Agilent Metabolomics Workflow Solutions Using LC/MS, GC/MS, and NMR

Technical Overview

Introduction

Agilent Technologies is an established leader in providing metabolomics solutions. Metabolomics refers to the identification and quantitation of metabolites, which are the small organic molecules that are used in biological processes within cells. These molecules are important indicators of biological function, and their study can provide unique insights into cell physiology, metabolism, and toxicity.

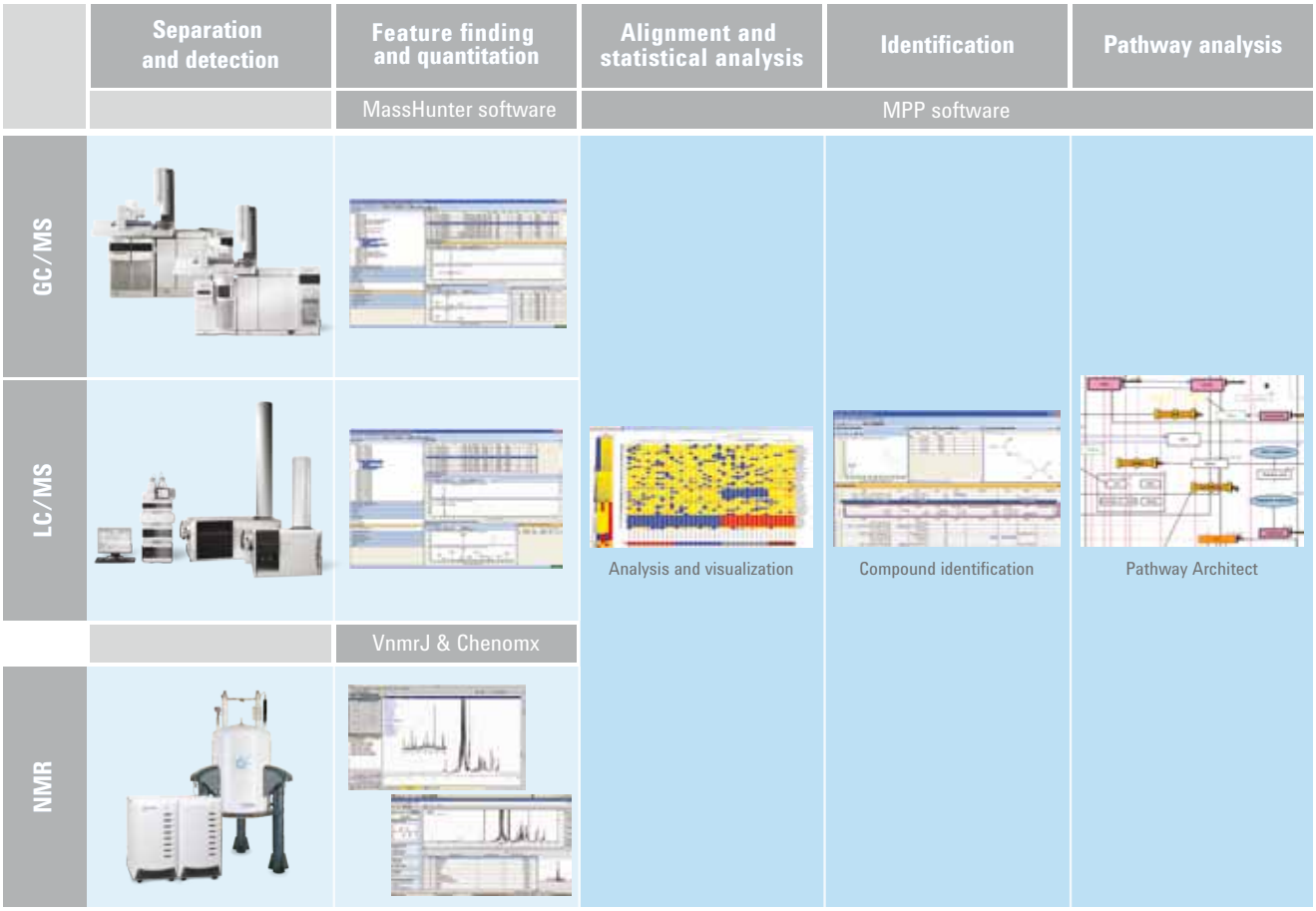
Because different kinds of metabolomics experiments require different technological approaches, Agilent provides several different workflow solutions. These workflows use liquid chromatography/mass spectrometry (LC/MS), gas chromatography/mass spectrometry (GC/MS), and nuclear magnetic resonance (NMR) instrumentation for data acquisition, metabolomics compound databases for compound identification, and software for data analysis and biological interpretation.

