Proteomics in Multi-omics Workflows Using Yeast as a Model System

Application Note

Authors

Christine A. Miller, Stefan Jenkins, Theodore R. Sana, and Steven M. Fischer

Agilent Technologies, Inc.

Santa Clara CA

Abstract

Pathway mapping and visualization allows protein scientists to focus research in active biological areas. Mapping the results of a traditional discovery proteomics experiment onto pathways can more rapidly narrow the list of protein targets. Because of the interconnectivity of metabolites, proteins and genes, any omic experiment can be the starting point for the next experiment using pathways as a framework to interpret the results and propose further experiments. This application note demonstrates both a protein discovery-based workflow and a metabolomics-directed targeted workflow.

Agilent Technologies

Introduction

A key step in proteomics research is identification of biologically significant differential proteins using LC/MS/MS. Improvements in instrumentation and fractionation techniques allow the efficient and comprehensive protein identification for many proteomes. However, transforming these protein identifications into a biological context is an essential step in understanding the results from differential expression experiments.

This application note describes a workflow that allows protein identification results to be analyzed using a variety of statistical tools, and then visualized on pathways. Based on this analysis and visualization, lists of proteins of interest can be exported for subsequent targeted experiments. Targeted proteomics methods can be developed based on information from genomics, metabolomic, or proteomic discovery experiments.

Using these results, an MRM-based assay can be created and the quantitative results from this targeted analysis can be imported into the same software for statistical analysis and pathway visualization, thus facilitating both discovery and verification. This application note demonstrates a protein discovery workflow and a metabolomic pathway-directed targeted workflow using a yeast sample set.

Experimental

Sample preparation

Baker's yeast, *Saccharomyces cerevisiae* strain BJ5459, was cultured (0.8 L) in parallel and grown to an OD_{600} of 0.8 before treatment. The immunosuppressants cyclosporin A and FK506 (4 µg/mL in 90:10 ethanol:Tween 20) were used to stress the cultures. Cultures were exposed to vehicle control (4 mL of 90:10 ethanol:Tween 20) for wild type (WT) and calcium control (CC) cultures, or FK506 (FK), or cyclosporin A (CA) for 1 hour. An equal fraction of media containing $CaCl_2$ was then added to FK, CA, and CC cultures to a final concentration of 100 mM. After 15 minutes of exposure to either vehicle control or calcium, the cultures were centrifuged and washed with phosphate buffered saline (PBS). Quenching was done with methanol, then the final samples were lyophilized and stored at –80 °C until use. Just prior to use, the pellets were resuspended to 20 mg/mL in 10 % glycerol in water and a 25 µL aliquot was subjected to protein extraction and digestion. The aliquot was diluted 1:10 in 4 % SDS in 100 mM TrisHCl with 100 mM dithiothreitol (DTT), then heated to 95 °C for 5 minutes. After cooling to room temperature, approximately 100 µL of acid-washed glass beads (Sigma) was added to each tube and the samples were vortexed vigorously for 30 seconds then placed on ice. This was repeated three times to ensure adequate lysis and protein extraction, then the extract was carefully removed to a fresh tube. A BCA protein assay (Pierce) was done for each culture condition to determine protein concentration. Aliquots from each culture lysate were subjected to trypsinization using the FASP II protocol[5]. After digestion, samples were stored at –20 °C until LC/MS analysis.

LC/MS analysis

For this study, protein discovery data was acquired on the Agilent 6550 iFunnel Q-TOF using data dependent acquisition while targeted analysis data was acquired on the Agilent 6490 Triple Quadrupole LC/MS system with retention-time scheduling using dynamic MRM mode. Both the 6490 Triple Quadrupole LC/MS system and 6550 iFunnel Q-TOF spectrometers incorporate iFunnel technology to achieve high sensitivity for detection of peptides. The microfluidic HPLC-Chip/MS technology, which allows routine and robust nano scale separations, was interfaced to the mass spectrometer. The HPLC-Chip system was comprised of an Agilent 1200 Series LC System consisting of a binary nano pump, thermostatted wellplate sampler, and binary capillary pump.

