

Rapid analysis and interpretation of metabolomics SWATH acquisition data using a cloud-based processing pipeline



Arianna Jones¹, Alexandra Antonoplis¹, Jason Causon², Christie Hunter¹
¹SCIEX USA, ²SCIEX CANADA

INTRODUCTION

SWATH acquisition is a data independent acquisition (DIA) workflow that has been demonstrated to improve metabolite coverage over traditional data dependent techniques for untargeted metabolomics.^{1,2} The workflow enables creation of a digitized record of the metabolome present in a sample, with full (MS1) and MS/MS scans capturing every detectable analyte in a single injection. The richness of SWATH acquisition data offers numerous data analysis opportunities, one of which is to identify differential metabolites across sample groups. While instrumentation and appropriate methods for collecting metabolomics SWATH acquisition data are well-established, the lack of software tools for processing large-scale metabolomics SWATH acquisition studies remains a challenge for wide-spread adoption of the workflow in laboratories.³ The OneOmics suite, a cloud-based solution for SWATH acquisition data processing, enables fast and confident identification of differential metabolites across experiment groups in large-scale metabolomics SWATH acquisition studies. The entire processing workflow, from meta data management to examining biological significance of results, is supported in the platform

In this work, the OneOmics suite was used to investigate analytes identified in the urine of Zucker Diabetic Fatty (ZDF) vs. control Sprague Dawley (SD) rats from SWATH acquisition data collected on the SCIEX ZenoTOF 7600 system. The OneOmics suite facilitated complete end-to-end processing of the metabolomics data set, starting with ion-library driven extraction of analyte peak groups from the SWATH acquisition data. Built-in false discovery rate (FDR) analysis and normalization algorithms facilitated accurate identification of differentially expressed analytes. Data sets were further interrogated using multivariate statistical analysis tools and viewed in a biological context with pathway mapping. Differentially regulated metabolites in the urine of ZDF versus healthy rats were identified and compared with previous metabolomic studies of diabetic rats to interpret the findings.

MATERIALS AND METHODS

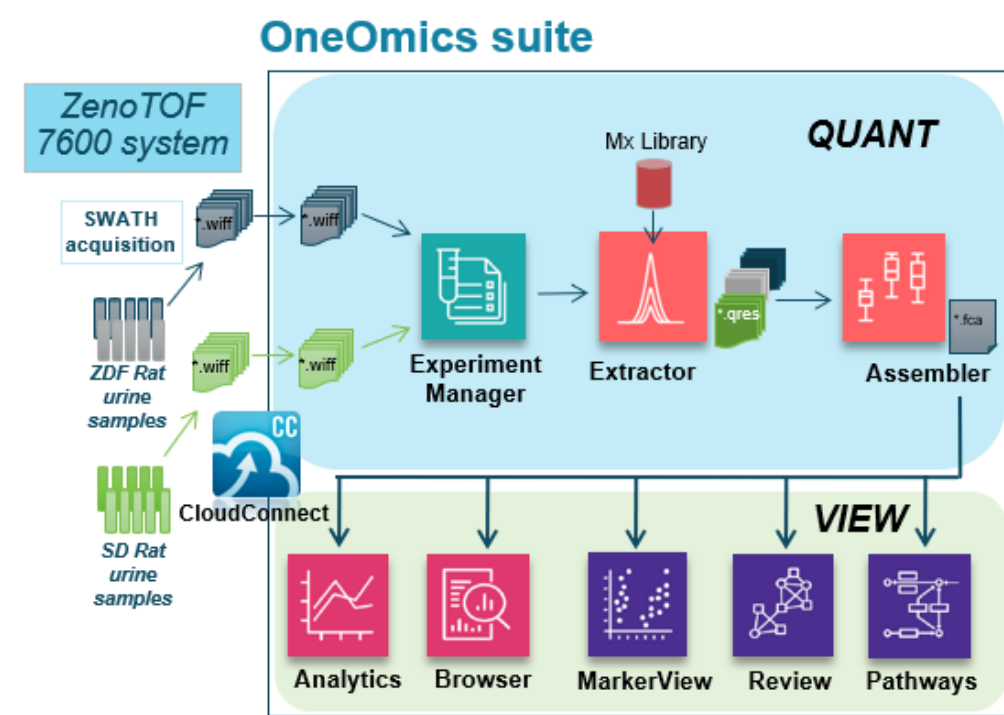
Sample preparation: Urine samples were collected from four distinct rat groups: male ZDF rats, female ZDF rats, male SD rats, and female SD rats. Urine was collected from N=5 rats per group and diluted 10-fold in mobile phase A prior to LC-MS/MS analysis.

