Extending the Depth of Coverage in SWATH® Acquisition with Deeper Ion Libraries

Joerg Dojahn[1], Nick Morrice[2], Dietrich Merkel[1], Christie Hunter[3]
[1]SCIEX Germany, [2]SCIEX UK, [3]SCIEX USA

INTRODUCTION AND OVERVIEW

SWATH® Acquisition has become a powerful tool in global protein discovery and quantitation and previous work showed that the extended dynamic range of the TripleTOF® 6600 system and its fast MS/MS acquisition speeds in combination with variable window acquisition increased the depth of coverage[1,2].

A less explored variable on depth of coverage is the impact of the size of the ion library. Previous studies have suggested that more quantitative data is available in a SWATH® Acquisition data files than is accessed from a small ion library generated from 1D data dependent acquisition experiments (Figure 1). Here, the impact of using larger and deeper ion libraries on the numbers of peptides and proteins detected was explored and the quality of quantitation observed by processing the same data files from digested yeast as well as human cell lysates with increasingly deeper ion libraries was assessed. Even as more data is extracted from the data files using the deeper libraries, the quality of quantitation is maintained even into the low abundant protein/peptide regime. Depending on the depth of the ion library needed and the available time for analysis, a researcher can balance the library generation effort required for a particular biological system.

MATERIALS AND METHODS

2D LC-MS/MS Fractionation:
A digest of a K562 human cell lysate (Promega) was fractionated using high pH reverse phase chromatography using a Shimadzu Nexera system (UV detection at 214nm). A Durashell RP column (250 x 4.6mm, 5um, Agela Technologies) operating at 1 mL/min was used, running a gradient of 2 – 90% acetonitrile in 2mM ammonium hydroxide. 15 fractions were collected every 2 mins.
Each subsequent peptide fraction was then separated using low pH reverse phase gradient on the NanoLC™ 425 system operating in microflow mode. A Triart C18 150 x 0.3mm column (YMC) was used at 5 µL/min flow rate with a 45 min gradient from 2-40% acetonitrile in 0.1% formic acid. The eluent was analyzed using the TripleTOF® 6600 system equipped with a DuoSpray™ Source and a 25 µm ID electrode. Data dependent acquisition was performed with 30 MS/MS per cycle, each with 50 msec accumulation.
1D LC-MS/MS Chromatography:
Separation of a trypsin digest of HEK human cell lysate or yeast cell lysate was performed on a NanoLC™ 425 with the cHiPLC® System operating in serial column mode. Two 75 µm x 15 cm ChromXP™ cHIPLC columns (SCIEX) were used with a long gradient (5-30% acetonitrile, 0.1% formic acid in water) at 300 nL/min. Gradient lengths for human analyses were 60 mins for 1D library generation and 180 mins for SWATH acquisition. For yeast, the gradient length for both 1D library generation and for SWATH acquisition was 120 mins.
Mass Spectrometry:
The MS analysis was performed on a TripleTOF® 6600 system (SCIEX) using a NanoSpray® Source (SCIEX). Information dependent acquisition was acquired on the fractionated and unfractionated lysates, using the 30 MS/MS per cycle with 50 msec accumulation time. Variable window SWATH® Acquisition method was built using Analyst® TF Software 1.7 using 100 variable windows and 25msec accumulation times.
Data Processing:
1D and 2D libraries were created by processing the IDA data using ProteinPilot™ Software. SWATH Acquisition data was processed using SWATH® 2.0 Micro App in PeakView® Software 2.2 using the various ion libraries excluding modified and shared peptides. Results analysis was performed in Excel using the SWATH Acquisition Replicates template. All protein and peptide numbers reported were determined at <1%FDR and <20% CV across the 5 replicates collected.

RESULTS

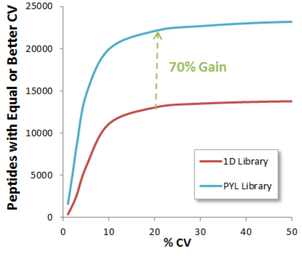

Figure 1. Gain in Peptides Quantified in Yeast Using a Deeper Ion Library. SWATH® Acquisition was collected on a yeast cell lysate using a 2 hour gradient and 3 µg total protein and using a 100 variable window method[1]. The data was analyzed using both a 1D ion library as well as the Pan Yeast library[4] and the numbers of peptides quantified was assessed. Using a 1% FDR filter at the peptide level and computing the %CV for 5 replicate injections, the cumulative %CV curves can be plotted. Using a 20% CV as a cutoff, a gain of 70% quantified peptides was observed when using the larger Pan Yeast library.

Table 1. SWATH® Acquisition Ion Libraries for Human. A number of ion libraries of increasing size were created for interrogation of SWATH acquisition data (at 1% FDR, excluding modified and shared peptides).