Since biological replicates were not available for this experiment, each cell lysate was analyzed four times. For the protein discovery workflow, a 120-minute gradient method was used while the targeted triple quadrupole method used a shorter 90-minute gradient method. The HPLC-Chip was the Agilent Polaris chip (G4240-6230), which is comprised of a 360 nL enrichment column and a 150 mm × 75 µm analytical column. Both were packed with Agilent Polaris C18-A, 180 Å, 3 µm material.

Data Analysis

The Q-TOF data was first processed by Agilent Spectrum Mill software against the UniProt yeast database (downloaded on January 24, 2013) and validated at the spectral level using 1.2 % FDR as the criteria. During the spectral extraction process, Spectrum Mill Software determines the area for each precursor, and then saves this information for use after peptide spectral matching is completed. These peptide areas are then assigned to the protein match and the peptide areas for a given protein are summarized as either the total or mean intensity. Protein results can then be summarized in a protein-protein comparison mode which first evaluates all replicates of all samples on the peptide level and groups proteins across the entire data set, then rolls the peptides up into protein groups and creates a table showing the summarized protein results with the protein intensities. These results were imported for evaluation in Mass Profiler Professional 12.6 (MPP), a data analysis and visualization software and pathway analysis was conducted using MPP's optional Pathways Architect software module.

Spectrum Mill Software includes tools to aid in the creation of triple quadrupole MRM-based methods. MRM Selector generates MRM methods from protein database search results, and Peptide Selector predicts peptides and product ions for proteins based on various filtering criteria. These tools were used in the development of the MRM-based assay. Skyline, a freely-available Windows client application[3], was used for method development, optimization, and quantitation in the targeted proteomics workflow. Multiple MRMs analyses were performed to develop the optimized dynamic MRM (DMRM) method. In Skyline, the SSRCalc retention-time modeling[2] and the dot product between the spectral library and MRM transition areas were employed to ensure confident peak assignment. Area results for all targeted peptides were imported into MPP for assessment of statistical significance and pathway visualization.

HPLC-Chip conditions	
HPLC-Chip	Agilent Polaris-HR-Chip-3C18 with a 360 nL enrichment column and 150 mm × 0.075 mm analytical column
Mobile phase A	0.1 % formic acid in water
Mobile phase B	0.1 % formic acid in 90 % acetonitrile in water
Loading	2 µL/min with 3 % B
Analytical flow rate	300 nL/min
Q-TOF analytical gradient (120 minutes)	3 % B initial to 25 % B at 90 minutes, 40 % B at 120 minutes, 90 % B at 125 minutes for 5 minutes, and 3 % B at 130.1 minutes
Triple quadrupole analytical gradient (90 minutes)	3 % B initial to 40 % B at 90 minutes for 10 minutes, 3 % B at 100.1 minutes
Injection volume	adjusted for each culture lysate based on protein assay to load 1 µg of total protein per injection
Agilent 6550 Q-TOF MS conditions	
Instrument mode	Extended dynamic range (2 GHz) with low mass range (1,700 *m/z*)
Drying gas	11 L/min, 250 °C
Acquisition rate	8 (MS) and 3 (maximum rate in MS/MS) spectra/sec
Acquisition range	275–1,700 (MS) and 50–1,700 (MS/MS)
Collision energy	slope of 3.1 and intercept 1 (+2) or slope of 3 and intercept 3.6 (+3 and higher)
Isolation	Narrow (~1.3 amu)
Data dependent acquisition	20 precursors per cycle using precursor abundance based acquisition rate with accumulation time limit enabled; active exclusion after one spectrum for 0.5 minutes
Agilent 6490 Triple Quadrupole MS conditions	
Drying gas	11 L/min, 150 °C
Cell Acc	4 V
Collision energy	optimized for each transition
Cycle time	750 ms
Delta EMV	100 V
Time filter	0.1 minutes

Results and Discussion

Protein Discovery Workflow

The protein discovery workflow summarized in Figure 1 was followed with the digested lysates from the treated yeast cultures. Spectrum Mill search results resulted in a total of 3,446 distinct proteins and 13,616 unique peptides being identified in the sample set using validation criteria of 1.2 % FDR at the spectral level. In order to see true differences in protein level between the culture conditions, the same amount of total protein was injected for each condition. In addition, the Spectrum Mill protein-protein comparison summary

