SWATH® acquisition enables precise label-free quantification on proteome-scale

Qiang Huang[1,2]; Lu Yang[1,2]; Lin Guo[1,2]; Zhiyuan Wang[1,2]; Xiangyun Yang[1,2]; Ji Luo[3]; Wenhai Jin[3] Juanying Ye[1,2] Bing Shan[1,2] Yaoyang Zhang[1,2]

[1]Interdisciplinary Research Center on Biology and Chemistry, Shanghai, China; [2]Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai, China ; [3]SCIEX, Shanghai, China

ABSTRACT

Mass spectrometry-based proteomics has been emerged as a powerful tool in biological studies. The shotgun proteomics strategy, in which proteolytic peptides are analyzed in data dependent mode, enables the most comprehensive proteome detection. The quantitative proteomics uses stable isotopes, or label-free method to measure relative protein abundance. SWATH acquisition is a recently developed technique, in which data independent acquisition is coupled with peptide spectral library match. In principle SWATH acquisition is able to do label-free quantification in a MRM-like manner, which has higher quantification accuracy and precision. Our study first time assessed the quantification performance of SWATH acquisition on proteome-scale using a complex mouse cell lysate sample. In total 3503 proteins got identified and quantified without sample prefractionation. The SWATH acquisition shows outstanding quantification precision, whereas the quantification accuracy becomes less perfect when protein abundance differs greatly. However, this inaccuracy does not prevent discovering biological correlates, because the measured signal intensities had linear relationship to the sample loading amount, thus the SWATH acquisition can predict precisely the significance of a protein. Our results prove that SWATH acquisition can provide precise label-free quantification on proteome-scale.

INTRODUCTION

Two complementary strategies, untargeted and targeted, are being employed widely for mass spectrometry-based protein analysis. Shotgun is the most powerful method in terms of throughput and proteome coverage. From a complex sample (e.g. whole cell lysate), shotgun method typically allows us to identify thousands of proteins, even over ten thousands proteins with sample prefractionation. By coupled with or without stable isotope labeling, shotgun proteomics also has been succeeded in quantifying proteins at proteome-scale. The isotope labeling strategies are more precise and accurate, but labeling procedures are complicated and expensive. By contrast, label-free method is easy to apply, and does not have sample limitation, but the quantification precision and accuracy are low.

Targeted approach usually refers to multiple reaction monitoring (MRM), in which a number of target peptides and their corresponding transition ions are selected manually for analysis with better reproducibility and sensitivity.

SWATH acquisition is a recently developed technique, in which data independent acquisition is coupled with peptide spectral library match. In principle SWATH acquisition is able to do label-free quantification in a MRM-like manner, which has higher quantification accuracy and precision. Our study first time assessed the quantification performance of SWATH method on proteome-scale using a complex mouse cell lysate sample.

MATERIALS AND METHODS

Sample Preparartion and Experimental design:
100 µg of Mouse fibrosarcoma L929 cell protein was precipitated with 5x volume of cold acetone then follow on further digestion. The digested peptides was diluted with 0.1% FA in water. Two microgram of peptides were analyzed by data dependent shotgun method for library generation. In SWATH acquisition experiments, four samples containing 0.25 µg, 0.5 µg, 1 µg, 2 µg of peptide mixtures derived from L929 cells respectively were analyzed using SWATH acquisition. Then The SWATH data files were loaded into SWATH MicroApp software, by which the XICs of all the transition ions were extracted to obtain relative area which used to quantitative ability of SWATH technology.

LC Conditions:
Peptides were separated on a nano column (75 µm x 15 cm, C18, 3 µm, 120 Å) using an Eksigent 425 HPLC system. The flow rate was set to 300 nL/min over 120 min multisegment gradienton solvent B (0.1% FA in 98% ACN): 0 min 5%, 1 min 11%, 80 min 28%, 104 min 50%, 105 min 80%, 110 min 80%, 111 min, 5%, 120 min 5%. The same chromatographic conditions were applied to all the measurements in this study. The separated peptides were injected directly into a TripleTOF® 5600 Plus mass spectrometer using positive ion mode for both shotgun and SWATH analysis..

