What’s Next in ’Omics: The Metabolome

The youngest sibling in the family of ’omics fields is growing up. Maturing right behind genomics, transcriptomics, and proteomics, metabolomics is the comprehensive analysis of small molecule metabolites. Since most metabolites are generated by enzymatic proteins that result from gene expression, and metabolites give organisms their biochemical characteristics, the metabolome links genotype with phenotype. Metabolomics is still developing, though, as vendors adapt separation and detection equipment to meet its challenges and the research community interprets and integrates the complex data they are acquiring. By Chris Tachibana

We’re just a bunch of chemical reactions, says David Wishart, professor of biological science and computing science at the University of Alberta, Canada, “so the genome and proteome evolved to support the metabolome, not the other way around.” Metabolomics offers a more immediate measure of physiology than other ’omics approaches, says Wishart. The metabolome responds to nutrients, stress, or disease long before the transcriptome or proteome. This makes metabolomics an attractive approach for multiple fields: environmental toxicology, evolution and development, disease diagnosis and treatment response, and the development of drugs, pesticides, and herbicides. Metabolomic flux analysis aids synthetic biologists by revealing how genetic changes affect pathways and products.

Metabolomics, the youngest field in the ’omics family, is growing rapidly. In 2014, the International Metabolomics Society held its 10th annual conference in the same city in Japan as the first conference: Tsuruoka. “The first year, we had about 150 people,” says Ute Roessner, University of Melbourne associate professor and leader of the Metabolomics Australia node at her university. “This year we had more than 500 registrants. And metabolomics is becoming part of every researcher’s toolkit. Young researchers now routinely do metabolomics as part of their Ph.D.”

The major challenges of metabolomics stem from its advantages. Metabolomic data are powerful because organisms have many metabolites including related precursors, derivatives, and degradation products with concentrations that vary dramatically and change rapidly. This complexity demands sophisticated separation and detection methods. Metabolomics includes analysis of lipids, a developing area with tremendous potential, says Yingying Huang, strategic marketing manager at Thermo Fisher Scientific. “Lipidomics is directly related to cardiovascular disease, cancer, diabetes, and obesity,” she says, “but lipids have huge diversity in structure and concentrations that vary 10-fold in the body.” Many physiological lipids and other metabolites share a chemical composition, differing only in arrangement, which led University of California Davis Professor and West Coast Metabolomics Center Director Oliver Fiehn to call metabolomics, “the science of isomers.”

The instrument decision tree

Rising interest in metabolomics has led vendors of molecular separation and detection equipment to develop hardware, software, and support specifically for metabolomics. Researchers now have a complex decision tree to navigate when adding this technology to their laboratory. The first branch is untargeted vs. targeted analysis. The untargeted approach gathers data on all possible compounds. Targeted analysis focuses on known compounds, often to quantitate changes, for example in response to stress or disease.

Nuclear magnetic resonance (NMR) spectroscopy is particularly powerful for targeted analysis because it is quantitative, reproducible, and suitable for complex samples such as blood, urine, or tissue extracts with little or no processing. Peter Würtz, head of molecular epidemiology, University of Oulu, Finland, recently coauthored a paper that applied NMR metabolomics using Bruker instruments to thousands of Estonian and Finnish biobank samples. The research team discovered that four biomarkers, when combined, are significantly associated with increased risk of short-term death from all causes. The study was observational, with no exploration of mechanisms, notes Würtz, but shows the power of NMR for biomarker analysis. However, use of NMR for metabolomics is limited because its sensitivity is orders of magnitude lower than the more popular metabolomics method: mass spectrometry (MS).

MS requires sample processing. Compounds are often ionized and usually undergo separation, with gas chromatography (GC) and liquid chromatography (LC) the most...
common methods. Capillary electrophoresis (CE) gives excellent separation of polar metabolites and can be used with samples as small as a single cell. CE-MS is a specialty of Human Metabolome Technologies (HMT), but is also offered by others such as Agilent and AB SCIEX. However, CE requires specialized expertise and powerful software to account for drift in migration times, says Theodore Sana, metabolomics program manager at Agilent.