Chromatography: An ExionLC AD HPLC system (SCIEX) with a Phenomenex Luna Omega Polar C18, 3 μ m 150 x 2.1 mm (00F-4760-AN) at 40 °C was used with a flow rate of 300 μ L/min. The injection volume was 2 μ L.

Mass spectrometry: SWATH acquisition data was collected using a SCIEX ZenoTOF 7600 system and 80 variable Q1 windows. For the TOF MS survey scan, an accumulation time of 50 ms was used, while an accumulation time of 10 ms was used for MS/MS scans.

Data processing: CloudConnect in PeakView software 2.2 was used to upload data files to the OneOmics suite. The study meta data was assigned using the Experiment Manager App. Analyte peak groups were then extracted in the Metabolomics App using an extracted ion chromatogram (XIC) extraction width of 40 ppm with 6 transitions required per analyte. Data were examined in the Bioreviews App for identification of enriched analytes and pathway mapping.

Figure 1. Overview of the metabolomics workflow for SWATH acquisition data processing in the OneOmics suite. Following collection of variable window SWATH acquisition data using the SCIEX ZenoTOF 7600 system, data files were uploaded to the OneOmics suite for processing. Experimental meta data was entered, and then the Extractor and Assembler Apps were used to extract and quantify metabolites across the samples. Additional apps were used to visualize the results and perform statistical analysis to determine the metabolite differences between ZDF vs. SD rat urine.



EXAMINING DATA QUALITY IN THE ANALYTICS APP

Using the Experiment Manager App, the SWATH acquisition data files were organized into groups of technical replicates to assess the analytical reproducibility of the LC-MS/MS workflow. The technical replicate groups were then processed using the Extractor App. The Extractor App extracts and integrates analyte precursor and fragment XIC areas using an ion library. The resulting peak groups are then scored according to chromatographic attributes, such as peak width and peak intensity ratio, and spectral attributes, including mass error of the monoisotopic peak and MS/MS m/z error. These target analyte scores are then incorporated into an FDR analysis within the application to assess the statistical accuracy of the detections. The FDR analysis follows a target-decoy approach, in which decoy analytes are generated in silico by selecting precursor and decoy fragment ions from the ion library used for data processing according to predetermined criteria.⁴ The resulting distribution of target vs. decoy is easily visualized in the Analytics App (Figure 2, top).

Figure 2. Assessing analytical reproducibility using the Analytics App. Top: for each data file, an FDR analysis is performed to identify analytes that pass a 1% FDR cutoff filter. Within the Analytics App, the distribution of scores for both target and decoy analytes can be viewed for each data file, along with the number of analytes passing the 1% FDR cutoff threshold. Bottom: the peak area reproducibility of analytes with at least one confident detection (<1% FDR) across the technical replicates was assessed. The vertical line marks the number of analytes passing a 20% CV filter.