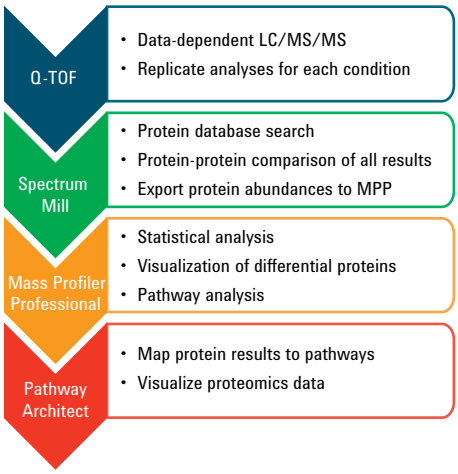

Figure 1. Workflow for mapping and visualizing protein discovery results on pathways.

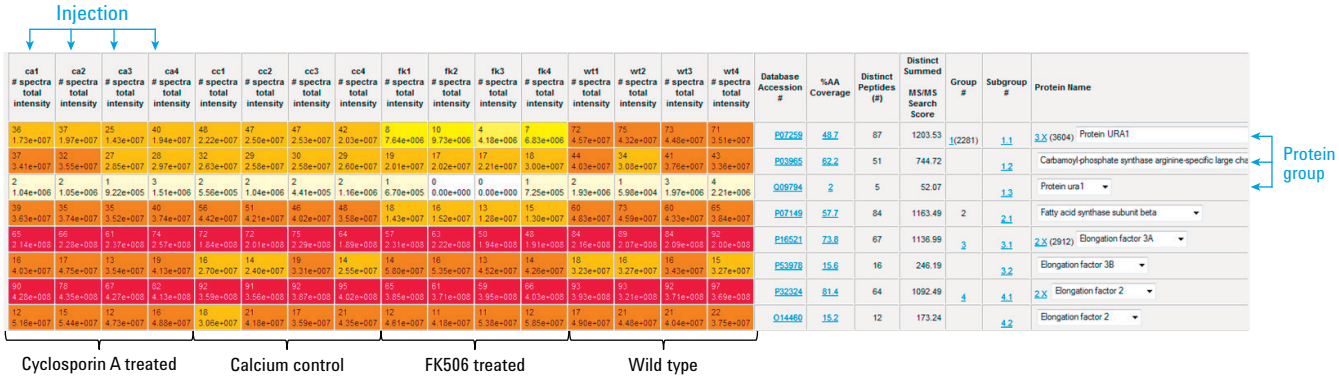

Figure 2. Spectrum Mill protein-protein comparison mode summarizes protein abundances across culture conditions. The results can be exported to MPP for further analysis.

(Figure 2) was used to verify that the abundance of some typical housekeeping proteins (heat shock proteins, elongation factor, enolase) were at roughly equivalent levels across the culture conditions and across several orders of magnitude in abundance. The protein abundances were then exported to MPP and assessed for statistical significance to determine the differential proteins. Principal component analysis (PCA) of the differential proteins showed clear separation of the culture treatments (Figure 3).

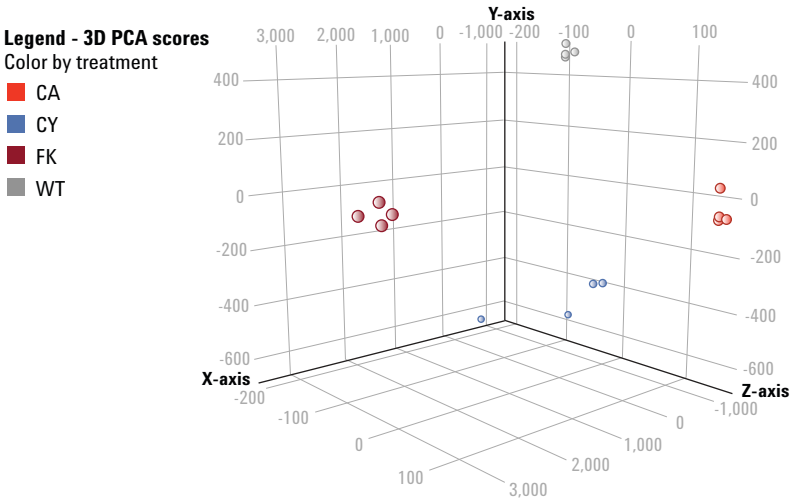

Figure 3. Principal component analysis of the differential proteins (ANOVA) between the culture conditions based on protein abundances from the Q-TOF protein discovery results. The results show clear separation of the four conditions.

