification of proteins by lcmse: A virtue of parallel ms acquisition* s. *Molecular & Cellular Proteomics*, 5(1):144–156, 2006.

[24] Adriaan Sticker, Ludger Goeminne, Lennart Martens, and Lieven Clement. Robust summarization and inference in proteome-wide label-free quantification. *Molecular & Cellular Proteomics*, 19(7):1209–1219, 2020.

[25] Johan Teleman, Hannes L Röst, George Rosenberger, Uwe Schmitt, Lars Malmström, Johan Malmström, and Fredrik Levander. Diana—algorithmic improvements for analysis of data-independent acquisition ms data. *Bioinformatics*, 31(4):555–562, 2015.

[26] M. The and L. Käll. Integrated identification and quantification error probabilities for shotgun proteomics. *bioRxiv*, 2018.

[27] Matthew The and Lukas Käll. Triqler for maxquant: Enhancing results from maxquant by bayesian error propagation and integration. *Journal of proteome research*, 20(4):2062–2068, 2021.

[28] Chih-Chiang Tsou, Dmitry Avtonomov, Brett Larsen, Monika Tucholska, Hyungwon Choi, Anne-Claude Gingras, and Alexey I Nesvizhskii. Dia-umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nature methods*, 12(3):258–264, 2015.

[29] John D Venable, Meng-Qiu Dong, James Wohlschlegel, Andrew Dillin, and John R Yates. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nature methods*, 1(1):39–45, 2004.

[30] John R Yates, Sung Kyu Robin Park, Claire M Delahunty, Tao Xu, Jeffrey N Savas, Daniel Cociorva, and Paulo Costa Carvalho. Toward objective evaluation of proteomic algorithms. *Nature methods*, 9(5):455–456, 2012.

[31] Bo Zhang, Mohammad Pirmoradian, Roman Zubarev, and Lukas Käll. Covariation of peptide abundances accurately reflects protein concentration differences. *Molecular & Cellular Proteomics*, 16(5):936–948, 2017.

[32] F. Zhang, W. Ge, G. Ruan, Cai X., and T. Guo. Data-independent acquisition mass spectrometry-based proteomics and software tools: A glimpse in 2020. *Proteomics*, 20(17-18), 2020.

