Communication

Data-Independent Acquisition (DIA) is Superior for High Precisely Phospho-Peptide Quantification in Fungi

Katharina Bersching [1,#], Thomas Michna [2,#], Stefan Tenzer [2,3] and Stefan Jacob [1,*]

[1] Institute of Biotechnology and Drug Research gGmbH (IBWF), Hanns-Dieter-Hüsch-Weg 17, Mainz, Germany; bersching@ibwf.de; jacob@ibwf.de
[2] Institute for Immunology, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany; thmichna@uni-mainz.de; tenzer@uni-mainz.de
[3] Helmholtz-Institute for Translational Oncology Mainz (HI-TRON), Mainz, Germany, German Cancer Research Center (DKFZ), Heidelberg, Germany; tenzer@uni-mainz.de
* Correspondence: jacob@ibwf.de
# Katharina Bersching and Thomas Michna contributed equally to this work

Abstract: The dynamic interplay of signaling networks in most major cellular processes is characterized by the orchestration of reversible protein phosphorylation. Consequently, analytic methods like quantitative phospho-peptidomics has been pushed forward from a highly specialized edge-technique to a powerful and versatile platform for comprehensively analyzing the phosphorylation profile of living organisms. Despite enormous progress in instrumentation and bioinformatics, a major problem remains a high number of missing values caused by the experimental procedure due to either a random phospho-peptide enrichment selectivity or borderline signal intensities, which both cause the exclusion for fragmentation using the commonly applied data dependent acquisition (DDA) mode. Consequently, an incomplete dataset reduces confidence in the subsequent statistical bioinformatic processing. Here, we successfully applied data independent acquisition (DIA) by using the filamentous fungus *Magnaporthe oryzae* as model organism and could prove that while maintaining data quality (such as phosphosite and peptide sequence confidence), the data completeness increases dramatically. Since the method presented here reduces the LC-MS/MS analysis from 3 h to 1 h and increases the number of phoshosites identified up to 10-fold in contrast to published studies in fungi, we pushed the phospho-proteomic technique beyond its current limits and could provide a sophisticated resource for investigation of signaling processes in filamentous fungi.

Keywords: Proteomics; LC-MS/MS; phosphopeptide enrichment; bioinformatics; cellular signaling; *Magnaporthe oryzae*; phosphorylation; DDA; DIA; phospho-peptidomics

1. Introduction

The phosphorylation of proteins is among the most prominent and significant post-translational modifications[1]. Protein kinases make this reversible modification possible by the addition of a phosphate group (PO_4) to the polar residual of amino acids. As a consequence, this addition modifies the protein from an apolar (hydrophobic) to a more polar (hydrophilic) state. The subsequent conformational changes enable interactions with other molecules[2]. The biochemical nature of phosphorylated amino acids facilitate interaction with other proteins what enables i.e. the assembly of proteins or protein complexes. The fundamental challenge in research of signal transduction pathways is the highly dynamic nature of reversible phosphorylation of the involved signaling proteins[3][4]. Apart from the static information whether a peptide, peptide fragment or a certain amino acid residue is phosphorylated or not ("on" or "off"), it is of utmost interest to understand the dynamic alteration of quantitative changes in phosphorylation levels over time associated with a given stimulus or cellular process[5].

In the last decade, liquid chromatography-tandem mass spectrometry (LC-MS/MS) and variations thereof were the method of choice to quantify thousands of proteins across

 doi:10.20944/preprints202211.0165.v1

multiple biological samples with high throughput, robustness and sensitivity[6]. An unsolved problem in quantitative phospho-peptidomics by mass spectrometry is still the low abundance of phosphorylated proteins as compared to the complete proteome and to the complement of a given protein, from which naturally only a small portion is phosphorylated in a particular way[7][8][9]. The drive to solve this problem and constantly improve the precision of protein measurements pushes MS-techniques and the related methods forward, following the overarching goal of proteomics to comprehensively identify and quantify all proteins and protein modifications in a biological system[10]. For example, the enrichment of phospho-peptides is absolutely necessary including immunoprecipitation (IP), metal oxide/immobilized ion affinity chromatography (MOAC/IMAC), fractionation strategies like high-pH reversed-phase chromatography (HpH RP), strong cation exchange (SCX), or electrostatic repulsion hydrophilic interaction liquid chromatography (ERLIC) is necessary[11].