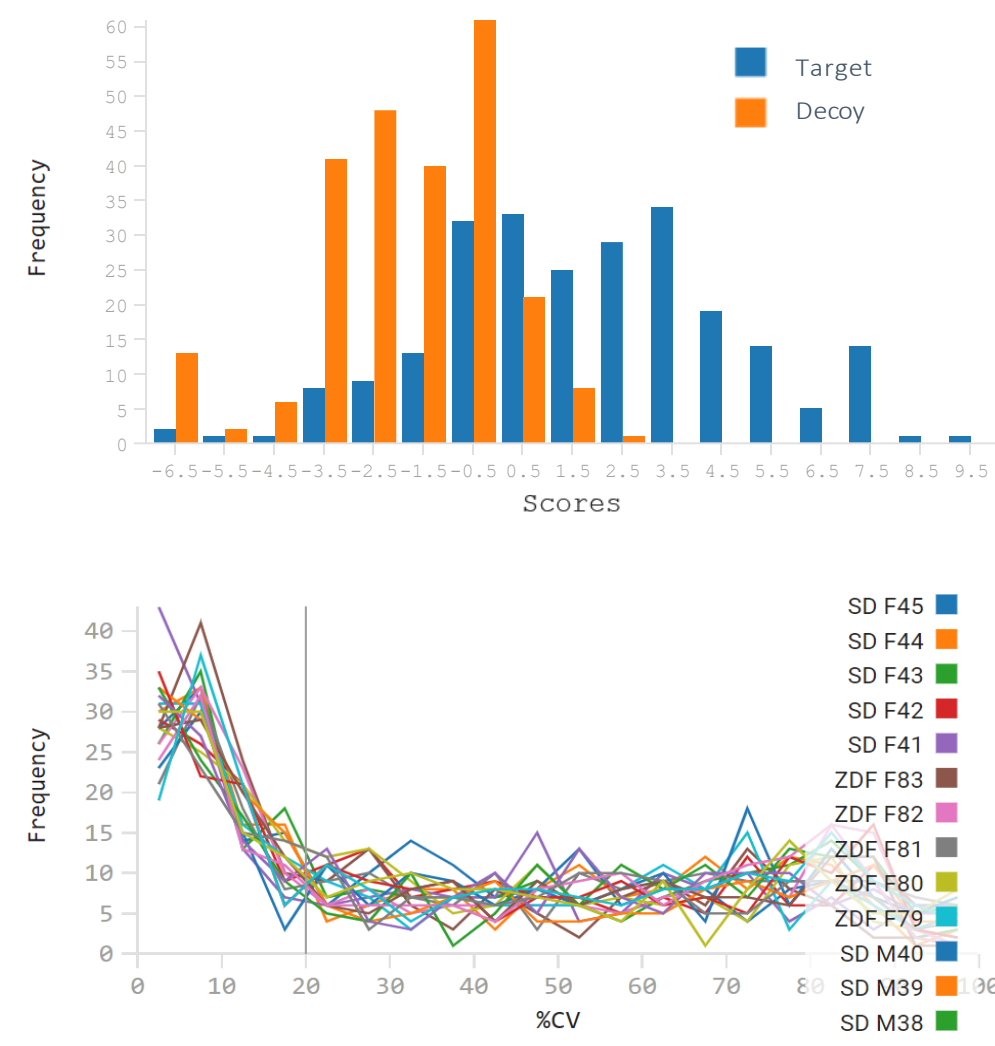


Figure 3. Analytes identified in urine of ZDF rats as compared to SD rats. (Top) Heatmap of analytes quantified can be generated in Browser for rapid visualization of the changes observed for the 236 detected analytes as compared to a control sample group. For visualization, a 75% fold-change confidence filter and MLR reproducibility filter of 0.15 were applied, 84 metabolites passed these filters and thus had significant change. The intensity of the color (orange for up-regulation, blue for down-regulation) indicates the magnitude of the log2 signed fold change, and a subset of the heatmap is shown, as visualized in the plot on the right. (Bottom) Selecting an individual analyte in the heat map (indoxyl) enables viewing of its associated normalized transition peak areas across all samples. Each color represents a different transition, and overall normalized analyte areas can also be viewed for each sample.

