# Compound Identification, Profiling and Pathway Analysis of the Yeast Metabolome in Mass Profiler Professional

## Application Note


### Authors

Stefan Jenkins, Steven M. Fischer, and Theodore R. Sana
Agilent Technologies, Inc.
Santa Clara, CA USA



### Abstract

A nontargeted, discovery-based approach for analyzing the yeast metabolome in response to calcium and immunosuppressant drug treatment was demonstrated using multiple LC conditions and the Agilent 6530 Accurate Mass Q-TOF mass spectrometer. Data analysis using Mass Profiler Professional (MPP) and the Personal Compound Database & Library (PCDL) yielded putative identifications for several hundred compounds, many of which were observed to be differential between controls and treated groups. Furthermore, canonical biological pathways analysis was performed using Pathway Architect to efficiently project the results of differential abundance results onto publicly available biological pathways, including BioCyc, MetaCyc, and WikiPathways. We inferred the importance of several metabolites using a pathway analysis strategy where we interactively visualized and analyzed metabolites that were over-represented in known biochemical pathways.


Agilent Technologies

## Introduction

Baker's yeast, *Saccharomyces cerevisiae*, is used extensively as a model organism for research into the biochemistry and biological pathways of more complex organisms. It is especially attractive because of the ability to grow it for extended periods under highly controlled conditions[1]. Moreover, its genome has been fully sequenced[2-4].

Metabolomics entails the study of an organism's complete set of small molecule metabolites comprised of a wide variety of compounds. An understanding of the yeast metabolome can shed light on the metabolic pathways of higher organisms, and their response to drugs and environmental changes. Methods are needed to provide as much coverage as possible of the yeast metabolome.

This application note describes a discovery-based, untargeted metabolomics analysis of hundreds of yeast metabolites under robust, controlled extraction conditions followed by identification. This is a prerequisite for studying biological pathways, especially when comparing response to environmental stress[5,6].

A metabolite extraction protocol was optimized to maximize metabolome coverage. The Agilent 1200 SL LC system was coupled to the Agilent 6530 Accurate Mass Q-TOF LC/MS System in multiple chromatographic and detection modes including reversed phase and aqueous normal phase to separate and detect large numbers of metabolites. Data analysis was performed using Agilent MassHunter Qualitative Analysis and Mass Profiler Professional software (MPP), as well as the Personal Compound Database and Library (PCDL) containing content derived from the METLIN database. Yeast cultures were exposed to calcium only or calcium and one of two immunosuppressant drugs (Cyclosporin A and FK 506), with the goal of perturbing the calcineurin and any other calcium/immunosuppressant responsive pathways[7,8]. Initial LC/MS accurate mass matches to METLIN, an accurate mass database of several thousand compounds, resulted in many compound matches with differential abundances. Subsequent confirmation by re-running samples on a LC/Q-TOF for LC/MS/MS analysis provided the highest-confidence identifications.

For pathway analyses, both single (SEA) or multi-omics (MOA) experiments were created in MPP. The results of differential analysis for each omics experiment was summarized in tabular form as a HeatMap, and dynamically mapped onto biological pathways. Since the products of enzymatic turnover are metabolites, and because the sum of all primary and secondary metabolites is expressed as an organism's phenotype, we used pathway-based metabolomics to query various public pathway databases such as WikiPathways and MetaCyc for matching one or more metabolites. To verify the representation of known biochemical pathways in our data, an inferential pathway analysis strategy mapped annotated metabolites onto curated biological pathways for interactive analysis, visualization, and interpretation.

## Experimental

### Standards and Reagents

| | |
|---|---|
| Methanol: | HPLC grade, Burdick & Jackson |
| Filtration filters: | Pall Corporation |
| YPD media: | MP Biomedical |
| FK506, Cyclosporin A: | Sigma-Aldrich |

External standards were added to monitor reproducibility of metabolite extraction in both ion modes. 1-naphthylamine (Sigma-Aldrich) and 9-anthracene carboxylic acid (Sigma-Aldrich) were added to the extraction solvent at a concentration of 5 µg/mL prior to milling/extraction.

### Instruments

This study was performed using an Agilent 1200 SL Series LC system with binary pump and degasser, well plate autosampler with thermostat, and thermostatted column compartment. The LC system was coupled to an Agilent 6530 Accurate Mass Q-TOF LC/MS System with an APCI source operated in positive ion mode, as well as an ESI source operated in positive and negative ion modes. Dynamic mass axis calibration was achieved by continuous infusion of a reference mass solution using an isocratic pump. Tables 1 and 2 give the instrument method parameters.