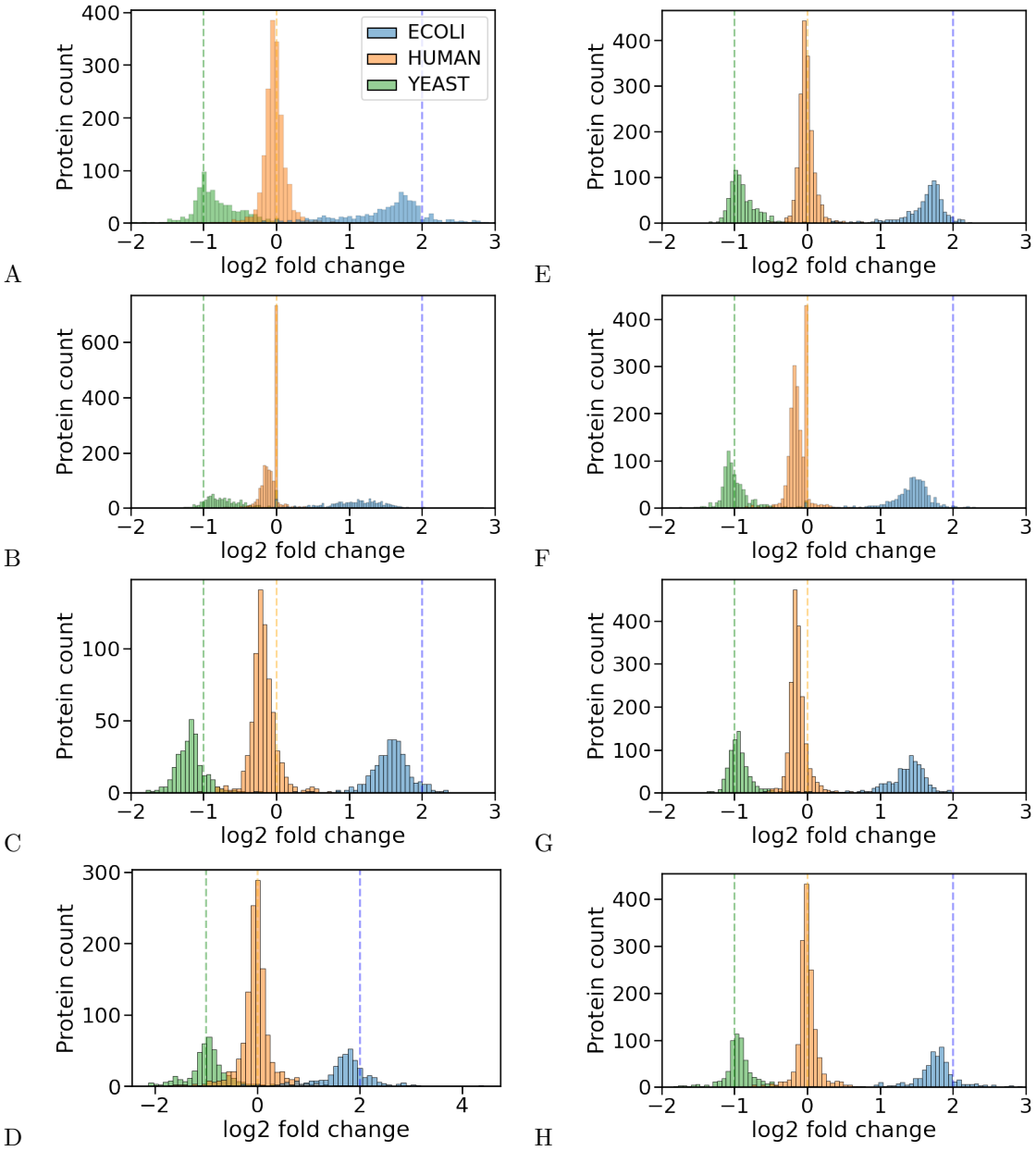


Figure 3: **Comparison of reported fold change distributions.** We used peptide data from (A-D) Spectrum library matching and (E-H) pseudo spectra workflow as generated by (A,E) Triqler, (B,F) MSqRobSum, (C,G) MSstats, and (D,H) Top3.