Agilent Technologies

This document provides an overview of the components of these metabolomics workflows (Figure 1). It also discusses considerations for choosing between approaches. For example, technical aspects of an experiment, such as sensitivity requirements, sample complexity, sample preparation, and the chemical properties of metabolites of interest can all influence technology decisions. These factors and others will be addressed in more detail in the sections that follow.

Comparing targeted and discovery metabolomics

Metabolomics experiments can encompass both discovery and targeted approaches. Discovery metabolomics refers to experiments that involve examining an untargeted suite of metabolites, finding the ones with statistically significant variations in abundance within a set of properly replicated experimental samples (as compared to controls), and determining their chemical identities. In contrast, targeted metabolomics refers to experiments that aim to accurately measure the abundance of previously identified metabolites, such as validation studies which typically include a larger number of samples. Targeted experiments can be highly quantitative and usually require the use of analytical standards.



Metabolomics via mass spectrometry

Mass spectrometry (MS) approaches are commonly used in metabolomics research due to their wide dynamic range, sensitivity, reproducibility, and ability to analyze complex biological fluids. To address sample complexity, separation (gas chromatography, liquid chromatography or capillary electrophoresis) is usually performed in series with MS analysis to detect as many metabolites as possible. Both the mass spectral and chemical information in the retention time (or migration time) of the separation are used to track and identify the metabolites undergoing analysis. This is necessary because MS alone often cannot distinguish isobaric metabolites, whereas chromatography can separate and identify isomers. Gas-phase approaches are more suited to more non-polar compounds than liquid-phase separations, but many classes of compounds can be resolved equally well by either approach.

Comprehensive resources for MS metabolomics

To support both discovery and targeted MS metabolomics, Agilent offers a range of GC/MS and LC/MS instrumentation that can be adapted to fit the specific application, including time-of-flight (TOF), quadrupole time-of-flight (Q-TOF), and triple quadrupole (QQQ) mass spectrometers (Figure 2). Agilent systems are designed to detect low-abundance targets, even in the presence of high-abundance metabolites, while maximizing reproducibility. A key design feature enabling this on Agilent Q-TOF instruments (when used in combination with LC or GC) is the use of dual signal gain processing which creates an instrument with an effective linear dynamic range covering five orders of magnitude.[1]

In addition to a comprehensive line of high-performance GC/MS and LC/MS instruments, Agilent has also developed the essential software tools needed to analyze, identify, and interpret the vast amount of data produced in metabolomics experiments. Each instrument uses the same data analysis software, creating a common user experience for experiments that require multiple MS configurations. Agilent has also developed metabolomics-specific, compound identification databases and libraries for evaluating both GC/MS and LC/MS data.[2] The processed data is analyzed using the Mass Profiler Professional (MPP) module of Agilent GeneSpring software, a statistical analysis and data visualization tool that allows researchers to quickly find and extract meaningful results from their data. The use of MPP is described in further detail later in this document.

Metabolite databases

Agilent collaborates with leading metabolomics researchers who are developing the next generation of metabolomics techniques and tools, and incorporates their advances into robust solutions that enhance and simplify metabolomics research. A key example of this is Agilent's effort with regard to creating curated databases and libraries for compound identification by MS in order to facilitate accurate compound identification. In particular, we have worked closely with Dr. Oliver Fiehn to develop searchable GC/MS electrospray ionization (EI) spectra and retention time indexes for approximately 700 common metabolites[2] and for LC/MS studies, we have worked with Dr. Gary Siuzdak to develop the METLIN personal compound database and library (PCDL), which contains over 25,000 compounds, including 8,000 lipids, as well as MS/MS spectra for 2,200 compounds, and retention times for about 700 standards.[3] For more information on Agilent database resources, please see "Database Resources in Metabolomics: an Overview".[4]

Figure 2. Agilent 6500 Series Accurate-Mass Q-TOF LC/MS and 7000 Series Triple Quadrupole GC/MS systems.