Table 1. LC Conditions for both reverse phase and aqueous normal phase chromatography.

| LC conditions | | |
|---|---|---|
| Reverse phase | | |
| Column | Guard column: Agilent ZORBAX C-8, 2.1 × 30 mm, 3.5 µm (p/n 873700-936)<br>Analytical column: Agilent ZORBAX C18 SB-Aq column 2.1 × 50 mm, 1.8 µm (p/n 827700-914) | |
| Column temperature | 60 °C | |
| Injection volume | 10 µL | |
| Autosampler temperature | 4 °C | |
| Needle wash | 3 seconds in wash port | |
| Mobile phase | A = 0.2% acetic acid in water<br>B = 0.2% acetic acid in methanol | |
| Flow rate | 0.6 mL/min | |
| Linear gradient | 2% B to 98% B in 13 minutes<br>6 minutes hold at 98% B<br>Stop time: 19 minutes<br>Post time: 5 minutes | |
| Aqueous normal phase | | |
| Column | Cogent Diamond Hydride HPLC Column, 100A 4 µm 2.1 x 150 mm Standard End Fittings (p/n 70000-15P-2)<br>Wash new columns with 10% water/90% isopropyl alcohol overnight. | |
| Column temperature | 60 °C | |
| Injection volume | 2 µL | |
| Autosampler temperature | 4 °C | |
| Needle wash | 3 seconds in wash port | |
| Mobile phase | Positive ion | Negative ion |
| | A = 50% water/50% isopropyl alcohol/0.1% formic acid | A = 50% water/50% isopropyl alcohol/0.025% formic acid and 5 µM EDTA |
| | B: 3% water/97% acetonitrile/0.1% formic acid | B = 10% water/90% acetonitrile with 5 mM ammonium formate with 5 µM EDTA adjusted to pH 7.0 using $NH_3$ |
| Flow rate | 0.6 mL/min | |
| Linear gradient | Positive ion | Negative ion |
| | 97% B to 20% B in 15 minutes<br>Stop time: 15 minutes<br>Post time: 5 minutes | 99% B to 20% B in 15 minutes<br>Stop time: 15 minutes<br>Post time: 5 minutes |

Table 2. Q-TOF MS and MS/MS conditions.

| MS conditions | |
|---|---|
| Ion mode | ESI Positive and negative, APCI positive |
| Drying gas temperature | 325 °C |
| Vaporizer temperature | 350 °C |
| Drying gas flow | 10 L/min ESI, 5L/min APCI |
| Nebulizer pressure | 45 psi |
| Capillary voltage | 4,000 V positive ion mode<br>4,000 V negative ion mode<br>3,500 V APCI |
| Spectra acquisition rate | 1.4 spectra/second |
| **MS/MS conditions** | |
| Quad resolution | High resolution |
| Ion mode | Both positive and negative |
| Drying gas temperature | 325 °C |
| Drying gas flow | 9 L/min |
| Nebulizer pressure | 45 psig |
| Capillary voltage | 4,000 V (positive mode) / 3500 V (negative mode) |
| Fragmentor | 140 V |
| Skimmer | 65 V |
| OCT1RFVpp | 750 V |
| Isolation width | ~ 1.3 *m/z* |
| Reference celivery | Agilent 1100 isocratic pump with 100:1 splitter (p/n G1607-60000) |
| Reference pump flow | 1 mL/min for 10 µL/min to nebulizer |
| Reference ions | Positive mode: 121.050873 and 922.009798<br>Negative mode: 119.036320 and 966.000725 |
| Instrument mass range | 1,700 Da |
| Acquisition rate | 3.35 spectra/s |
| TOF spectra mass range | 25 to 1,000 *m/z* |
| Collision energy (eV) | 10 and 20 |
| Data storage | Centroid |
| Threshold | 100 (MS) and 5 (MS/MS) |
| Instrument mode | Extended Dynamic Range |

## Sample preparation

*S.cerevisiae* strain BJ5459, a gift from Dr. Martha Cyert of Stanford University, was cultured in parallel and at an $OD_{600}$ of 0.8 exposed to vehicle control (4 mL of 90:10 ethanol:Tween 20) for wild type (WT) and calcium control (CA) cultures, or 4 µg/mL FK506 immunosuppressant (FK), or 4 µg/mL cyclosporinA (CY) for 1 hour, suspended in 90:10 ethanol:Tween 20. An equal fraction of media containing $CaCl_2$ was added to FK and CY cultures as well as a calcium-only exposed culture (calcium control-CA), 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) to remove any residual media. Quenching was done with 1 mL of methanol added to the final pellet at –40 °C. After quenching, the sample was lyophilized.