MS Conditions:
A SCIEX TripleTOF® 5600 Plus LC/MS system with NanoIII™ source and Electrospray Ionization (ESI) probe was used. For library generation, A cycle of one full-scan mass spectrum (350-1250 m/z) with accumulation time of 0.25 s followed by 40 data dependent MS/MS spectra (100-1500 m/z) with accumulation time of 50 ms was repeated continuously throughout the whole gradient. For SWATH acquisition, A consecutive data independent acquisition with 14 m/z increment in precursor isolation window resulted in 58 MS2 across the 350-1150 m/z range. The accumulation time was set to 0.1 s for MS1 scan and 65 ms for MS2 scan. The total cycle time was approximately 3.8 s.

RESULTS

Library generation:
The peptide spectral library employed in this study was generated by shotgun analysis using the same cell lysate which was analyzed by SWATH acquisition. From the shotgun analysis, the quantified peptides distribute along the whole chromatographic gradient, indicating the chromatographic conditions were good for these samples(Figure 1) A peptide spectral library consisting of 4035 protein groups and 29945 unique peptides (FDR<1% on both peptide and protein leve ls) was established(Figure 2).

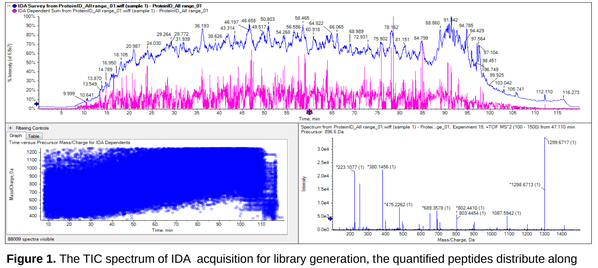

Figure 1. The TIC spectrum of IDA acquisition for library generation, the quantified peptides distribute along the whole chromatographic gradient.

Proteins Identified at Critical False Discovery Rates				Peptides Identified at Critical False Discovery Rates			
Number of Proteins Detected				*Number of Peptides Identified*			
Critical FDR	**Local FDR**	*Global FDR*	**Global FDR from Fit**	**Critical FDR**	**Local FDR**	*Global FDR*	**Global FDR from Fit**
1.0%	*3588*	*4021*	**4035**	**1.0%**	*24097*	*29874*	**29945**
5.0%	**3767**	*4368*	*4388*	**5.0%**	**27745**	*34527*	*34553*
10.0%	**3883**	*4619*	*4655*	**10.0%**	**29337**	*37762*	*37729*

** It is recommended you use numbers in bold and avoid using numbers in italics.* ** It is recommended you use numbers in bold and avoid using numbers in italics.*

Figure 2. The unique proteins and peptides identified from IDA acquisition

Quantification Reproducibility and Variability
To assess the reproducibility of SWATH acquisition, we first compared the overall chromatographic elution profiles. Based on the total ion chromatogram overlay (Figure 3), the samples with the same sample loading amount generated quite similar chromatographic profiles, overlapping completely. (Figure 3) The total ion intensities got increased with the growth of injected peptide quantity（0.25ug, 0.5ug, 1ug, 2ug peptides loaded on column, 3 replicates for each amount level ）.

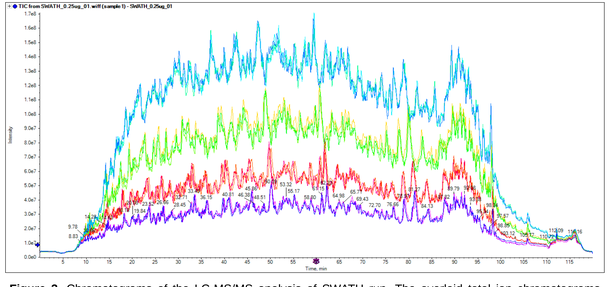

Figure 3. Chromatograms of the LC-MS/MS analysis of SWATH run. The overlaid total ion chromatograms (TICs) of all twelve measurements, including four samples with varied protein amounts and three technical replicates for each samples. The reproducibility between different technical replicates is good

Next looked into the peptides and transition ions used in the quantificationIn, total 3503 proteins got quantified across all twelve measurements, including four different sample loadings and corresponding three technical replicates. Over 70% of the total proteins have more than two peptides quantified, and on average, 3.2 peptides can be matched to a protein. By further evaluated the quantification reproducibility and variability on individual proteins between three technical replicates. The peak area values obtained in three replicates were plotted in 3-D mode. (Figure 4 a-d)