One of the major bottlenecks to obtain nowadays a comprehensive and precise analysis is not the accuracy of the instruments and measurements used but rather important processes like data acquisition and data processing[12]. Most of MS-based proteomic work-flows use the "data-dependent acquisition" (DDA) strategy[13][14][15][16], often in combination with "dynamic exclusion" (DE), which rules out a selection of fragmented peptides within a specific time window[17]. In DDA, precursor ions are stochastically selected on the basis of their signal intensity and subsequently fragmented, separated and finally detected by a mass analyzer such as a "time-of-flight" (TOF) or an Orbitrap[18]. In more detail, the top N most intensive m/z ions are identified from the MS1 scan (precursor spectrum, in proteomics typically precursors are peptides) by the operating software of the mass spectrometer and sequentially selected with very narrow window (e.g. ±0.5 Dalton) by the quadrupole for fragmentation, so their MS2 spectra (fragment spectra) can be collected. The resulting fragment m/z values vary by the corresponding masses for amino acids according to their sequence. This way, the processing software (or the analyzing scientist) can compare the obtained amino acid sequence with the measured m/z of the intact peptide. Depending on the amount of amino acid sequence evidence and the congruence between theoretical and measured precursor m/z, a score for the probability of a correct identification is calculated[19]. The selected number N of most intense ions is typically between 10 to 25 and can be chosen depending on the instrument speed and on the analytical need. When short LC gradients and highly complex MS1 spectra are present, a high N is needed for deep peptide coverage. On the other hand, a high N costs measurement time and MS1 quantification accuracy. In general, the DDA strategy decides depending on the MS1 information, which precursors are selected for fragmentation. It provides clean and high-quality spectra, that can also be used for *de-novo* sequencing with certain prerequisites. In addition to that, the data processing is not computationally intensive and implements easy and straight forward algorithms that are accessible to a broad community.

In contrast, to alleviate the limitations associated with DDA and DE, strategies on unbiased "data-independent acquisition" (DIA) are available, in which every peptide within a specific time window is fragmented[20]. That means, in data independent acquisition strategy, no preselection is performed. The fragmentation is independent of any MS1 information. Instead of choosing a very narrow window for selecting the precursors for fragmentation, a wide window of precursor m/z are allowed to pass through the quadrupole[21]. This way, multiple precursor co-fragment and create chimeric MS2 spectra, where the assignment of precursor and their corresponding fragments is not easily possible. More complex bioinformatic algorithms have to be applied to elucidate the amino acid evidence for each precursor[22]. This also includes the use of spectral libraries, which are either labor intensive or computationally intensive to create. Recent developments in the proteomics community show improvements in algorithms and software to be able to process DIA generated raw-data in a comprehensive and user-friendly way. The accessibility to high performing computer systems have paved the way for increasing use of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 November 2022 doi:10.20944/preprints202211.0165.v1

DIA[23]. The major advantage of DIA is a robust and accurate quantification as well as the decrease of missing values, due to the fact that no selection of precursors is performed. Instead also borderline signal intensities are fragmented and have the chance to be identified and quantified.

Prior to this study, it was generally assumed that DIA can quantify the same number of proteins as typically identified by DDA methods, but with better accuracy and reproducibility across many samples[24]. In DDA, one major problem was the high number of missing values caused by the experimental procedure due to either a random phosphopeptide enrichment selectivity or borderline signal intensities, which both cause the exclusion for fragmentation. From this follows an incomplete dataset reducing confidence in the subsequent statistical bioinformatic processing.

Here, we successfully developed a method including DIA for data acquisition by using the filamentous fungus *Magnaporthe oryzae* as model organism. Application of this method resulted in absolutely reliable datasets with high data quality (such as phosphosite and peptide sequence confidence) while at the same time data completeness increases dramatically. We are convinced, that this is an excellent basis for further research on the dynamic processes of phosphorylation in signaling networks in a high quality as never seen before.