Wet milling was performed with 5 mg of dry sample in 2-mL Eppendorf tubes in an extraction solvent of 1.1 mL of 5:3:3 chloroform:methanol:water. A Retsch MM301 ball mill mixer was used with a single 5-mm ball bearing as a tube insert to facilitate mechanical rupture of the yeast. External standards were added prior to milling. Nine replicate samples for each culture condition were processed for 3 × 1 minute cycles at 30 Hz. The milling process resulted in polar and nonpolar solvent phases. The polar phase supernatants were filtered first through 0.2-µm and then 10-kDa ultrafiltration membranes to ensure removal of any residual protein/cellular debris.

### Data Analysis

Compounds were extracted from the raw data files using an unbiased, Molecular Feature Extraction (MFE) algorithm in Agilent MassHunter Qualitative Analysis B.04.00 Software. The processed data files were subsequently analyzed using Agilent Mass Profiler Professional (MPP), a multivariate statistical analysis and data visualization software. Provisional compound identification was performed by matching accurate mass results to content from:

- METLIN Personal Compound Database and Library (PCDL)
- Additional, orthogonal RT matching using Agilent Accurate Mass Retention Time (AMRT) database

In addition, MS/MS sample spectra were queried against a library of LC/MS/MS spectra acquired from over 2,700 individual standards.

### Pathway Analysis

Following statistical analysis of yeast proteomics data by ANOVA, a yeast metabolomics pathway experiment was created and analyzed using the Single Experiment Analysis (SEA) wizard in MPP. An example of metabolite profile results that can be projected onto pathways is represented by the salvage pathways of adenine, hypoxanthine and their nucleosides from YeastCyc (Figure 6). The colored HeatStrips denote the average normalized abundances of each metabolite for the different treatments: calcium control, Cyclosporin A, FK506 and Wild Type respectively.

## Results and Discussion

### Feature Extraction

The MFE algorithm in MassHunter finds all related ions, including isotopes and any adducts, such as Na+ or K+ and dimers. It then combines all ion signals into one value (feature), in a fully automated mode. Table 3 shows the results for the four cultures: wild type (WT), calcium-control (CA), FK 506 treated (FK), and Cyclosporin A treated (CY). Each culture was analyzed using both reverse phase (RP) for semipolar and nonpolar compounds and aqueous normal phase (ANP) chromatography for polar compounds. QTOF MS analysis was done with an electrospray ionization (ESI) source operated in positive and negative modes, as well as an atmospheric pressure chemical ionization (APCI) source operated in positive ion mode. For the most part, the results for all four culture types were similar, with ANP-ESI + generating the most features. By incorporating two chromatographic separation techniques, two MS sources (ESI and APCI), as well as both positive and negative ionization polarities, we achieved broader detection coverage of the yeast metabolome.

### Data Analysis in MPP

Data filtering for the biological replicates resulted in a reproducible list of frequently detected compounds, which were then assessed for statistical significance to determine differential abundance of metabolites according to the culture conditions. Chemical entities (features) that were present in at least six of nine data files, and in at least one of the four culture conditions were retained (Table 4). The filtered list of compounds was then queried against the METLIN database of > 25,000 compounds in PCDL (Table 4) for identification. The filtered features were then evaluated for statistical significance between culture conditions at a cutoff of $p < 0.05$ (Table 4). Features that passed this cutoff were again queried in terms of accurate mass against the METLIN PCDL (Table 4).

Table 3. MFE results by treatment and analysis type. Average number of features found for samples in each culture condition (n=9) and analysis type.

| | WT | CA | FK506 | CY |
|---|---|---|---|---|
| RP-ESI+ | 418 | 417 | 411 | 416 |
| RP-ESI– | 439 | 488 | 442 | 462 |
| RP-APCI+ | 260 | 277 | 381 | 569 |
| ANP-ESI+ | 992 | 966 | 961 | 924 |
| ANP-ESI– | 267 | 211 | 191 | 183 |

RP: reverse phase chromatography
ANP: aqueous normal phase chromatography
ESI+: electrospray ionization, positive polarity
ESI–: electrospray ionization, negative polarity
APCI+: atmospheric pressure chemical ionization, positive polarity

Table 4. MPP results by analysis type. Average number of entities found for samples in each culture condition (n=9) and analysis type.

| | Compounds passing frequency filter | METLIN PCDL identified compounds | Compounds at $p < 0.05$ | METLIN PCDL identified compounds at $p < 0.05$ |
|---|---|---|---|---|
| RP-ESI+ | 300 | 112 | 158 | 79 |
| RP-ESI– | 523 | 50 | 418 | 32 |
| RP-APCI+ | 364 | 48 | 333 | 37 |
| ANP-ESI+ | 398 | 81 | 129 | 25 |
| ANP-ESI– | 276 | 88 | 213 | 63 |

RP: reverse phase chromatography
ANP: aqueous normal phase chromatography
ESI+: electrospray ionization, positive polarity
ESI–: electrospray ionization, negative polarity
APCI+: atmospheric pressure chemical ionization, positive polarity

## PCA

Principal Component Analysis (PCA) resulted in spatial separation for all four culture treatments (Figure 1).