The differential proteins were also mapped by Pathway Architect Software to the curated yeast pathways from BioCyc (http://www.biocyc.org). By mapping these proteins and the abundance results, it is possible to view pathways impacted by the treatment. For example, 18 of the 22 proteins in the superpathway of ergosterol biosynthesis (Figure 4) were found to be significantly down regulated in the FK506 treated culture. The individual heat strips adjacent to each differential protein summarize the $\log_2$ transformed abundances for each condition, thus allowing a quick view of the experiment. In addition, a heat map table for all proteins in the pathway is matched with the BioCyc pathway protein list.

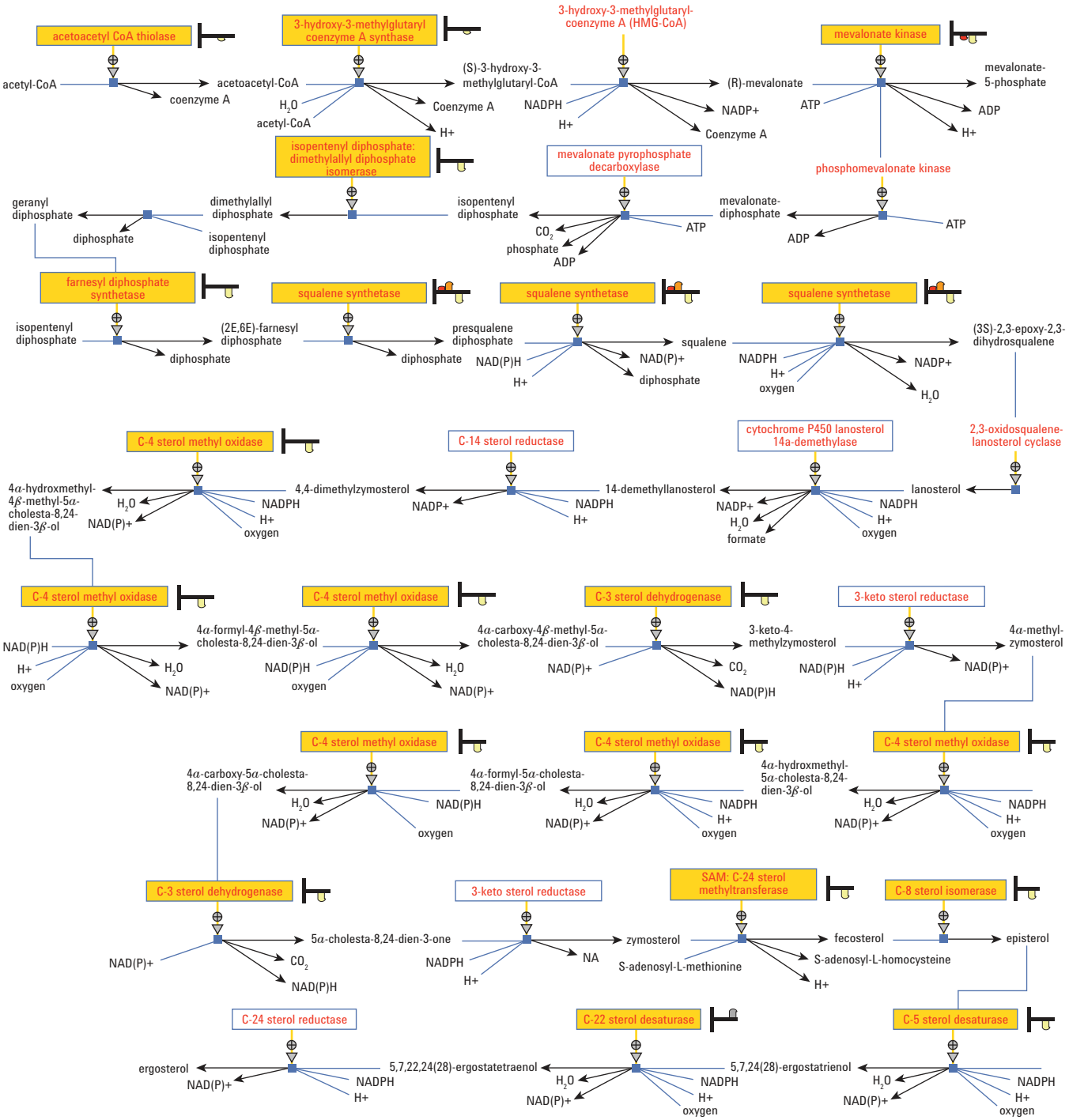

Figure 4. Differential protein results from pathway-targeted QQQ analysis displayed on the superpathway of ergosterol biosynthesis.

Metabolomics pathway-directed targeted workflow

The targeted proteomics workflow (Figure 5) was demonstrated using previously acquired metabolomics results[1] for the same yeast samples as the starting point for generating the list of target proteins. All protein accession numbers were exported for three pathways:

- Salvage pathways of adenine, hypoxanthine and their nucleosides