Metabolomics via nuclear magnetic resonance

NMR provides another powerful tool for metabolomics research. This is because NMR is inherently quantitative (a unit response factor can be measured for any molecule, even without standards), highly reproducible, and requires minimal sample preparation. Moreover, NMR provides completely orthogonal compound detection and identification mechanisms to LC/MS or GC/MS because the spectroscopic properties interrogated by NMR are unrelated. An Agilent 750 Mz NMR system is shown in Figure 3.

NMR is complementary to MS for metabolomics research in two important ways: first, NMR is unique in that it does not rely on separation of the sample matrix prior to data collection; second, all of the small molecule metabolites in a sample are measured simultaneously. Consequently, the results are not biased by the method itself. Moreover, NMR is non-destructive, so samples remain available for further analysis. NMR metabolomics studies can be applied to *in situ*, *in vivo*, *in vitro*, and *in viro* samples, such as those as indicated below:

- MRI
- Biopsies
- Excised and perfused whole organs
- Perfused cell cultures
- Intact plants, seeds, and tissues
- Unadulterated, unmodified biofluids and biota
- Solvent and/or aqueous extracts

Figure 3. Agilent DD2 MR System with a 750 MHz magnet.

In terms of technology, metabolomics research is very demanding of the NMR spectrometer, but Agilent NMR systems are designed to provide the stability, flexibility, and performance required for these studies. Quantitative measurement by NMR requires that the spectrometer response is linear over a large concentration range. The Agilent DD2 receiver system meets this requirement because it directly samples the NMR signal, rather than using outdated methods such as quadrature detection. This yields a flat baseline without phase distortions and a highly linear response, both of which are needed for accurate quantitative measurements. For example, when applied to the quantification of tetracycline, the Agilent DD2 MR system provides direct measurement of sample concentrations down to very low µM levels with high precision and accuracy (Figure 4).

High-quality solvent suppression is also critical for metabolomics NMR research and Agilent NMR systems support the full complement of solvent suppression strategies, making it possible to tailor the solvent suppression method to fit the specific situation and then automatically apply the same method to all samples in a given study.

Agilent's NMR instrumentation is supported by the VnmrJ software package which allows complete user control and customization for sample-specific methods. Acquisition parameters, shimming methods, data archiving, and automatic processing tools can be specified independently for a given sample type, and then recalled and applied using a single mouse click. Data collection on a large number of samples is simplified with spreadsheet-driven submission tools, and automated shimming routines can be configured and applied as needed. Following data collection, NMR metabolomic profiles are computed using the Chenomx software suite.

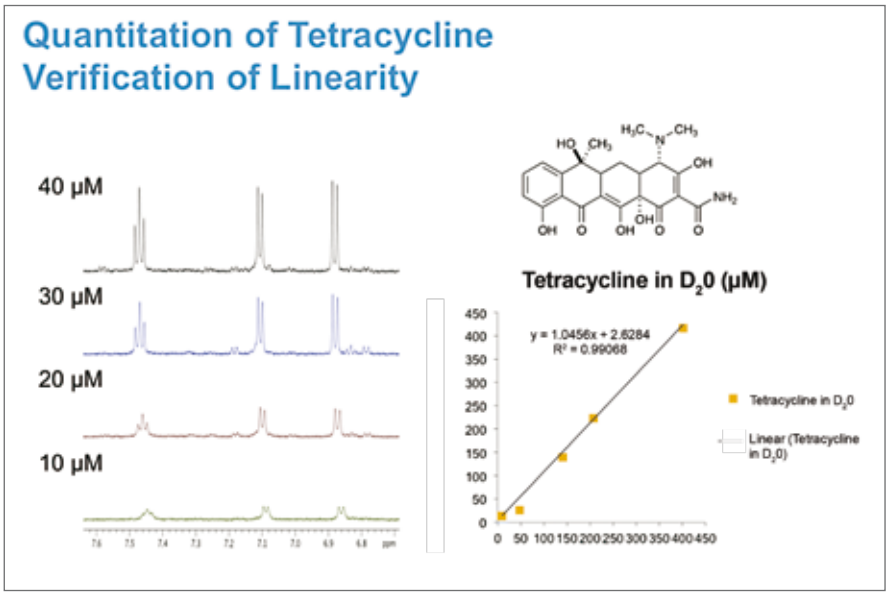


Figure 4. Quantitation of tetracycline and verification of linearity.