## Profile Plot

Analysis of the differential abundances for several compounds revealed the effect of the various treatments on several metabolic pathways. One example was the metabolites 5′-methylthioadenosine and S-adenosylhomocysteine which are both substrates of the enzyme 5′-methylthioadenosine nucleosidase (EC 3.2.2.9) in the cysteine and methionine metabolism pathway. Profile plots of the abundances for these two metabolites were increased in the calcium treated data set (CA), indicating possible inhibition of 5′-methylthioadenosine nucleosidase in a calcium dependent manner (Figure 2). A similar calcium dependent effect upon 5′-methylthioadenosine nucleosidase has been previously reported for an orthologous pathway in *Arabidopsis thaliana*[9]. Good reproducibility for the biological replicates was reflected by the low relative standard deviations (RSDs).

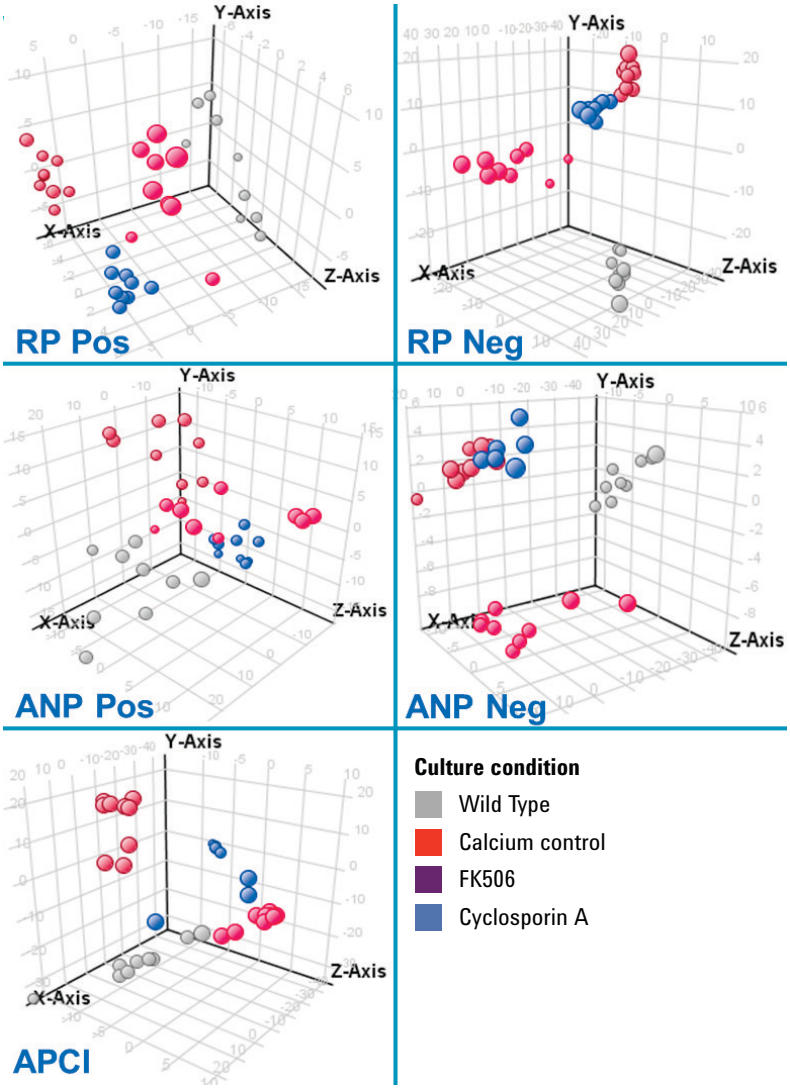


Figure 1. PCA plots generated using MPP for each analysis type. Spatial separation for each culture condition (n=9) was observed, indicating the unique effect of each culture treatment.

