HUPO 2022, PP02.045

LFQ Workflows for Quantitative Results on the PaSER Platform

Christopher Adams[2], **Sven Brehmer**[1], **Robin Park**[2], **Matt Willetts**[2], **Tharan Srikumar**[2], **Jonathan Krieger**[2]
[1]Bruker Daltonik GmbH, Fahrenheitstraße 4, 28359 Bremen, Germany [2]Bruker Daltonics Inc., 40 Manning Road, Manning Park, Billerica, MA 01821, USA

Introduction

Data dependent acquisition (DDA) where peptide precursors are chosen based primarily on their intensity remains a cornerstone in mass spectrometry-based bottom-up proteomics. As the abundance of these precursors can dramatically change from sample to sample, so can peptide identifications. Identification alone is not enough in current proteomic applications, where the peptide intensities should be extracted to report on protein abundances. Label free quantitation (LFQ) relies on the extraction of peptide precursor ion intensities of one or more peptides to be correlated to protein abundance. Match between runs (MBR) represents an approach where identifications based upon MSMS in one run can be transferred to the next. MBR represents a big step in filling the missingness gap often observed in DDA experiments. Here we present LFQ functionality on the PaSER platform. PaSER provides real-time search for DDA acquisitions, with searches performed in real-time the LFQ step is expedited. We show that doing LFQ on PaSER results in database search and quantitation up to 30 times faster than using standard state-of-the-art algorithms.

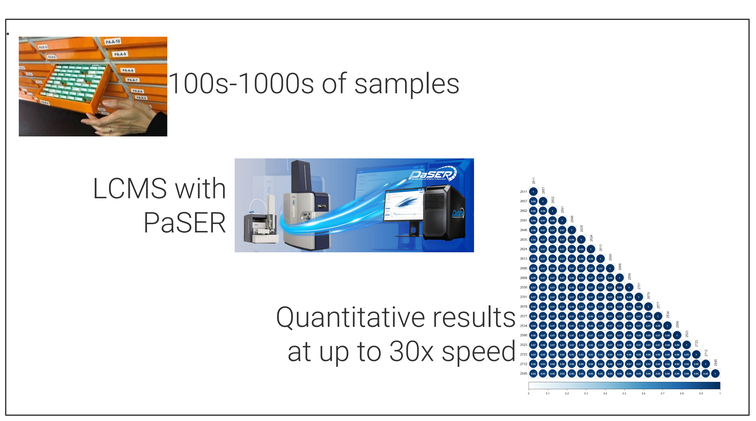


Fig. 1 timsTOF with PaSER demonstrating samples to LCMS to immediate results including LFQ findings at up to 30x traditional speeds

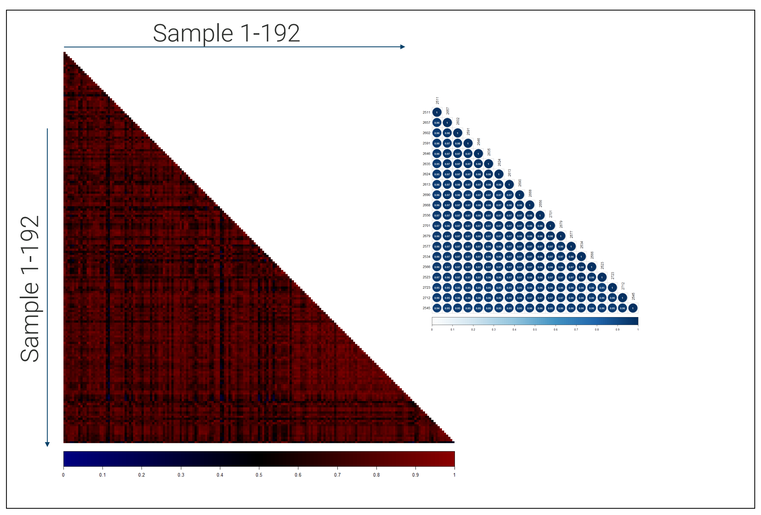
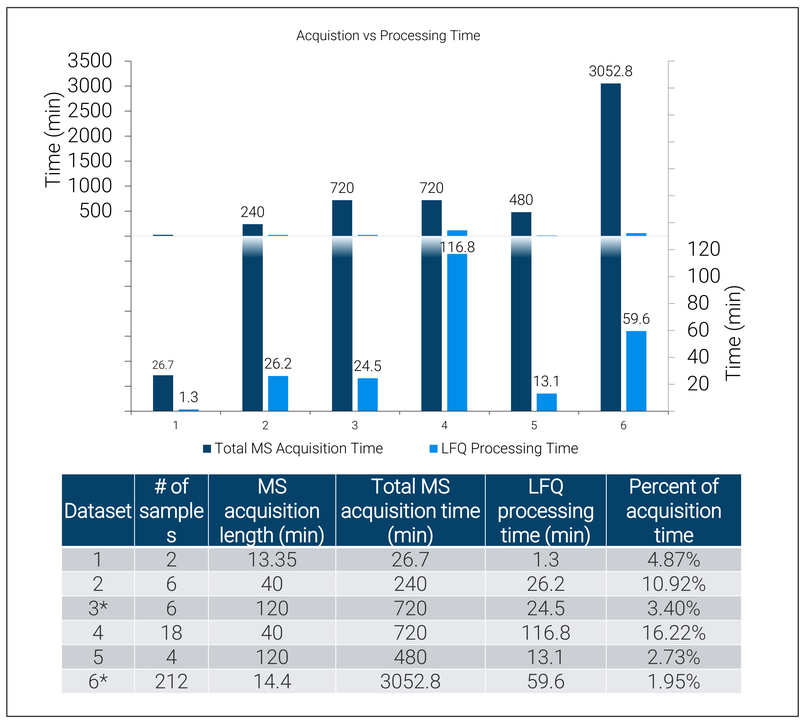


Fig. 2 Spearman correlation plot of 192 un-depleted plasma samples acquired on timsTOF Pro utilizing an EvoSep One LC operating on 100 SPD method. *Inset:* Spearman correlation plot of the 20 pooled samples for QC that were dispersed throughout the experiment.