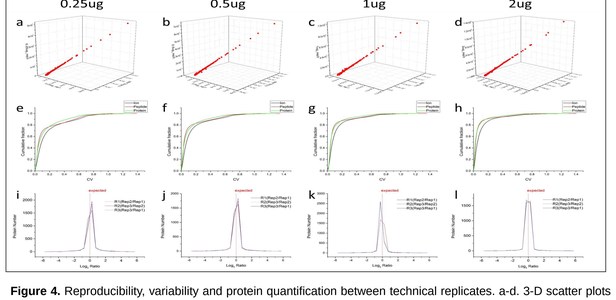

Figure 4. Reproducibility, variability and protein quantification between technical replicates. a-d. 3-D scatter plots for protein ion intensities measured in three replicates. For each sample, any two replicates show a very good linear correlation. e-h. Coefficient of variation (CV) of ion intensities on transition ion (black), peptide (red) and protein (green) levels. In general, larger sample loading results in smaller variance. i-l. The distributions of quantification results between any two technical replicates. The frequencies of Log2 ratios derived from three comparisons (replicate2 vs. replicate1, replicate3 vs. replicate2, and replicate3 vs. replicate1) were plotted for each samples. 1:1 ratios were observed on all four samples.

Quantification Precision and Accuracy
The comparison between three technical replicates for each sample represents the quantification scenario when protein abundances have 1:1 ratios. The Log_2 ratios derived from three comparisons (replicate2 vs. replicate1, replicate3 vs. replicate2, and replicate3 vs. replicate1) for each sample were plotted in Figure 3 i-l. The averages of Log_2 ratios are 0.03, 0.01,-0.05, and 0.00 respectively for each samples, which are all perfectly close to the theoretical value of 0. The result shown above demonstrates that the SWATH acquisition has very high quantification precision and accuracy when protein abundance ratios are about 1:1. We next investigated the quantification accuracy of SWATH-based label-free analysis in which protein abundances differ between samples. The result shows below (Figure 5). The practical fold changes are underestimated due to matrix effects caused by different sample load amount , but the result still can reflect the trending of protein amount across different samples

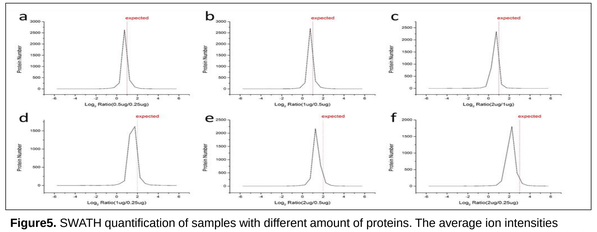

Figure5. SWATH quantification of samples with different amount of proteins. The average ion intensities derived from three technical replicates were used in these plots. The red dashed lines indicate the theoretical ratios. a-f. The distributions of Log_2 ratios between samples of 0.5 µg and 0.25 µg (a), 1 µg and 0.5 µg (b), 2 µg and 1 µg (c), 1 µg and 0.25 µg (d) 2 µg and 0.5 µg (e), 2 µg and 0.25 µg (f)

CONCLUSIONS

Our results demonstrated that the SWATH-based quantitative strategy could provide outstanding measurement reproducibility. The quantitative data showed that the protein abundances could be relatively quantified very precisely, and the reproducibility of the measurements was great. Therefore the SWATH acquisition can still result in reliable and valuable protein quantitative information. In summary, SWATH acquisition can provide precise label-free quantification on proteome-scale.

REFERENCES

1Leitner, A.; Aebersold, R., SnapShot: mass spectrometry for protein and proteome analyses. *Cell* **2013,** 154, (1), 252-252 e1.
2Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M. C.; Yates, J. R., 3rd, Protein analysis by shotgun/bottom-up proteomics. *Chem Rev* **2013,** 113, (4), 2343-94.
3Mann, M.; Kulak, N. A.; Nagaraj, N.; Cox, J., The coming age of complete, accurate, and ubiquitous proteomes. *Mol Cell* **2013,** 49, (4), 583-90.
4Nagaraj, N.; Wisniewski, J. R.; Geiger, T.; Cox, J.; Kircher, M.; Kelso, J.; Paabo, S.; Mann, M., Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol* **2011,** 7, 548.

TRADEMARKS/LICENSING