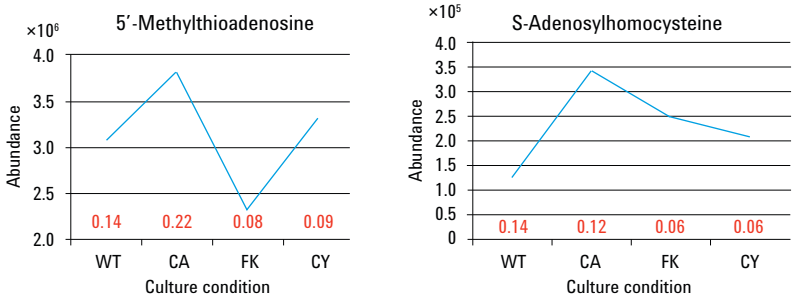


Figure 2. Differential abundances calculated in MPP at $p<0.05$ for compounds identified by accurate mass in PCDL as 5′-Methylthioadenosine and S-Adenosylhomocysteine. Evaluation of extracted ion chromatograms (EICs) confirmed differential abundance. The charts represent the average abundance according to culture condition (n=9), with the RSD in red.

## MS/MS Spectral Matching

MS/MS spectral matching currently provides the highest confidence compound identifications in discovery-based MS analyses. Moreover, MS/MS spectral matching in PCDL provides a level of confidence that cannot be obtained using publicly available databases that use accurate mass only. The addition of MS/MS library searching capability against 2,700 standards with MS/MS spectra in PCDL provided a chromatography-independent means for compound identification.

Several compounds in this application note that showed differential abundance with statistical significance ($p < 0.05$), were identified using MS/MS spectra matching. These are represented by mirror image plots of the acquired sample spectrum (top) with library standard spectra (bottom) (Figure 3 A–F). Both forward and reverse searches were performed for each compound. The collision energy scores are shown in red. As the collision energy was increased from of 20 eV to 40 eV, decreased abundance of parent ion was observed, in conjunction with higher abundance of lower mass fragments.

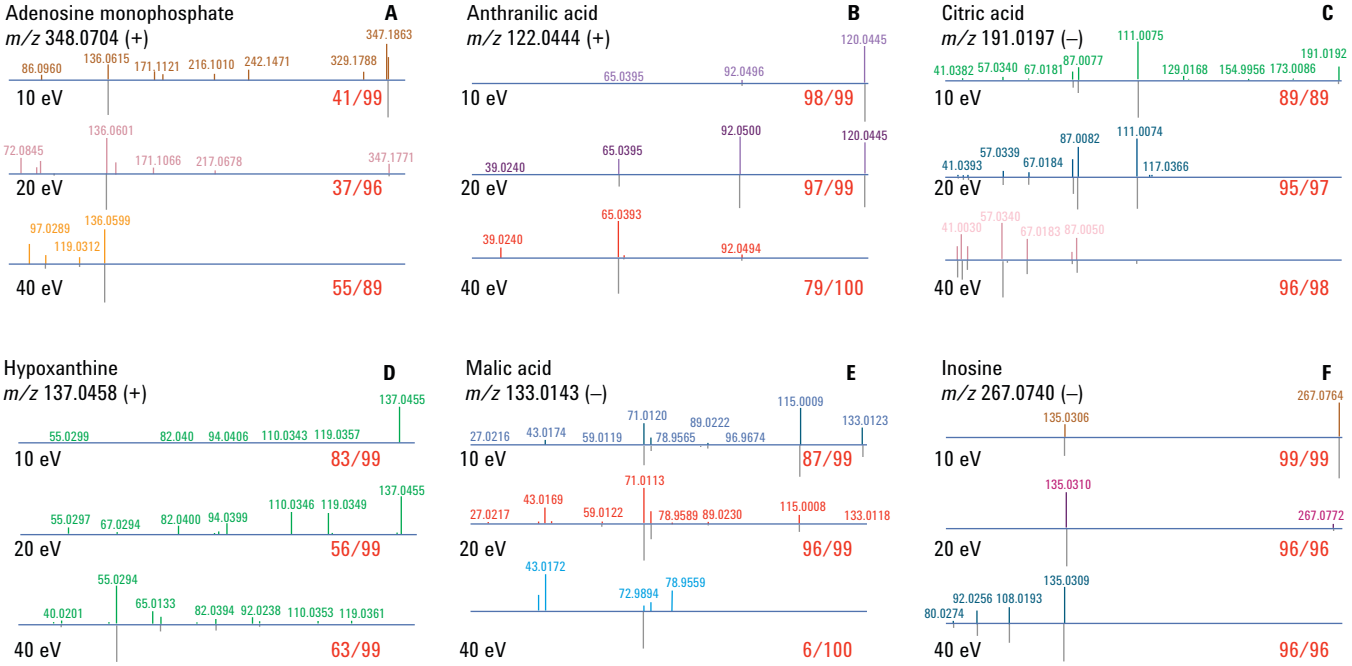


Figure 3 A–F. MS/MS spectra were collected for each compound in the samples at 10, 20, and 40 eV. The spectra were matched to the METLIN PCDL library and scored using a Forward and Reverse library match scoring result (in red) for each collision energy with a maximum possible score of 100. The acquired spectrum is displayed above the library spectrum. The Forward score reflects isotope matches for all the peaks from the sample spectrum that are reflected in the library of standards spectrum, with a penalty for each non-matching peak. The Reverse score is the opposite; only the peaks from the library standards were used to match those spectra acquired from the sample. The Reverse score is typically higher.