2. Materials and Methods

2.1 Sample preparation

2.1.1 Cultivation of *Magnaporthe oryzae*

The fungal strain used in this study was *Magnaporthe oryzae* (*M. oryzae* 70-15 strain (MoWT), Fungal Genetics Stock Center). The strain was maintained at 26°C on complete medium (CM) according to[25]. For protein isolation, the *M. oryzae* cultures were grown in 250 ml liquid CM in 500-ml glass flasks for 96 h at 26°C and 120 rpm. Samples were then taken and the mycelium was immediately separated from the culture fluid and ground into powder with the TissueLyserII (Qiagen) according to the user manual.

2.1.2 Cell lysis and protein digest

If not stated otherwise, all reagents were used in LC-MS/MS grade from common vendors. The sample preparation for all *Magnaporthe oryzae* samples has been performed as described in[11]. In short, a sample aliquot of lyophilized and grinded mycelium was suspended in boiling SDS/DTT lysis buffer with following treatment of ultrasound. Proteins were precipitated by chloroform/methanol precipitation and resolubilized in urea containing buffer. DNA/RNA removal by benzonase and tryptic digest was performed overnight followed by desalting and lyophilization. An aliquot of lyophilized peptides was used for proteome analysis, 1000 µg were subjected to phospho-peptide enrichment by TiO_2 spin tips.

2.1.3 Phosphopeptide enrichment

Phospho-peptide enrichment of *M. oryzae* samples was performed as described in[11].

2.2 Peptide identification

2.2.1 LC-MS/MS of *M. oryzae* samples

2 µL of the reconstituted phospho-peptides were separated on an Ultimate 3000 nanoUPLC (Thermo Scientific, USA) with 300 nL/min by a reversed phase C18 column (HSS-T3 C18 1.8 µm, 75 µm x 250 mm, Waters Corporation) at 55°C using a 45 min linear gradient from 5 % Eluent A (0.1 % TFA, 3 % DMSO in water) to 35 % Eluent B (0.1 % TFA, 3 % DMSO in ACN) followed by ionization in positive mode using a Nanospray Flex electrospray ionization source (Thermo Scientific). Mass-to-charge analysis of the eluting peptides was performed using an Orbitrap Exploris 480 (Thermo Scientific) in data independent acquisition (DIA) mode. MS1 scans were acquired with a resolution of 120.000 at 200 m/z in a range of 345-1250 m/z. RF lens was set to 40 % and AGC target to 300 % (i.e. corresponding to $3x\ 10^6$ charges). DIA MS2 scans were acquired with a resolution of 30.000 at 200 m/z with a variable window scheme (as shown in supplementary Table S1). The

 doi:10.20944/preprints202211.0165.v1

normalized collision energy was set to 27 %, RF lens to 40 % and AGC target to 1000 % (i.e. corresponding to 10x 10^6 charges).

2.2.2 Data processing parameters

Peptides were identified and label-free quantification of proteins was performed using DIA-NN (v1.8). Full proteome samples from *M. oryzae* were processed using library free mode with standard parameters, except for tryptic cleavage sites considering no cleavage before proline. The FASTA protein database contained 12.790 protein entries of the *M. oryzae* reference proteome and 172 common contaminant proteins (Both from Uniprot). For phospho-peptide analysis of *M. oryzae*, a phosphopeptide spectral library was predicted *in-silico* using the built-in library free prediction algorithm provided by DIA-NN. For *M. oryzae*, the aforementioned FASTA database was used as basis.

The spectra library was predicted with the precursor charge range set between 1-4 and the range for fragment ions and precursor mass to charge ratio was limited to 250-1250 m/z. The peptide length was set to 7-30. Tryptic cleavage considering no cleavage after the lysine or arginine is followed by proline, maximum one missed cleavage was allowed. N-terminal methionine excision was enabled and cysteine carbamidomethylation was set as fixed modification. The maximum number of variable modifications was set to 3, allowing exclusively UniMod:21 modifications, i.e. mass delta of 79.9663 corresponding to phosphorylation at serine, threonine and tyrosine. The generated spectral libraries were used for follow up identification and quantification in DIA-NN using the standard settings.

2.2.3 Availability of raw data

All raw data have been uploaded via JPOST[26] to be available on proteomeXchange[27] and can be accessed with the identifier PXD034481.