GC-MS is powerful for resolving compounds that are volatile or can be volatilized through derivation such as the small molecules of primary central metabolism. LC-MS is particularly suitable for large, thermo-unstable organic molecules including many secondary metabolites, larger carbohydrates, and lipids. Advances such as ultraperformance LC (UPLC)—originally developed by Waters—supercritical fluid chromatography, and ion mobility separation have increased speed and separating power, including for nonpolar molecules and compounds with many isomers and isobars such as lipids. Separation methods are coupled to mass analyzer options such as time-of-flight (TOF) detection MS, which is fast and has high resolution over a large mass range. Adding quadrupole or linear ion traps allows selection of particular compounds or fragments of compounds generated during ionization. Once researchers have sorted their metabolomics priorities, they can choose between buying equipment and sending samples to a service.

**Vendor and service options**

Agilent offers GC-MS, LC-MS, CE-MS, and NMR instruments. “We are in the unique position of offering all these technologies from a single vendor,” says Sana, “including hardware, software, and technical support.” (Shimadzu also offers multiple systems that researchers recommend as inexpensive starter setups.) Users compliment Agilent on their intuitive software and Sana confirms this is a company priority. An example is the Agilent GeneSpring-Mass Profiler Professional software for integrated analysis of multiple "omics data. "GeneSpring lets you take disparate data from RNA, proteins, and metabolites and metabolically map them on biological pathways like those from KEGG [a metabolic pathway tool]," says Sana. GeneSpring supports preprocessed results from other vendor platforms, for example as a spreadsheet. Visualizing results from a large number of experiments using a pathway tool is useful for gleaning biological insights and planning subsequent experiments, says Sana.

AB SCIEX specializes in LC-MS instruments that researchers praise for fast data acquisition, excellent dynamic range, and quantitation and accurate mass capabilities (generally defined as accuracy to five parts per million or lower). The company also offers CE and ion mobility options. For metabolomics, AB SCIEX recommends the Triple TOF and QTrap systems. “Our Triple TOF is a high-resolution, accurate mass instrument for qualitative and quantitative analysis, suitable for discovery metabolomics,” says Fadi Abd, senior market manager. “The Qtrap is for sensitive, targeted, quantitative workflows.” The high acquisition speed of the Triple TOF series allows simultaneous MS and tandem MS (MS-MS) from a single injection in an LC timescale, explains Abd, for identification, confirmation, and quantitation of compounds. Qtrap multiple reaction monitoring provides data on a parent compound and fragments successively generated by MS, simplifying data processing and giving researchers confidence about their data, says Baljit Ubhi, staff scientist in metabolomics and lipidomics applications at AB SCIEX.

Researchers looking for broad metabolite coverage can consider packages that combine high-resolution LC-MS systems from AB SCIEX with high-resolution GC-MS systems from LECO. The partnership offers clients “the best of both technologies,” says Jeff Patrick, LECO director of marketed technologies. The packages include software from Genedata that integrates data from the two platforms for comprehensive metabolomics analysis.

Other popular, versatile LC-MS options are the Thermo Fisher Scientific orbitrap-based instruments. Orbitrap mass analyzers are commonly used for proteomics but the high resolving power is particularly suitable for metabolomics when combined with an advanced separation system such as UPLC, says Huang. Parallel detection eliminates multiple runs for faster sample analysis. The newest system in the Thermo Scientific line is the Orbitrap Fusion Tribrid with three mass analyzers. After LC separation, sample filtering occurs in a quadrupole with parallel analysis possible in the subsequent orbitrap and linear ion trap. The Tribrid system was designed to offer researchers flexibility along with accurate mass capability, says Huang.

Thermo Fisher, AB SCIEX, and Bruker also offer instruments for imaging MS, also called MALDI imaging. Roessner explains that this method overcomes a metabolomics limitation: metabolite analysis is usually done on homogenized samples or extracts, but specialized cells and tissues have different metabolite profiles. Imaging MS gives data on the spatial distribution of metabolites from a thin tissue section, for example from an organ or plant tissue. The process is laser desorption of a matrix-treated sample to ionize compounds that are then analyzed by MS. A reconstructed two-dimensional image of the compounds provides the location of metabolites across the tissue.

“It’s still early, but it’s an exciting method,” says Roessner, who uses a Bruker system for this technique. “It can be used to see the distribution of drug-related metabolites in a liver sample or to follow alterations in metabolite distribution in tissues after environmental stimuli.” The Thermo Scientific Orbitrap is advantageous for this method, Huang says. “Since the ions are coming from the tissue with no LC separation, you need an ultrahigh-resolution instrument like the Orbitrap.”