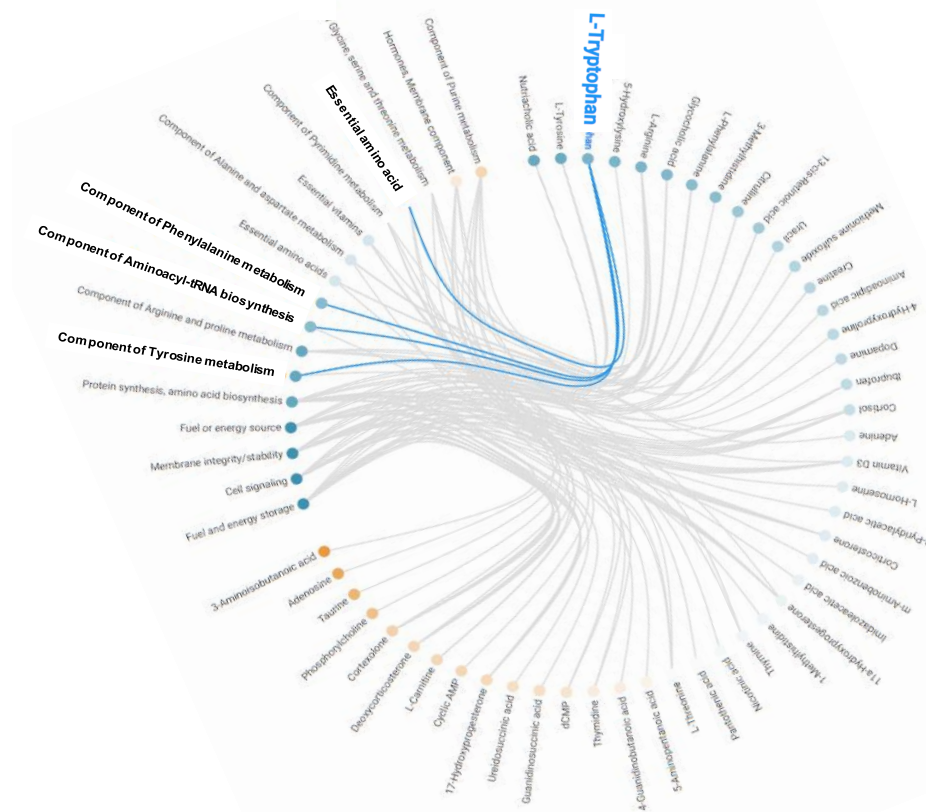
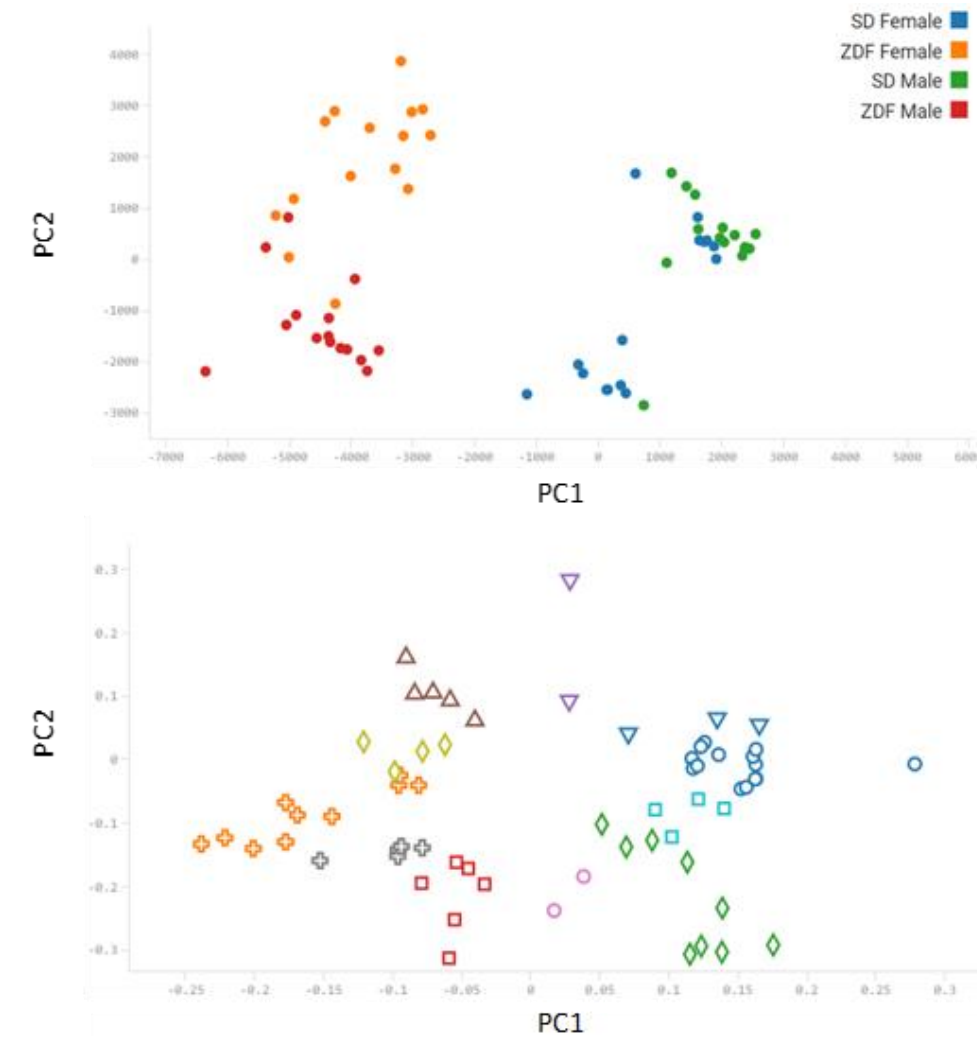


Figure 4. Exploring individual analytes in Browser. Shown at the top are biofunction ontology terms for L-tryptophan, a differential analyte in ZDF vs. SD rats. Hovering over an analyte of interest in the ontology wheel reveals its associated ontology terms, which are provided by the HMDB. Hovering over an ontology will highlight all the metabolites found within that ontology. The color of the dot highlights the direction of the fold change (positive fold change = orange, negative fold change = blue, no fold change = white).