3. Results and Discussion

Comparison of DDA vs. DIA approach for phospho-peptide identification

A promising approach to gain more confidence in phospho-peptide data is the data independent acquisition (DIA) approach. Per definition, DIA generates MS2 spectra of higher complexity compared to DDA. Especially the identification of the phosphosites requires sophisticated bioinformatic methods that had not been available in the past. Recent implementations in proprietary software such as Spectronaut[28] and developments of open source software such as DIA-NN[23] in combination with affordable high-performance compational resources made the analysis of phospho-peptides in DIA possible with sufficient confidence within a reasonable time frame. There are only few publications describing the use of DIA for phospho-peptides[28][29][30] and thus the differences in the data quality have not been reviewed yet comprehensively, especially in the context of predicted spectral libraries. Therefore, a dataset of three biological replicates of wildtype *M. oryzae* was measured in DDA and in DIA with a Bruker nanoElute coupled to a timsTOF Pro 2, processed with PEAKS and DIA-NN, respectively, and the results summarized in figure 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 9 November 2022 doi:10.20944/preprints202211.0165.v1

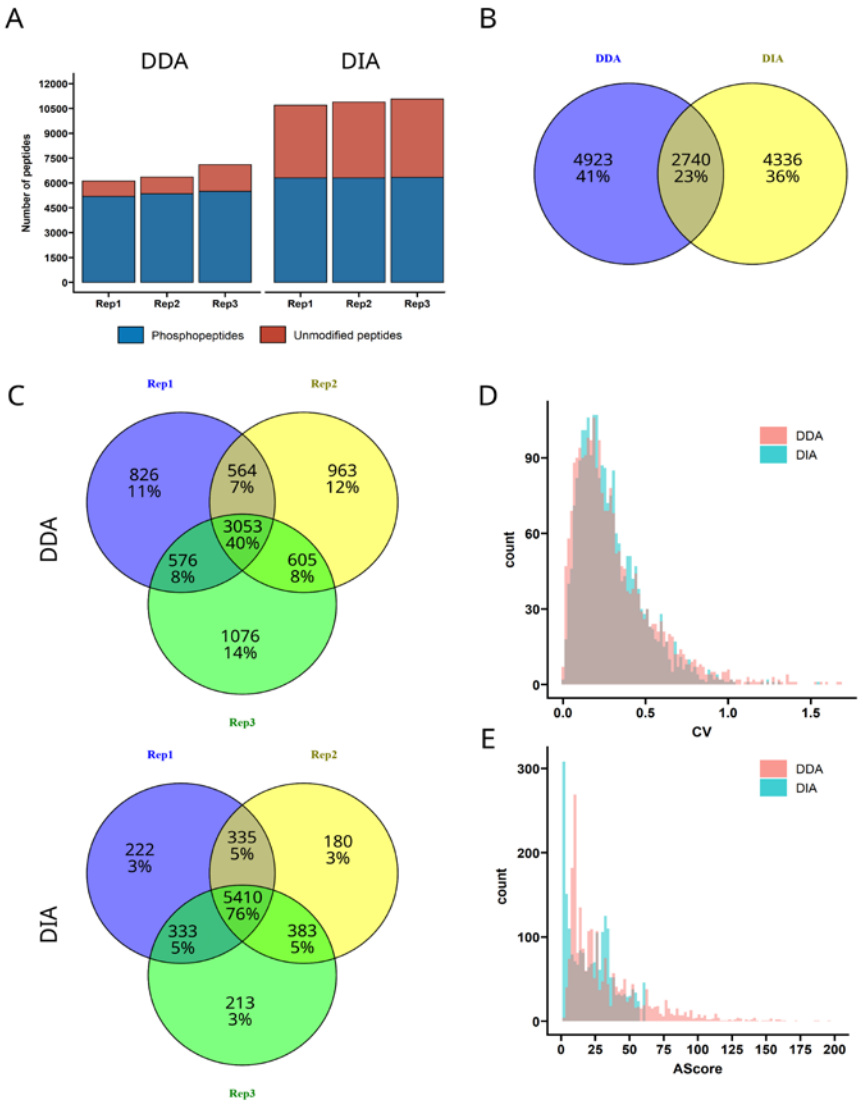

Figure 1: Performance comparison of three *M. oryzae* biolocigal replicates measured in DDA and DIA regarding A) peptide counts B) overlap of identified phospho-peptides C) overlap withing DDA and DIA replicates D) precursor quantity reproducibility and E) phosphosite identification confidence.