- Superpathway of purine nucleosides salvage
- Guanosine ribonucleotides *de novo* biosynthesis

This list of 15 proteins was then used to generate proteotypic peptides and possible MRM transitions for the triple quadrupole method. As a first step, the discovery data from the Q-TOF was used to create MRM transitions where possible, and the Spectrum Mill results were exported as a spectral library for Skyline import. The SRM Atlas for yeast[4] provided possible peptides and transitions for those proteins not identified in the Q-TOF experiment. Skyline software was used to further develop and refine the MRM method including *in silico* prediction of peptides and transitions, peak identification, and comparison to the spectral library. The final retention-time scheduled DMRM method had 2-4 peptides per protein (43 total peptides) and 3-4 transitions per peptide (166 total transitions). The final results were reviewed in Skyline (Figure 6), and areas were exported for evaluation in MPP. Peak areas were summed at the protein level prior to import, then subjected to statistical significance testing and visualization.

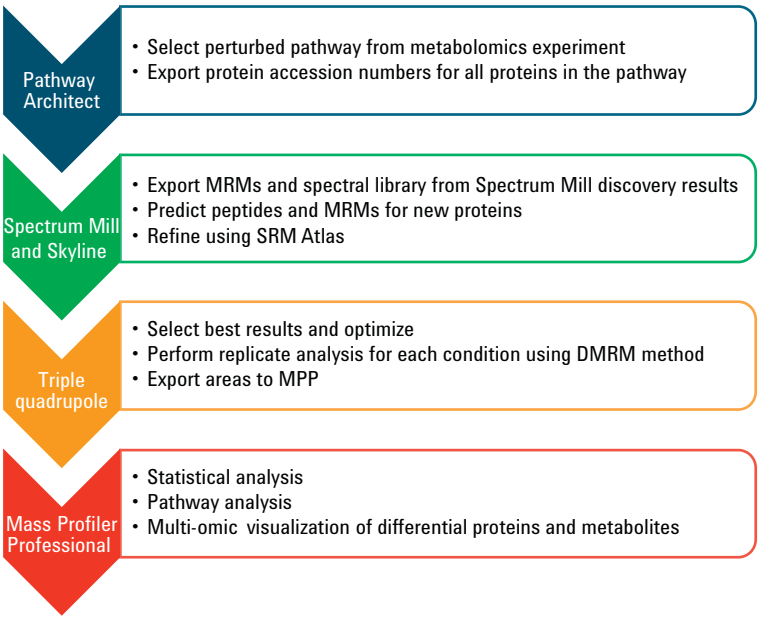

Figure 5. Targeted proteomics workflow using pathway-directed metabolomics information.