## Pathway Analysis in MPP

Pathway analysis was performed in Mass Profiler Professional 12.5 (MPP 12.5), providing a means for the visual analysis of metabolite abundances. By mapping the compound IDs and overlaying the abundance results onto curated pathways, such as those sourced from WikiPathways and YeastCyc (http://pathway.yeastgenome.org/), it was possible to view which pathways were enriched with metabolites. For example, the results for several metabolites associated with methionine degradation were mapped onto the corresponding pathway in WikiPathways, facilitating a quick view of whether the pathway was significantly perturbed in response to various drug treatments (Figure 4). Individual HeatStrips adjacent to each compound are summarized as the log2 transformed abundance values for each condition. Furthermore, a HeatMap table (Figure 5) of all compounds from the compound list is matched with those from the Pathway compound list from WikiPathways. The HeatMap is dynamically linked to the HeatStrips in the pathway figures. This provides an additional means for reviewing specific compounds, and for projecting the abundances of their potential isomers across the different treatment conditions. A separate Yeast Biochemical Pathway Database (YeastCyc), from a Level 3 BioCyc pathway collection for *Saccharomyces Cerrevisiae*, shows the salvage pathways of adenine, hypoxanthine, and their nucleosides. The Simple graphical layout was selected for rendering associations between different entities in the pathways. The metabolites and associated HeatStrips that were matched in the pathways were boxed and highlighted with a yellow background.