The number of identified phospho-peptides is comparable between DDA and DIA, while the number of unmodified peptides in the DIA samples is significantly higher. Consequently, the apparent enrichment efficiency decreases from around 80 % in DDA to 50 % in DIA (Fig.1A). This observation is explained by the DIA scheme, as no criterion for fragmentation is applied, also unmodified peptides with low signal intensity are selected for fragmentation. Interestingly, all unique phospho-peptide identifications of the three replicates combined is roughly 10 % higher in DDA (7.663 peptides) compared to DIA (7076 peptides) and the overlap of peptide IDs is small (23 %) as shown in figure 1B. The overlap of peptide sequences without considering the phosphosite was slightly increased with 44 %, so roughly 20 % differ in the assigned phosphosite. It has also not been shown yet, to which extend the software DIA-NN actually provides false positive identifications. To exclude a higher false positive rate as reason for the low number of overlapping identifications, both datasets (DDA and DIA) were searched in either PEAKS or DIA-NN against a connected database of *Mus musculus* and *Magnaporthe oryzae* proteome. As the sample was generated from *M. oryzae*, the number of identified *Mus musculus* proteins are expected to be not more than the previously set up false-discovery rate of 1 %. For the

 doi:10.20944/preprints202211.0165.v1

DDA dataset, from 36.006 total identifications were 112 peptides identified from *Mus musculus* (i.e. 0.3 %) and in the DIA dataset, from 40.817 total identifications were only 65 identified from *Mus musculus* (i.e. 0.2 %). In conclusion, the false identification rate can be excluded as a reason for the low overlap between the data acquisition strategies. Comparing the intra-sample group overlap of the identifications within the replicate measurements (Fig. 1C), reveals another possible reason for the difference in peptide numbers. DIA provides consistently more reproducible identifications, while the overlap for DDA measurements is much less. When accepting only peptides with at least two out of three identifications, the number of quantifiable peptides is 35 % higher in DIA (6.461 peptides) compared to DDA (4.798 peptides), while the number of complete peptide data (three out of three) is also increased in DIA measurement. Thus, not only the number of quantifiable peptides but also data completeness is increased.

In order to understand the biology, not only the number of quantifiable peptides is important, but also the reproducibility and quality. Therefore, the coefficients of variation (CVs) for every quantifiable peptide (at least two out of three replicates) have been calculated from the replicate measurements and plotted as histogram (Fig.1D). The difference between both datasets is not significant with median CVs around 25 %, which is reasonable due to technical variability in LC-MS/MS measurement. A beneficial effect of DIA on data quality has been shown on proteome level[31], which results from the higher number of peptides that are available for quantification. A second important aspect in phosphopeptide identification is the correct localization of the phosphosite. Both approaches, DDA and DIA offer a confidence measure for the correct site. Nevertheless, even when no evidence for the correct phosphosite is present in the spectrum, the peptide still harbours a phospho-group at some amino acid, otherwise the peptide precursor mass would not be correct. Thus, we can be confident due to common quality control measures (e.g. false discovery rate calculation at peptide level) that there is a phospho-group somwhere in the peptide present, but the correct phosphosite identification can remain ambiguous. Therefore, DIA-NN calculates a site localization probability and PEAKS provides the Ascore, which is calculated by multiplying the negative decadic logarithm of the p-value for incorrect identification by 10. Consequently, the higher the AScore the more confident is the identification with a maximum possible value of 1000. Typically, a confidence of at least 75 % (for calculation of AScore: 25 % probability of false localization) is desired[32]. Therefore, a common cut of value for the AScore is a value of 6, corresponding to 25 % false localization probability. In Figure 1E, the distribution of AScores obtained from both acquisition strategies is shown. The DDA AScores peak is around an value of 10, whereas DIA data seems to provide two different peaks, the first peak with an AScore below 6 and the second peak with an AScore around 30, which equals a site confidence of 99.9 %. Thus, the median site confidence is roughly the same, due to the inhomogenious distribution of the DIA-NN confidences. The reason for this difference is presumably the higher complexity on MS2 in DIA data. There, confidence is only achieved in presence of strong fragment evidence, whereas the algorithm of PEAKS for processing DDA MS2 spectra seems to have a more refined algorithm to assign also calculate variances in probability with high sensitivity. Therefore, the assumption that DDA data provides more confidence in the site localization by higher quality spectra is only partly true. Nevertheless, in discovery phospho-proteomics, the correct phosphorylation site is anyway of less importance. More importantly, both algorithms provide equally high confidence that these peptides are phosphorylated (regardless the phosphosite). Conclusions about active/inactive pathways or protein phosphorylation with approximate protein sites can be drawn anyway.