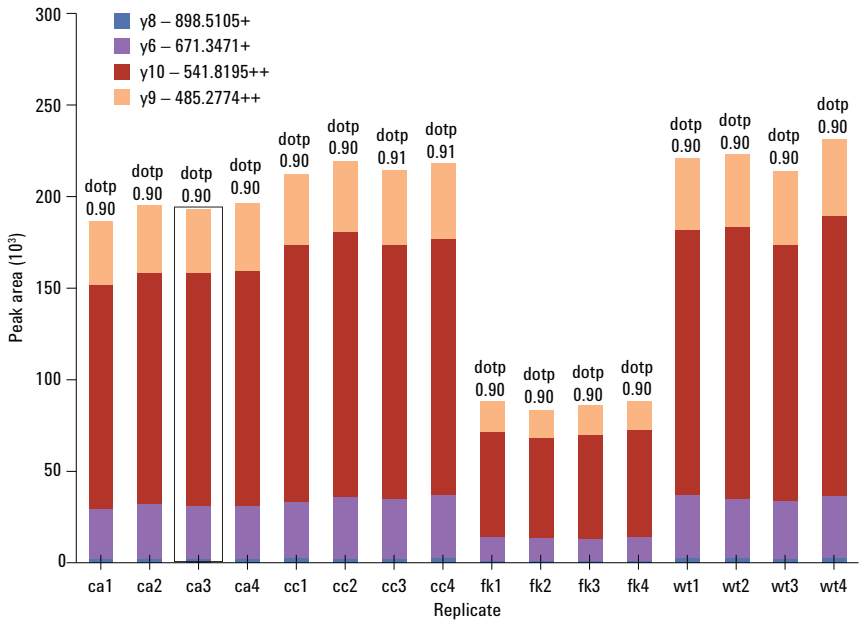

Figure 6. Skyline display of peak areas for all transitions from the peptide TFIAVKPDGVQR (P36010, nucleoside diphosphate kinase). The dot product score to the spectral library is also displayed.

The PCA plot (Figure 7) shows the culture conditions are well separated, and all 15 proteins were found to be significant statistically. MPP 12.0 introduced multi-omic capability whereby two different omic experiments can be mapped and seen on the same pathway. Figure 8 shows the pathway results of a multi-omics experiment created using the metabolomics and targeted proteomics data for the salvage pathways of adenine, hypoxanthine, and their nucleosides. Differentially detected metabolites and proteins are highlighted in color and have an adjacent heat strip for the relative abundances across the culture conditions. In addition, separate heat map tables are produced for the metabolomic and proteomics data.

The advantage of a targeted workflow, as opposed to an extensive fractionation-based discovery workflow, is speed. Using pathway information to focus the targeted experiment gives greater coverage of these biologically interesting proteins. For the three targeted pathways, the discovery experiment only confidently identified three of the 15 proteins in the pathways, whereas the targeted workflow gave 100 % coverage.

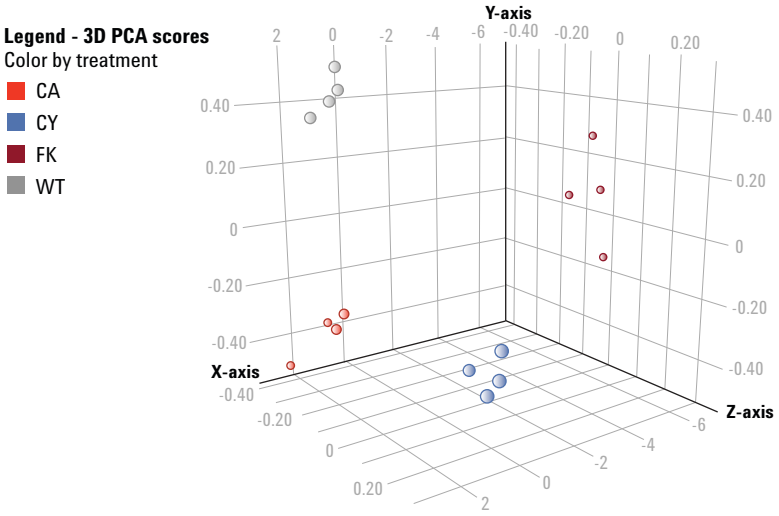

Figure 7. Principal component analysis of the differential proteins (ANOVA) between the culture conditions based on protein abundances from targeted triple quadrupole results. The results show clear separation of the four conditions.