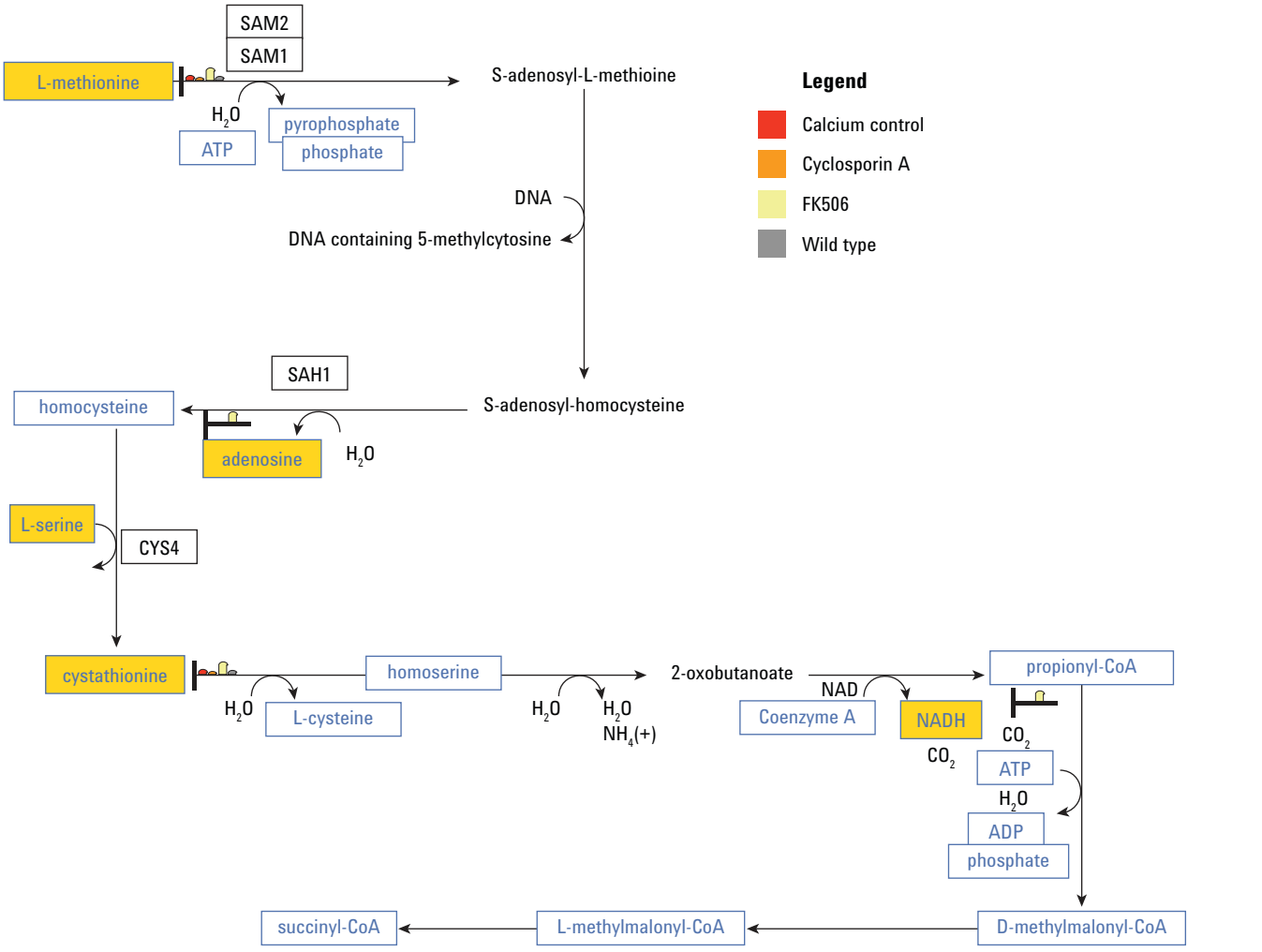


Figure 4. A WikiPathways map of the Yeast methionine degradation pathway, showing HeatStrips corresponding to several metabolites and their corresponding log2 transformed average abundances for the different conditions: Calcium control, Cyclosporin A, FK506, and Wild Type.

| Review | Entity Name | [Calcium] | [Control] | [Cyclosporin A] | [FK506] | DB | DB ID |
|---|---|---|---|---|---|---|---|
| ☑ | adenosine | 0 | 0 | 6.018 | 0 | CAS Number | 58-61-7 |
| ☑ | cystathionine | 12.033 | 0 | 9.678 | 9.655 | CAS Number | 56-88-2 |
| ☐ | L-methionine | 9.063 | 5.079 | 18.007 | 6.662 | CAS Number | 63-68-3 |
| ☑ | L-serine | 2.287 | 0 | 0 | 2.268 | CAS Number | 56-45-1 |
| ☐ | NADH | 0 | 0 | 2.304 | 0 | CAS Number | 53-84-9 |
| | ATP | | | | | | |
| | pyrophosphate | | | | | | |
| | phosphate | | | | | | |
| | homocysteine | | | | | | |
| | L-cysteine | | | | | | |
| | homoserine | | | | | | |

Find: Find Next Find Previous Match Case

Figure 5. A representative HeatMap table for detected compounds that are present in the methionine degradation pathway from WikiPathways. Normalized abundance values are shown with color coding to provide an additional means for interpreting trends across the different conditions: Calcium control (Calcium), Wild Type (Control), Cyclosporin A and FK506.