Library	# of Proteins	# of Peptides
1D - HEK	*3371*	*12809*
2D – K562	*6889*	*62329*
Pan Human[3]	*14425*	*197585*

Library	# of Proteins	# of Peptides
1D - Yeast	*2772*	*15627*
Pan Yeast[4]	*4658*	*87535*

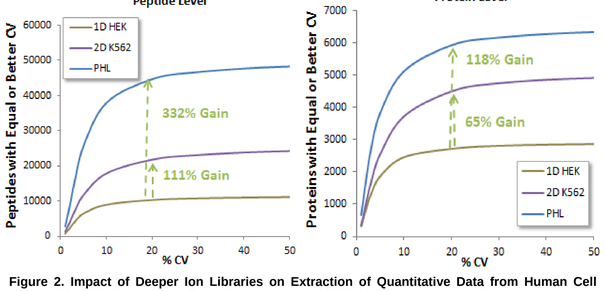

Figure 2. Impact of Deeper Ion Libraries on Extraction of Quantitative Data from Human Cell Lysate SWATH® Acquisition Data. The same set of SWATH replicates from a HEK cell lysate were processed with increasingly larger libraries and the proteins/peptides quantified with <1% FDR peptides and CVs were plotted. Significant gains in both peptides and proteins quantified at <20% CV were observed as the size of the ion library used for targeted data processing increased.

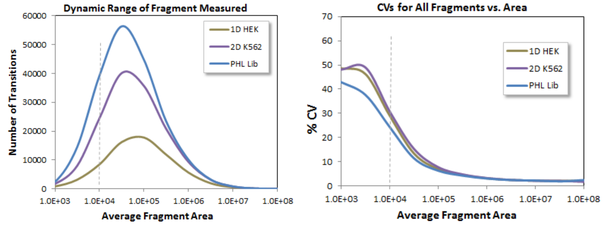

Figure 3. Dynamic Range and Reproducibility of the Increased Quantification with Deeper Libraries. As more peptides are quantified at lower and lower abundance due to the use of increasingly larger ion libraries, it is important to monitor variance in quantitation (%CV across 5 replicates). Here the bulk of fragment ions quantified are above a peak area of 1e4 and the CV vs intensity plot shows that the bulk of the fragment ions have CVs in the 10-20% CV range.

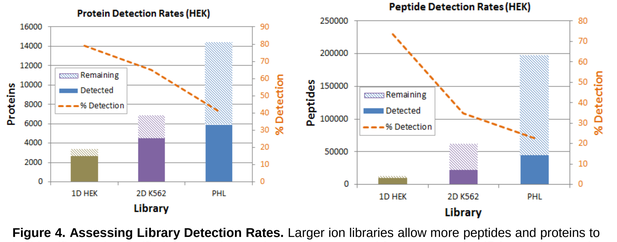

Figure 4. Assessing Library Detection Rates. Larger ion libraries allow more peptides and proteins to be identified and quantified, but need more time to be generated from consecutive IDA experiments. As the size of the ion library increases, the number of peptides and proteins quantified increases. Here the total # of peptides/proteins quantified at <1%FDR and <20% CV (solid bars) are plotted along with the portion of the library that was not quantified (not detected at 1% FDR or failed the stringent quantitation filters, hashed bars). The % detection of the entire library used can then be computed (orange). On the small library, 70-80% of proteins in the library are quantified, whereas only 40% of the protein forms are quantified from the pan human library.

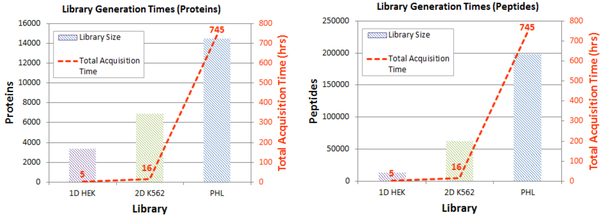

Figure 5. Balancing Peptide and Protein Library Size with Library Generation Time. Based on the library depth needed, the total library generation time can be balanced by the number of IDA experiments, the amount of fractionation used, the number of related cell types, nanoflow vs microflow LC, etc. *Note* that the computation of library generation time includes the load, wash and re-equilibration times of the LC system.

CONCLUSIONS

- High quality SWATH® Acquisition can be acquired using Variable Window acquisition and 100 Q1 windows on TripleTOF® Systems
  - Contain a tremendous amount of quantitative data
  - To fully mine the data, deep ion libraries are key for targeted data processing
- Here, significant gains in peptides and proteins quantified (332% and 118% gains, respectively) when transitioning from a simple 1D ion library, to a 2D library, to a pan human library (PHL)
  - Choose optimal experimental conditions for SWATH acquisition
  - Choose an appropriate ion library strategy that matches the depth of coverage required for the biology under study and the time needed to create library
- For the HEK nanoflow data file investigated here, 5922 proteins were quantified at <20% CV when processed using the very comprehensive pan human library.

REFERENCES

1. Improved Data Quality Using Variable Q1 Window Widths in SWATH® Acquisition, SCIEX Technical note RUO-MKT-02-2879-A.
2. Increasing Depth of Coverage in Data Independent Acquisition with Higher Sample Loads and Smaller Q1 Windows, SCIEX Technical note RUO-MKT-02-3245-A.
3. Rosenberger G *et al.* (2014) *Scientific Data*, **1**, 140031.
4. Selevsek, N., Chang, C.-Y., Gillet, L.C., Navarro, P., Bernhardt, O.M., Reiter, L., Cheng, L.-Y., Vitek, O., Aebersold, R., (2015) *Mol Cell Prot.* **14**, 739-749.

TRADEMARKS/LICENSING