 Capillary electrophoresis gives excellent separation of polar metabolites and can be used with samples as small as a single cell.
Researchers who aren’t ready to purchase a system or prefer the convenience of a service have several options. Biocrates and Metabolon offer MS metabolomic analysis worldwide. Metabolon is also involved in large, long-term collaborations, such as Craig Venter’s Human Longevity project, says Chris Bernard, Metabolon senior vice president, sales and marketing. Collaborating is important for Metabolon, says Bernard, for seeing how metabolomics is practically implemented and how it is used in a range of projects.

Scientists interested in the advantages of NMR metabolomics—true quantitation without standard curves, sample preservation and minimal sample processing, and detection of all metabolites with structural information in a single run—might be hindered by the startup costs. Researchers can access this method through services, for example from Chenomx. Würtz and colleagues also offer NMR metabolomic services and collaborations through Brainshake, founded on their experience using NMR for biomarker research and linking metabolomics data to genome-wide association studies that connect genetic variation with disease risk.

Metabolomics services are increasingly available through national and academic core facilities. In Australia, Roessner established one of the first government-funded national metabolomics service facilities. Canada is also funding national metabolomics initiatives and other countries are following suit. In the U.S., the National Institutes of Health (NIH) Common Fund has been establishing regional metabolomics cores since 2012. Fiehn leads the West Coast program. "The cores increase metabolomics capacity and services for both academia and industry and they have different specialties," he says, "such as metabolic flux in the Kentucky center and NMR in North Carolina." The NIH core facilities also offer short courses for scientists new to metabolomics.

The next challenges

When asked about the major challenges facing metabolomics, researchers and vendors consistently name three issues: identification of unknowns, development of standardized data repositories that can be queried like the genomics resource GenBank, and integration of metabolomics and other systems-wide data.

“MS is so sensitive that we get lots of compounds and we don’t know what they all are,” says Wishart. To help researchers with metabolite identification, Agilent has curated GC-MS and LC-MS spectral libraries for metabolite identification with MS-based techniques. LECO offers a GC-TOF library that Patrick says can be used with any GC data, although it is optimized for LECO instruments. LECO is currently exploring opportunities to collaborate on high-resolution accurate mass libraries.

Wishart notes that most unknowns are metabolites of metabolites: “breakdown products and molecules transformed by enzyme or microbial activity.” This molecular similarity can be the key to identifying them. The Fiehn group, which developed several of the commercially available MS libraries, is also producing virtual libraries. Starting with known spectral data on compounds, the researchers predict spectra of similar or modified variants. LipidBlast is the Fiehn group’s free downloadable MS-MS virtual spectra library for more than 200,000 lipids.

To meet challenges of searchability and data integration, the metabolomics community has several initiatives to establish data repositories. Examples are Metabolights in the United Kingdom, supported by the European COSMOS (COordination in Standards in MetaboLomicS) consortium that is developing metabolomics data standards, and Metabolomics Workbench, which aims to be the database for NIH-funded metabolomics projects. Thermo Fisher collaborates with Fiehn on mzcloud.org, a free community database that includes actual and virtual MS spectra with unknown compounds to be annotated as they are identified.

Even after the metabolomics community has databases to corral its data, Patrick notes that “an overarching challenge for systems biology is biological inference”—giving context and meaning to large amounts of data. The solution, he says is software, which “is driving the direction of metabolomics.” Bernard agrees, saying that the number-one request of Metabolon’s clients is help in understanding the biological meaning of results. “MS is wonderful for generating data,” he says. “The challenge is separating signals from noise: 85% of MS data points is noise.” Metabolon clients have access to the MetaboLync portal which includes a list of metabolites in their sample with comparison to an in-house library of over 14,000 compounds for automatically verifying metabolite identification, says Bernard, adding that the most powerful feature of the portal is the ability to visualize and explore data at the pathway level.

As the field grows, vendors and service providers are eager to collaborate with and hear from users about their metabolomics research, not just for sales but to stay current. “We collaborate closely with several investigators,” says Sana. “Researchers are frequently the source of novel ideas for advancing the field of metabolomics.”

Chris Tachibana is a science writer based in Seattle, USA, and Copenhagen, Denmark.

DOI: 10.1126/science.opms.p1400088