EXPLORING DIFFERENTIAL ANALYTES IN THE BROWSER APP

Next, the data files were processed according to rat type (diabetic vs. healthy) and sex (female vs. male) to explore up-regulated and down-regulated metabolites. The Browser App features a dashboard of visuals to showcase the direction, extent, and confidence of detected metabolite fold changes, as well as ontology information from the Human Metabolome Database (HMDB) for each differentially expressed analyte. Metabolite fold changes in the Browser App can further be interrogated in heatmap form, with processing filters to visualize detections with high confidence and reproducibility (Figure 3). As compared to control rats, several highly differential analytes were identified in the ZDF rats across both sexes. The analytes exhibiting the greatest log2 signed fold change values were indoxyl (6.521), xanthosine (4.536), and L-tryptophan (-3.02). Indoxyl sulfate is a known uremic toxin that has been detected in the urine of type 2 diabetes patients with low renal function, and xanthosine has been identified as a urinary metabolite biomarker of type 2 diabetic nephropathy.⁵ In a metabolomics analysis of the urine of ZDF rats, tryptophan was also found to be down-regulated.⁵ The identified analytes can also be correlated to their biofunction in the Ontology section of Browser (Figure 4). The analyte L-tryptophan was linked to its role in amino acid metabolism using the ontologies provided by HMDB, highlighting processes that might be affected by its down-regulation in ZDF rat urine as compared to SD rat urine.

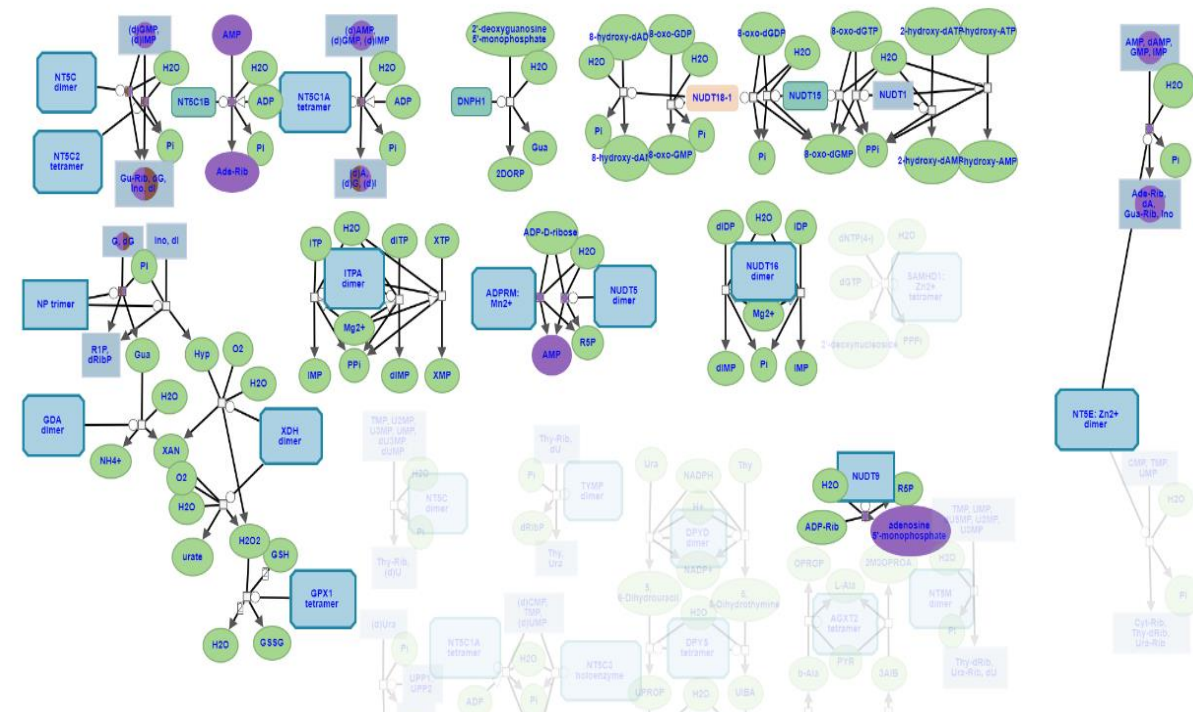
Figure 5. Using the Bioreviews App for multivariate statistical analysis. (Top) PCA analysis revealed distinct clustering of the sample groups, with healthy SD rats forming one distinct subset and ZDF rats forming a second subset. (Bottom) The loadings plot shows the distinct PCVG groups of analytes, clustered according to similar patterns of change across samples.