In conclusion, the application of DIA is a promising strategy for the comprehensive description of a phospho-proteomics dataset. We have shown, that data completeness increases dramatically while the data quality remains at least equal. The downside of the DIA application are a resource intensive and time consuming bioinformatic processing and the lack for intuitive spectra visualization. A possible solution to this is provided by

 doi:10.20944/preprints202211.0165.v1

the proprietary software Spectronaut, that is able to visualize XICs of precursors and fragments in a user-friendly way[28]. Nevertheless, DIA-NN has been shown to provide superior identification performance utilizing neuronal networks while being open source at the same time. A direct phospho-peptide ID benchmark of both software has not been described in the literature yet and would serve as interesting starting point for further bioinformatics research.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1: title; Table S1: title; Video S1: title.

Table S1: Variable window sizes foe DIA acquisition with Orbitrap Exploris 480.

Window	center (m/z)	Isolation Window (m/z)	m/z start	m/z end
1	392.58	95.2	344.98	440.18
2	459.56	39.8	439.66	479.46
3	494.25	30.6	478.95	509.55
4	522.78	27.5	509.03	536.53
5	548.53	25	536.03	561.03
6	572.89	24.7	560.54	585.24
7	596.9	24.3	584.75	609.05
8	620.81	24.5	608.56	633.06
9	645.09	25.1	632.54	657.64
10	670.1	26	657.1	683.1
11	696.1	27	682.6	709.6
12	723.35	28.5	709.1	737.6
13	752.37	30.6	737.07	767.67
14	783.12	31.9	767.17	799.07
15	816.29	35.4	798.59	833.99
16	852.84	38.7	833.49	872.19
17	893.43	43.5	871.68	915.18
18	939.95	50.6	914.65	965.25
19	997.57	65.7	964.72	1030.42
20	1070.86	81.9	1029.91	1111.81
21	1180.66	138.7	1111.31	1250.01

Author Contributions: "Conceptualization, T.M., S.T. and S.J.; methodology, T.M. and K.B.; software, S.T.; validation, T.M. and K.B.; formal analysis, T.M. and S.T.; resources, S.T. and S.J.; writing—original draft preparation, K.B. and T.M; writing—review and editing, S.J. and S.T.; supervision, S.J. and S.T.; project administration, S.J. and S.T.; funding acquisition, S.J. and S.T. All authors have read and agreed to the published version of the manuscript."

Funding: "This research was funded by DFG (German Research Foundation, Bonn), grant numbers 403841309 and 426554840".

Institutional Review Board Statement: "Not applicable"

Informed Consent Statement: "Not applicable."

Data Availability Statement: All raw data have been uploaded via JPOST[26] to be available on proteo-meXchange[27] and can be accessed with the identifier PXD034481.

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] X. Li, M. Wilmanns, J. Thornton, and M. Köhn, "Elucidating human phosphatase-substrate networks," *Sci. Signal.*, vol. 6, no. 275, May 2013, doi: 10.1126/SCISIGNAL.2003203/SUPPL_FILE/6_RS10_TABLES_S1_TO_S16.ZIP.

[2] F. Ardito, M. Giuliani, D. Perrone, G. Troiano, and L. Lo Muzio, "The crucial role of protein phosphorylation in cell signalingand its use as targeted therapy (Review)," *Int. J. Mol. Med.*, vol. 40, no. 2, pp. 271–280, Aug. 2017, doi: 10.3892/IJMM.2017.3036/HTML.

[3] J. Dengjel *et al.*, "Quantitative proteomic assessment of very early cellular signaling events," *Nat. Biotechnol. 2007 255*, vol. 25,

no. 5, pp. 566–568, May 2007, doi: 10.1038/nbt1301.

[4] K. T. G. Rigbolt and B. Blagoev, "Quantitative phosphoproteomics to characterize signaling networks," *Semin. Cell Dev. Biol.*, vol. 23, no. 8, pp. 863–871, Oct. 2012, doi: 10.1016/J.SEMCDB.2012.05.006.