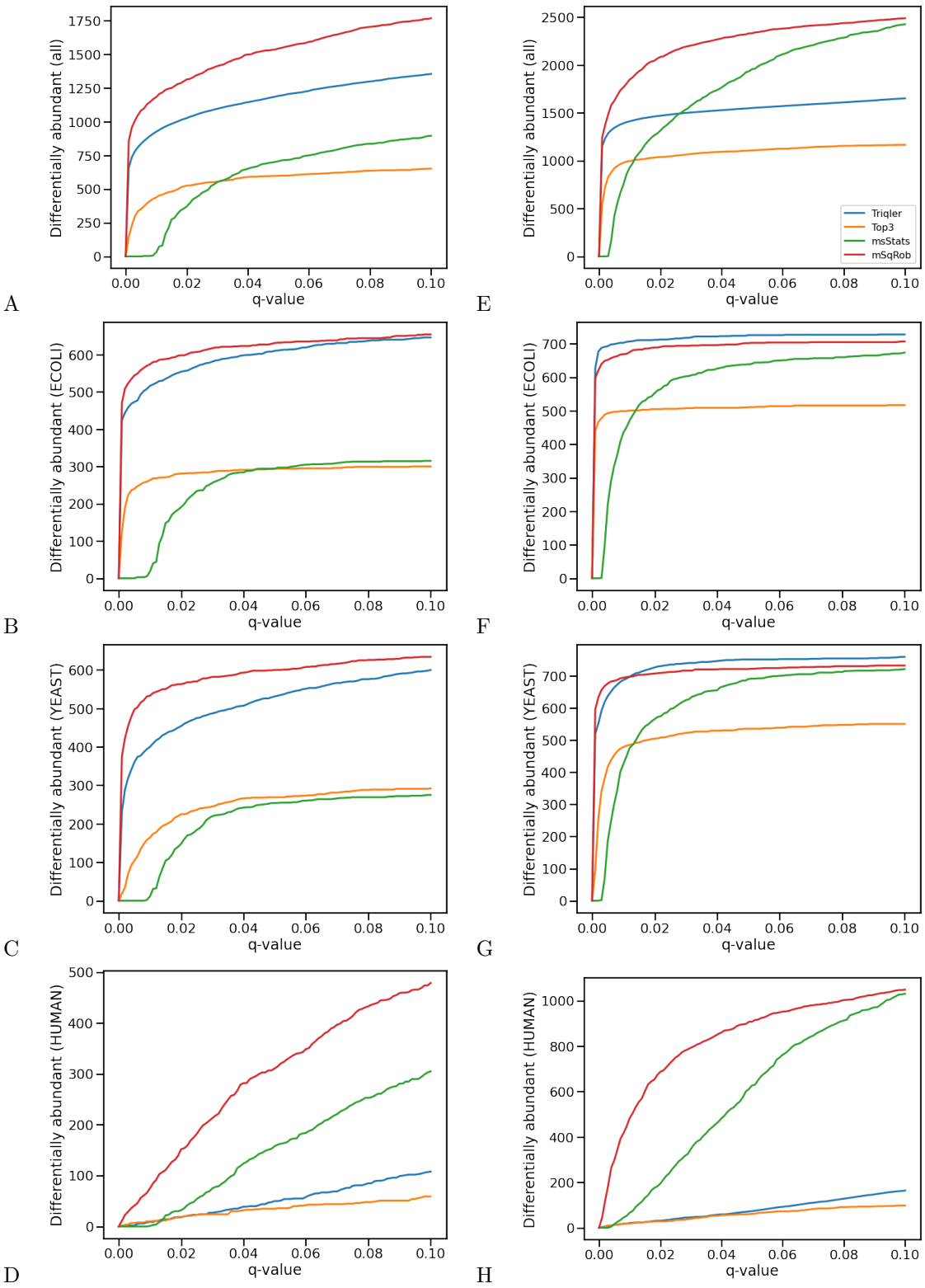


Figure 4: **Comparison of reported differential abundance.** Differential abundance is reported for (A-D) Spectrum library matching and (E-H) pseudo spectra workflowsf for each proteome (A,E) All, (B,F) *E. Coli*, (C,G) Yeast, and (D,H) Human.

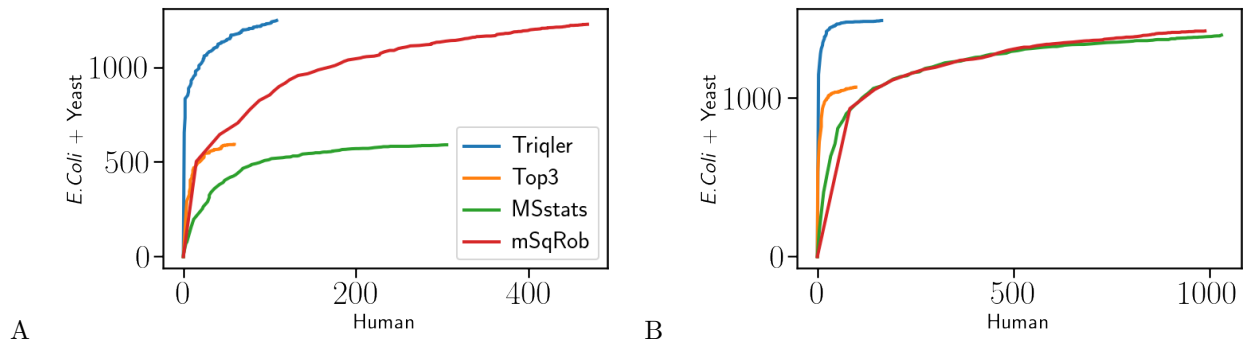


Figure 5: **Comparison of ability to differentiate differentially abundant proteins** We plotted the number of reported differentially abundant *E. Coli* and Yeast proteins as a function of number of proteins from the HeLa background when sorting according to significance for (A) DDA generated spectral libraries and (B) DIA-Umpire geneated Pseudo spectra. For the test we selected a fold-change treshold of 0.4 for Triqler, because it is the rounded lower bound for fold-change computed by Triqler (Supplement Figure S1 shows differential abundance for each specie).

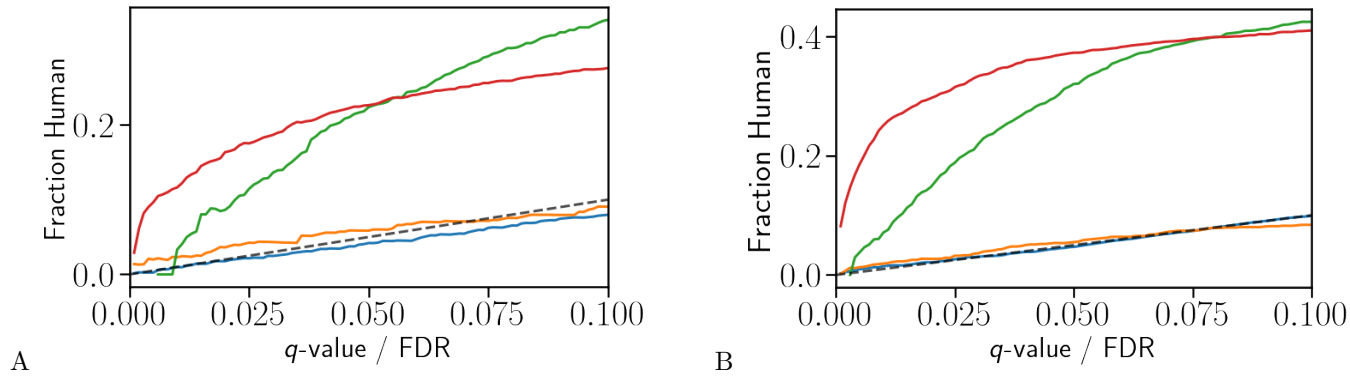


Figure 6: **Comparison of calibration of the compared summarization methods.** We plotted the fraction of reported differentially abundant HeLa proteins as a function of $q$ value treshhold for (A) DDA generated spectral libraries and (B) DIA-Umpire geneated Pseudo spectra. The multiple test correction is performed with $q$ value for Triqler and Top3, while Benjamini-Hochberger [Add citation] is used for MSstats and MSqRobSum.