Statistical analysis and compound identification

In order to help understand the biological meaning of MS or NMR-derived metabolomics data, Agilent provides the GeneSpring/ Mass Profiler Professional (MPP) software suite. MPP includes several useful visualization and analysis tools for metabolomics experiments, as well as capabilities for the co-analysis of metabolomics, proteomics, transcriptomics, and genomics data.

Regarding metabolomics-specific analyses and visualization tools, MPP enables data normalization, alignment, and comparison of data collected using multiple instruments (Figure 5). In addition, experimental metadata can be added to the analysis to highlight relationships in complex experimental designs.

With respect to biomarker discovery or biological interpretation, MPP facilitates identification of important differences in the type and relative amount of metabolites present in different samples using statistical techniques such as analysis of variance (ANOVA), principal component analysis (PCA), clustering, volcano plots, hierarchical clustering, enrichment, and class prediction. Moreover, adding customized analyses is straightforward using the MPP interface to the R statistical computing environment.

For LC/MS and GC/MS-based metabolomics, MPP couples the statistical tools above with those necessary for determining chemical identities. Depending upon the analytical technique used to acquire the data, different levels of confidence can be achieved in compound identification. The most common approach is through metabolite-specific database matching, using accurate-mass information, in order to identify candidate compounds. The addition of isotope pattern matching can increase the specificity of the assignment to confirm the empirical formula. Moreover, if retention time information is also included, confident compound identification can be achieved. Alternatively, searching MS/MS or EI libraries can produce similar identification results. Combining retention time information with MS/MS or EI library searching gives the highest level of confidence. Agilent provides databases and libraries for compound identification by either GC/MS or LC/MS, as described earlier in this document.[3]

For NMR-based metabolomics, Agilent's workflow leverages the Chenomx NMR software suite to produce metabolic profiles from NMR spectra. Once created, these metabolomics profiles can be imported into MPP for statistical analyses and further biological interpretation, including integrated analysis.

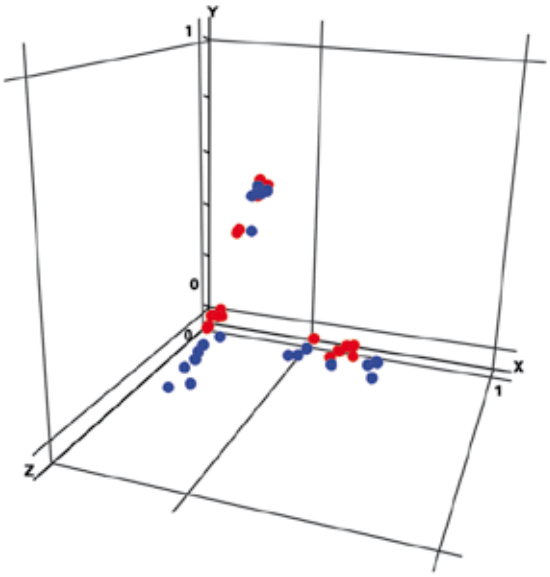


Figure 5. MPP software allows the comparison of multiple samples and/or multiple MS analysis platforms in a single project.

Integrated biological analysis

Integrated or multiomics analysis is critically important for metabolomics research in two distinctly different ways. First, some of the most important biological insights can only be discovered by co-analyzing metabolomics data in the context of external database and in comparison to other kinds of biological measurements. To enable this, Agilent currently offers an approach based on a simple premise: The most intuitive way to link metabolite measurements to protein or transcript levels is by using metabolic pathways. When used in conjunction with the Agilent Pathway Architect module, MPP enables co-visualization and co-analysis of metabolites and gene products in the context of such pathways (Figure 6). Currently, WikiPathways is the primary pathway source, but exports to MetaCore, IPA, BioCyc, and KEGG pathway resources are also supported.