USING THE BIOREVIEWS APP

In the Bioreviews App, the MarkerView App can be used to perform multivariate statistical analysis of quantitative metabolomics results, including principle components analysis-principle component variable grouping (PCA-PCVG) analysis. For these analyses, a PCA Scores plot of the sample groups indicated significant differentiation between the ZDF rat and SD rat sample groups (Figure 5, top). The Loadings plot shows the PCVG groups, clustering metabolites with similar patterns of expression (Figure 5, middle). After performing multivariate statistical analysis, analyte cluster groups can be mapped to biological pathways in the Pathways App (Figure 6). Pathways analysis is conducted using Reactome, and the application enables visualization of the enriched biological and chemical pathways from the results. The Pathway Browser revealed significant enrichment of the purine catabolism pathway in the data set, in which purine bases are converted to xanthine. This pathway was previously identified as enriched in a study of the urinary metabolome of Zucker diabetic rats as compared to healthy control rats.⁵ Increases in the analyte xanthosine in urine have been linked to increased purine catabolism, and xanthosine has been observed to be significantly increased in the urine of diabetic rats.⁵ Increased purine catabolism has been proposed as a homeostatic response of mitochondria to oxidative stress.⁵

Figure 6. Pathway mapping with Bioreviews App. Mapping of PCA-PCVG cluster groups that contained metabolite differences to biological pathways revealed enrichment of the purine catabolism pathway (p = .0049). Green ovals represent small molecules and outlined blue boxes represent proteins and protein complexes. Analytes identified from the data set, including adenosine, guanosine, and deoxyguanosine, are shown in purple and brown. Xanthosine triphosphate (XTP – marked with star) is also an intermediate in this pathway. Xanthosine was found to be significantly enriched in the urine of ZDF rats in the SWATH acquisition data set.



CONCLUSIONS

In this study, the OneOmics suite was used to process a metabolomics data set acquired using SWATH acquisition on the SCIEX ZenoTOF 7600 system and to detect differentially regulated metabolites in the urine of ZDF versus healthy rats. Results were compared with previous metabolomic studies of diabetic rats to interpret the findings. The complete integration of data processing applications in the cloud-based platform enabled rapid processing of the study data. Built-in FDR analysis enabled library-driven peak identifications to be assessed for statistical accuracy. The platform also features tools for exploring the biological relevance of detected analytes, including ontology enrichments and pathway mapping.

REFERENCES

- SWATH acquisition improves metabolite coverage over traditional data dependent techniques for untargeted metabolomics. [SCIEX technical note, RUO-MKT-02-7128-A](#).
- Improved metabolite identification using data independent analysis for metabolomics. [SCIEX technical note, RUO-MKT-02-10617-A](#).
- Raetz M, Bonner R, Hopfgartner G. (2020) SWATH-MS for metabolomics and lipidomics: critical aspects of qualitative and quantitative analysis. [Metabolomics 16, 71](#).
- How does FDR scoring work in the OneOmics suite for metabolomics? [SCIEX community post RUO-MKT-18-13219-A](#).
- Rapid analysis and interpretation of metabolomics SWATH acquisition data using a cloud-based processing pipeline. [SCIEX technical note, RUO-MKT-02-13056-A](#).

TRADEMARKS/LICENSING

The SCIEX clinical diagnostic portfolio is For In Vitro Diagnostic Use. Rx Only. Product(s) not available in all countries. For information on availability, please contact your local sales representative or refer to www.sciex.com/diagnostics. All other products are For Research Use Only. Not for use in Diagnostic Procedures. Trademarks and/or registered trademarks mentioned herein, including associated logos, are the property of AB Sciex Pte. Ltd. or their respective owners in the United States and/or certain other countries (see www.sciex.com/trademarks).
© 2021 DH Tech. Dev. Pte. Ltd. RUO-MKT-02-13056-A