[5] H. K. Kweon and P. C. Andrews, "Quantitative analysis of global phosphorylation changes with high-resolution tandem mass spectrometry and stable isotopic labeling," *Methods*, vol. 61, no. 3, pp. 251–259, Jun. 2013, doi: 10.1016/J.YMETH.2013.04.010.

[6] Y. Bian, C. Gao, and B. Kuster, "On the potential of micro-flow LC-MS/MS in proteomics," *https://doi.org/10.1080/14789450.2022.2134780*, 2022, doi: 10.1080/14789450.2022.2134780.

[7] R. Wu *et al.*, "A large-scale method to measure absolute protein phosphorylation stoichiometries," *Nat. Methods 2011 88*, vol. 8, no. 8, pp. 677–683, Jul. 2011, doi: 10.1038/nmeth.1636.

[8] J. V. Olsen *et al.*, "Quantitative phosphoproteomics revealswidespread full phosphorylation site occupancy during mitosis," *Sci. Signal.*, vol. 3, no. 104, Jan. 2010, doi: 10.1126/SCISIGNAL.2000475/SUPPL_FILE/SUPPLEMENTARYTABLE_S9.XLS.

[9] D. W. Greening, E. A. Kapp, and R. J. Simpson, "The Peptidome Comes of Age: Mass Spectrometry-Based Characterization of the Circulating Cancer Peptidome," *Enzymes*, vol. 42, pp. 27–64, Jan. 2017, doi: 10.1016/BS.ENZ.2017.08.003.

[10] A. J. Link *et al.*, "Direct analysis of protein complexes using mass spectrometry," *Nat. Biotechnol. 1999 177*, vol. 17, no. 7, pp. 676–682, 1999, doi: 10.1038/10890.

[11] T. Michna and S. Tenzer, "Quantitative Proteome and Phosphoproteome Profiling in Magnaporthe oryzae," *Methods Mol. Biol.*, vol. 2356, pp. 109–119, 2021, doi: 10.1007/978-1-0716-1613-0_9/FIGURES/1.

[12] M. U. Tariq, M. Haseeb, M. Aledhari, R. Razzak, R. M. Parizi, and F. Saeed, "Methods for Proteogenomics Data Analysis, Challenges, and Scalability Bottlenecks: A Survey," *IEEE Access*, vol. 9, pp. 5497–5516, 2021, doi: 10.1109/ACCESS.2020.3047588.

[13] M. Mann, R. C. Hendrickson, and A. Pandey, "Analysis of proteins and proteomes by mass spectrometry," *Annu. Rev. Biochem.*, vol. 70, pp. 437–473, 2001, doi: 10.1146/ANNUREV.BIOCHEM.70.1.437.

[14] S. E. Ong *et al.*, "Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.," *Mol. Cell. Proteomics*, vol. 1, no. 5, pp. 376–386, May 2002, doi: 10.1074/mcp.M200025-MCP200.

[15] P. L. Ross *et al.*, "Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents," *Mol. Cell. Proteomics*, vol. 3, no. 12, pp. 1154–1169, Dec. 2004, doi: 10.1074/mcp.M400129-MCP200.

[16] N. W. Bateman *et al.*, "Maximizing peptide identification events in proteomic workflows using data-dependent acquisition (DDA)," *Mol. Cell. Proteomics*, vol. 13, no. 1, pp. 329–338, Jan. 2014, doi: 10.1074/mcp.M112.026500.

[17] Y. Zhang, Z. Wen, M. P. Washburn, and L. Florens, "Effect of dynamic exclusion duration on spectral count based quantitative proteomics," *Anal. Chem.*, vol. 81, no. 15, pp. 6317–6326, Aug. 2009, doi: 10.1021/AC9004887/SUPPL_FILE/AC9004887_SI_001.PDF.

[18] J. Zhou, Y. Li, X. Chen, L. Zhong, and Y. Yin, "Development of data-independent acquisition workflows for metabolomic analysis on a quadrupole-orbitrap platform," *Talanta*, vol. 164, pp. 128–136, Mar. 2017, doi: 10.1016/J.TALANTA.2016.11.048.