Second, integrated biology is critical for identifying the most promising follow-up experiments. For example, consider a situation in which the analysis of a proteomics (or transcriptomics) experiment reveals that a particular pathway is active. This knowledge can be used to design a targeted metabolomics assay by using information about the chemical moieties known to be present in the pathway. Analogously, metabolomics experiments can suggest follow-up proteomics, genomics, or transcriptomics experiments. To facilitate this kind of feedback loop, the Pathway Architect module helps researchers find active pathways and can then export identifiers from selected pathways directly to open source tools like Skyline to help design and execute targeted proteomics analyses, for instance.

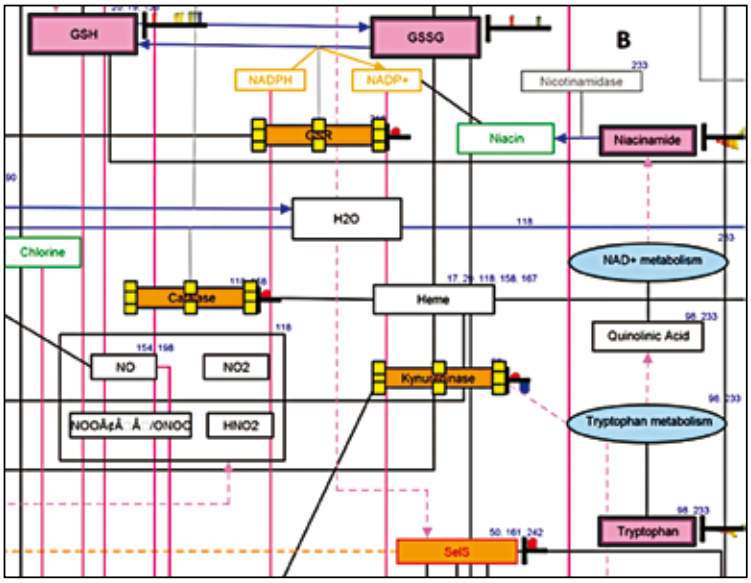


Figure 6. MPP software finds relationships beyond pair-wise comparison in complex experimental designs and enables biological interpretation through Pathway Architect.

Conclusions

This document has provided an overview of the Agilent selection of instrumentation, databases, and software tools that can be assembled to support a wide variety of metabolomics workflows. Sample characteristics including complexity, quantity, fidelity, and chemical characteristics such as polarity, are examples of some of the factors that help determine the best workflow choice for your research. In addition, there are often situations that require a combination of approaches, such as when both discovery and targeted metabolomics are required, or when the chemical diversity of compounds of interest cannot be adequately addressed by a single technology.

For all of these cases, Agilent's GeneSpring/MPP software suite can bring together a variety of biological data, including that from transcriptomics, proteomics, or genomics experiments, and analyze it in a biological context. These analyses can both produce new biological insights and provide guidance as to which follow-up experiments are likely to be the most fruitful.

References

1. Agilent Solutions for Metabolomics. Agilent publication 5990-6048EN.
2. The Fiehn GC/MS database, http://fiehnlab.ucdavis.edu/metabolite-library-2007/
3. http://metlin.scripps.edu/
4. Database Resources in Metabolomics: An Overview. www.springerlink.com/content/r3n3744t5504v058/
5. Chen, R., *et al*. Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell*, **2012**, 148(6):1293-1307.
6. Fischer, R., *et al*. Discovery of candidate serum proteomic and metabolomic biomarkers in ankylosing spondylitis. *Mol Cell Proteomics*, **2012**, 11(2):M111.013904.
7. Vareed, S.K., *et al*. Metabolites of purine nucleoside phosphorylase (NP) in serum have the potential to delineate pancreatic adenocarcinoma. *PLoS One*, **2011**, 6(3):e17177.
8. Shin, J.H., *et al*. 1H NMR-based metabolomic profiling in mice infected with Mycobacterium tuberculosis. *J Proteome Res*., **2011**, 10 (5):2238-2247.
9. Garcia, E. *et al*., Diagnosis of early stage ovarian cancer by 1H NMR metabolomics of serum explored by use of a microflow NMR probe. *J Proteome Res*., **2011,** 10 (4):1765-1771.
10. Woo, D.C., *et al*. Regional metabolic alteration of Alzheimer's disease in mouse brain expressing mutant human APP-PS1 by 1H HR-MAS. *Behav Brain Res*., **2010**, 211:125-131.

www.agilent.com/chem/metabolomics

Published in the USA, March 1, 2013
5991-1893EN

Agilent Technologies