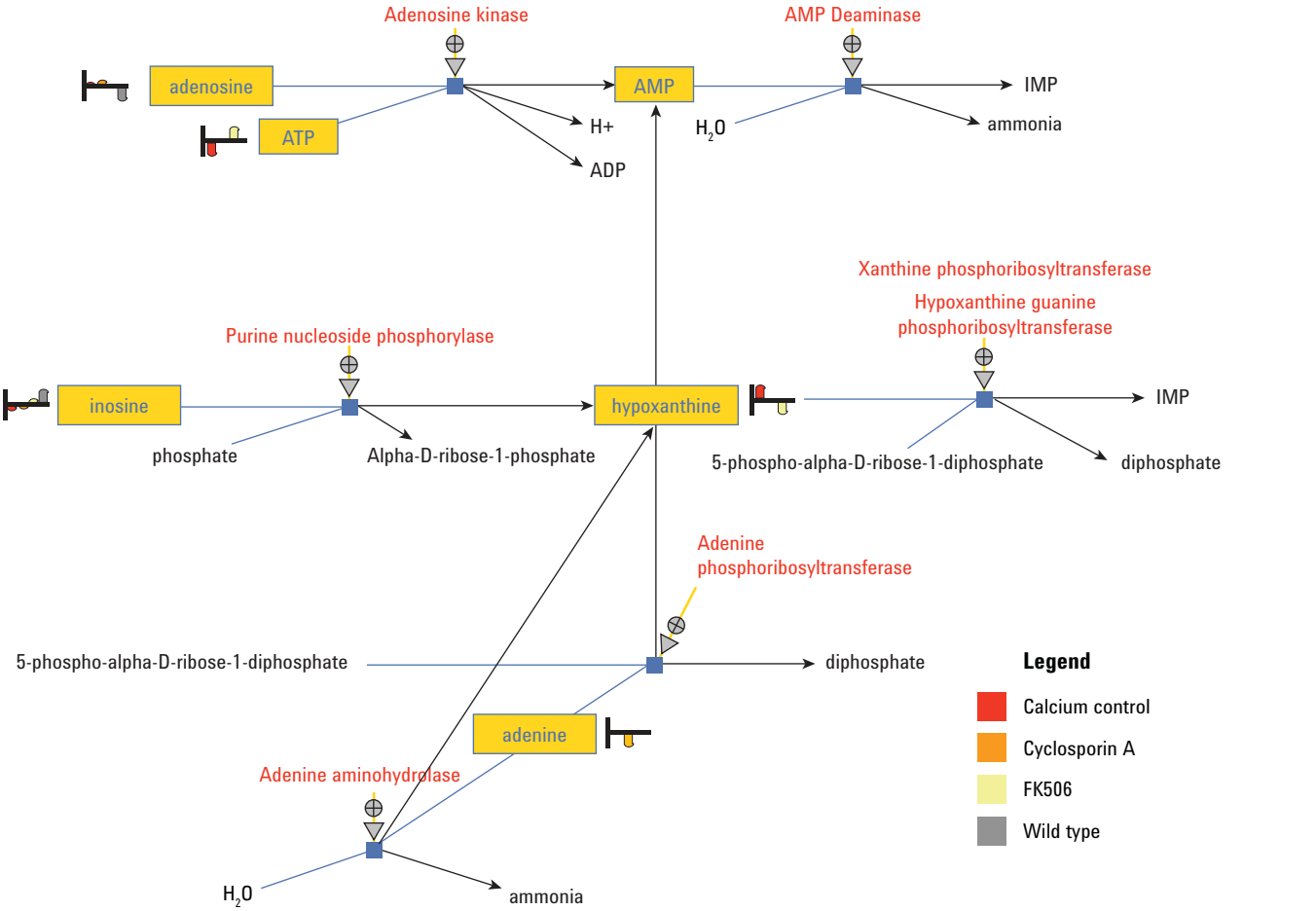


Figure 6. The salvage pathways of adenine, hypoxanthine, and their nucleosides from YeastCyc rendered in the Pathway Architect reveals associations between the different metabolite and protein entities.

## Conclusions

The utility of a nontargeted, discovery based accurate mass approach for analyzing the yeast metabolome in response to calcium and immunosuppressant drug treatment was demonstrated. A combination of new software tools, databases, and libraries yielded several hundred compounds, many of which were observed to be differential between controls and treated groups.

Agilent MPP was used for data alignment, filtering, and statistical analysis to determine differential relative abundances of hundreds of compounds. The Agilent METLIN PCDL was used for provisional identification of compounds using accurate mass analysis, with subsequent confirmation of metabolite identity by matching acquired sample MS/MS spectra to spectra of standard compounds in the library. In order to visually represent metabolite abundances onto pathways, a Pathway Architect module in MPP was used to map compounds onto pathway content in curated databases, according to treatment condition. This discovery-based, pathway-informed workflow is an efficient, cost-effective tool for finding and identifying differential compounds that are over-represented in response to calcium and immunosuppressant drug treatment. Furthermore, the results can be used to design the next series of experiments, such as using metabolomics pathway results to design a proteomics experiment. This pathways driven approach to multiomics can provide further insights into the possible biological relevance of one's experimental results in an intuitive, visual and interactive way.

## References


1. Herrgard *et al.*, "A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology", *Nat Biotechnol*. 26(10):1155-1160, **2008**.
2. Botstein *et al.,* "Yeast as a model organism", *Science* 277(5330):1259-1260, **1997**.
3. Saccharomyces Genome Database- http://www.yeastgenome.org
4. European Saccharomyces Cerevisiae Archive for Functional analysis- http://web.uni-frankfurt.de/fb15/mikro/euroscarf
5. Villas-Boas *et al*., "Global metabolite analysis of yeast: evaluation of sample preparation methods", *Yeast* 22(14): 1155-1169, **2005**.
6. Martins *et al*., "Comparison of sampling techniques for parallel analysis of transcript and metabolite levels in Saccharomyces cerevisiae", *Yeast* 24(3), 181-188, **2007**.
7. Aramburu *et al.,* "Calcineurin: From Structure to Function", *Curr Top Cell Regul*, 36:237-295, **2000**.
8. Cyert, "Calcineurin signaling in Saccharomyces cerevisiae: how yeast go crazy in response to stress", *Biochem Biophys Res Commun*, 311(4):1143-1150, **2003**.
9. Seung-Ick Oh *et al*., "The Arabidopsis Calcium Sensor Calcineurin B-Like 3 Inhibits the 5'-Methylthioadenosine Nucleosidase in a Calcium-Dependent Manner", *Plant Physiol*, 148(4):1883-1896, **2008**.


www.agilent.com/chem/mpp



Published in the USA, June 18, 2013
5991-2470EN


Agilent Technologies