[19] J. Zhang *et al.*, "PEAKS DB: De novo sequencing assisted database search for sensitive and accurate peptide identification," *Mol. Cell. Proteomics*, vol. 11, no. 4, Apr. 2012, doi: 10.1074/mcp.M111.010587.

[20] G. Nys, C. Nix, G. Cobraiville, A. C. Servais, and M. Fillet, "Enhancing protein discoverability by data independent acquisition assisted by ion mobility mass spectrometry," *Talanta*, vol. 213, p. 120812, Jun. 2020, doi: 10.1016/J.TALANTA.2020.120812.

[21] K. W. Li, M. A. Gonzalez-Lozano, F. Koopmans, and A. B. Smit, "Recent Developments in Data Independent Acquisition (DIA) Mass Spectrometry: Application of Quantitative Analysis of the Brain Proteome," *Front. Mol. Neurosci.*, vol. 13, p. 248, Dec. 2020, doi: 10.3389/FNMOL.2020.564446/BIBTEX.

[22] L. Krasny and P. H. Huang, "Data-independent acquisition mass spectrometry (DIA-MS) for proteomic applications in oncology," *Mol. Omi.*, vol. 17, no. 1, pp. 29–42, Feb. 2021, doi: 10.1039/D0MO00072H.

[23] V. Demichev, C. B. Messner, S. I. Vernardis, K. S. Lilley, and M. Ralser, "DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput," *Nat. Methods 2019 171*, vol. 17, no. 1, pp. 41–44, Nov. 2019, doi: 10.1038/s41592-019-0638-x.

[24] K. Barkovits *et al.*, "Reproducibility, specificity and accuracy of relative quantification using spectral librarybased data-independent acquisition," *Mol. Cell. Proteomics*, vol. 19, no. 1, pp. 181–197, Jan. 2020, doi: 10.1074/mcp.RA119.001714.

[25] S. Bohnert *et al.*, "Fungicide resistance toward fludioxonil conferred by overexpression of the phosphatase gene MoPTP2 in Magnaporthe oryzae," *Mol. Microbiol.*, vol. 111, no. 3, pp. 662–677, Mar. 2019, doi: 10.1111/MMI.14179.

[26] S. Okuda *et al.*, "jPOSTrepo: an international standard data repository for proteomes," *Nucleic Acids Res.*, vol. 45, no. D1, pp. D1107–D1111, Jan. 2017, doi: 10.1093/NAR/GKW1080.

[27] J. A. Vizcaíno *et al.*, "ProteomeXchange provides globally coordinated proteomics data submission and dissemination," *Nat. Biotechnol. 2014 323*, vol. 32, no. 3, pp. 223–226, Mar. 2014, doi: 10.1038/nbt.2839.

[28] D. B. Bekker-Jensen *et al.*, "Rapid and site-specific deep phosphoproteome profiling by data-independent acquisition without the need for spectral libraries," *Nat. Commun. 2020 111*, vol. 11, no. 1, pp. 1–12, Feb. 2020, doi: 10.1038/s41467-020-14609-1.

[29] A. Martinez-Val, D. B. Bekker-Jensen, A. Hogrebe, and J. V. Olsen, "Data Processing and Analysis for DIA-Based Phosphoproteomics Using Spectronaut," *Methods Mol. Biol.*, vol. 2361, pp. 95–107, 2021, doi: 10.1007/978-1-0716-1641-3_6/FIGURES/1.

[30] R. B. Kitata *et al.*, "A data-independent acquisition-based global phosphoproteomics system enables deep profiling," *Nat. Commun. 2021 121*, vol. 12, no. 1, pp. 1–14, May 2021, doi: 10.1038/s41467-021-22759-z.

[31] R. Bruderer *et al.*, "Optimization of experimental parameters in data-independent mass spectrometry significantly increases depth and reproducibility of results," *Mol. Cell. Proteomics*, vol. 16, no. 12, pp. 2296–2309, Dec. 2017, doi: 10.1074/mcp.RA117.000314.

[32] J. V. Olsen *et al.*, "Global, In Vivo, and Site-Specific Phosphorylation Dynamics in Signaling Networks," *Cell*, vol. 127, no. 3, pp. 635–648, Nov. 2006, doi: 10.1016/J.CELL.2006.09.026.