Methods

LFQ datasets include depleted blood plasma, human-yeast-E. Coli (HYE) mixtures and K562 dilution series.

- The blood plasma (dataset 6 Fig. 2, 3) was collected from severe infection patients and was depleted for the 12 most abundant proteins. The depleted plasma proteome was digested with trypsin and diluted to a final concentration of 5 ng/µL in 0.1% FA. A total of 100 ng of each sample was loaded onto individual Evotips for desalting and then washed with 20 µL 0.1% FA followed by the addition of 100 µL storage solvent (0.1% FA) to keep the Evotips wet until analysis.
- The HYE mixture (dataset 3 Fig. 3, 4) was prepared from commercially available tryptic protein digests of H. sapiens (HeLa, Pierce), S. cerevisiae (Promega) and E. coli (Waters) and were mixed in two different experiments leading to a ratio of 1:1 (HeLa), 1:2 (S. cerevisiae) and 1:4 (E. coli) between the two samples. (https://doi.org/10.1074%2Fmcp.TIR119.001720)
- In these studies, a timsTOF Pro, Pro 2, and HT were used in PASEF mode. The ramp times varied from 50 ms to 100 ms depending on gradient length and sample complexity.
- LCs were either a nanoElute (Bruker) or and Evosep One (Evosep) with Ionopticks or PepSep columns.
- Data was searched in real-time using PaSER where search output was stored locally until LFQ processing, in which "Batch LFQ" was used to perform quantitation.



Dataset	# of samples	MS acquisition length (min)	Total MS acquisition time (min)	LFQ processing time (min)	Percent of acquisition time
1	2	13.35	26.7	1.3	4.87%
2	6	40	240	26.2	10.92%
3*	6	120	720	24.5	3.40%
4	18	40	720	116.8	16.22%
5	4	120	480	13.1	2.73%
6*	212	14.4	3052.8	59.6	1.95%

Fig. 3 Time savings using a LFQ workflow with PaSER. The total MS acquisition time vs processing time for quantitation in PaSER. (Note: search time = 0 min as it is done in real-time). Lower panel shows a zoomed in portion of the y-axis. Also shown is the data table, where Total MS acquisition time and LFQ processing time are utilized for the plots above. *Dataset 3 is utilized for figure 4 and dataset 6 is utilized in figure 2.

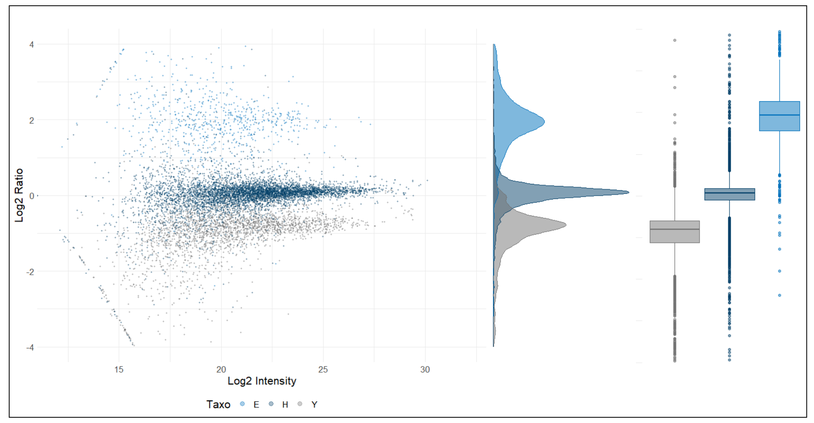


Fig. 4 Scatter plot of the ratio vs precursor intensity for Human-Yeast-E.Coli mixed dataset. The marginals show the density and boxplot representation of the same data.

Results

- All database searching is done in real-time during the MS acquisition, therefore no time cost
- Search results are stored locally on the PaSER box meaning no additional cost in file transfer time
- 6 different datasets of varying gradient lengths and sample complexity were used for LFQ demonstration where LFQ processing time is correlated to single run MS acquisition length
- The ratio between LFQ processing time and MS acquisition length varied from 6-30x's faster than MS acquisition
- For 212 plasma samples, requiring >2 days of LCMS acquisition time, the LFQ processing step was performed in under 1hr
- Spearman correlation plots for 20 controls and 192 plasma samples had a mean correlation > 0.96
- Scatter plot of Log2 ratio vs. intensity of a 3-proteome (HYE) mixture is within the expected values, showing quantitative accuracy

Summary

From our analysis of 6 different datasets of different matrices and quantitative experiments we show that LFQ with PaSER real-time search is fast (up to 30x instrument acquisition), scalable as tested with 212 plasma samples and both precise and accurate as indicated by the Spearman correlations and scatter plot analysis of multi-species mixtures. With automated sample processing, fast and sensitive timsTOF mass spectrometers, and now automated and real-time data processing, the use of deep and quantitative proteomics to solve complex biological problems becomes widely accessible. Quantitative capabilities on PaSER now include TIMS DIA-NN for dia-PASEF workflows and LFQ for DDA workflows covering most all experimental types.

Conclusion

- LFQ for DDA acquisition on PaSER is automated, scalable, fast, sensitive and accurate
- PaSER can dramatically expedite the ability to convert data to information
- Large scale bottom-up proteomics studies providing quantitative insight are now seamless to perform

For Research Use Only. Not for use in diagnostic procedures.

Innovation with Integrity