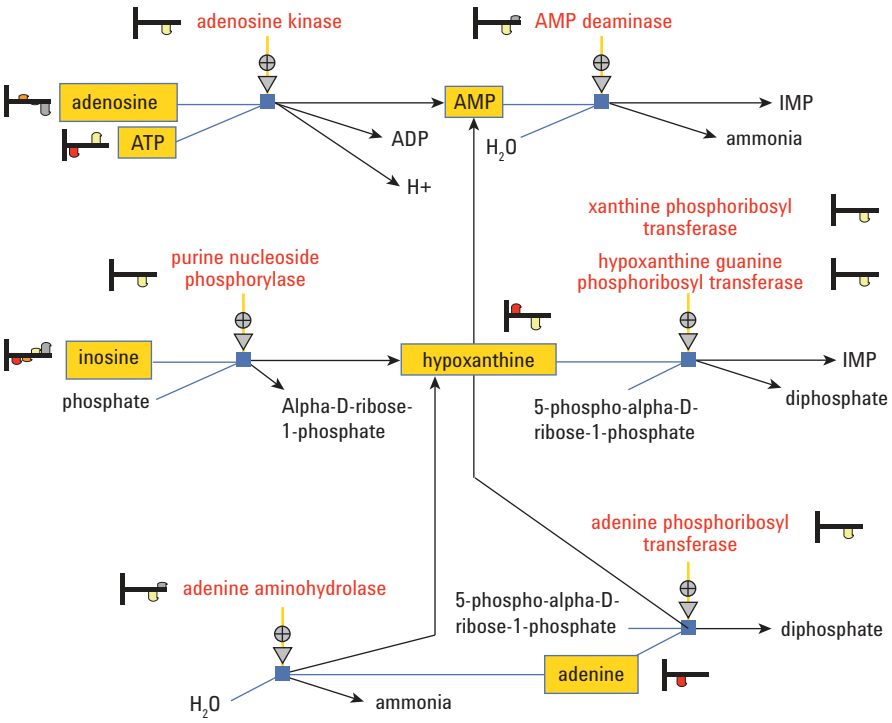

Figure 8. Pathway visualization and mapping of a multi-omics experiment created using metabolomics and targeted proteomics data. The pathway results are for the salvage pathways of adenine, hypoxanthine and their nucleosides.

Conclusions

This application note demonstrates the power of mapping differential protein features to pathways. Mapping the results of a discovery proteomics experiment onto pathways focuses the research to active biological areas. Pathway mapping and visualization allows the protein researcher to identify proteins that were missed in a discovery proteomics experiment.

A metabolomics experiment was the starting point for the next experiment using pathway mapping and visualization to propose a targeted proteomics experiment. The results of the targeted proteomics experiment were then merged with the previously obtained metabolomics experiment to produce a more complete view of the biology. This powerful multi-omics approach was made possible by using MPP and Agilent Pathway Architect Software.

References

1. Jenkins *et al*., Mass Profiler Professional and Personal Compound Database and Library software facilitate compound identification for profiling of the yeast metabolome: Agilent Technologies 5990-9858EN, **2013**.

2. Krokhin, O. V., , Sequence-Specific Retention Calculator. Algorithm forPeptide Retention Prediction in Ion-Pair RP-HPLC: Application to 300- and 100-Å Pore Size C18 Sorbents: Analytical Chemistry, v. 78, p. 7785-7795, **2006**.

3. MacLean, B., *et al*., Skyline: an open source document editor for creating and analyzing targeted proteomics experiments: Bioinformatics, v. 26, p. 966-968, **2010**.

4. Picotti, P., *et al*., A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis: Nature, v. 494, p. 266-270, **2013**.

5. Wisniewski, J. R., *et al*., Universal sample preparation method for proteome analysis: Nat Meth, v. 6, p. 359-362, **2009**.

www.agilent.com/chem/proteomics

Published in the USA, June 20, 2013
5991-2484EN

Agilent Technologies