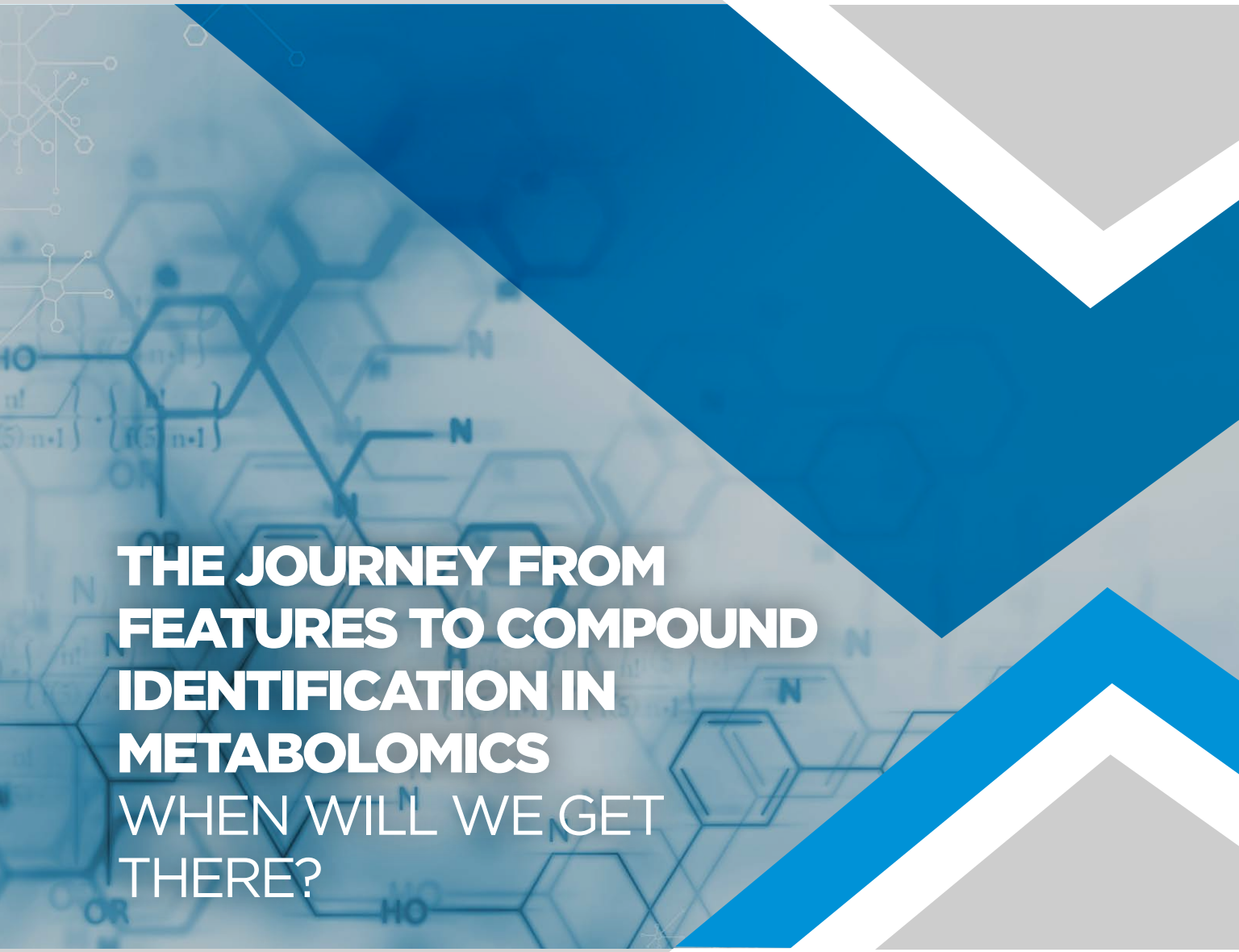
Custom Publishing From:

TheScientist
EXPLORING LIFE, INSPIRING INNOVATION

Sponsored By:

ThermoFisher
SCIENTIFIC



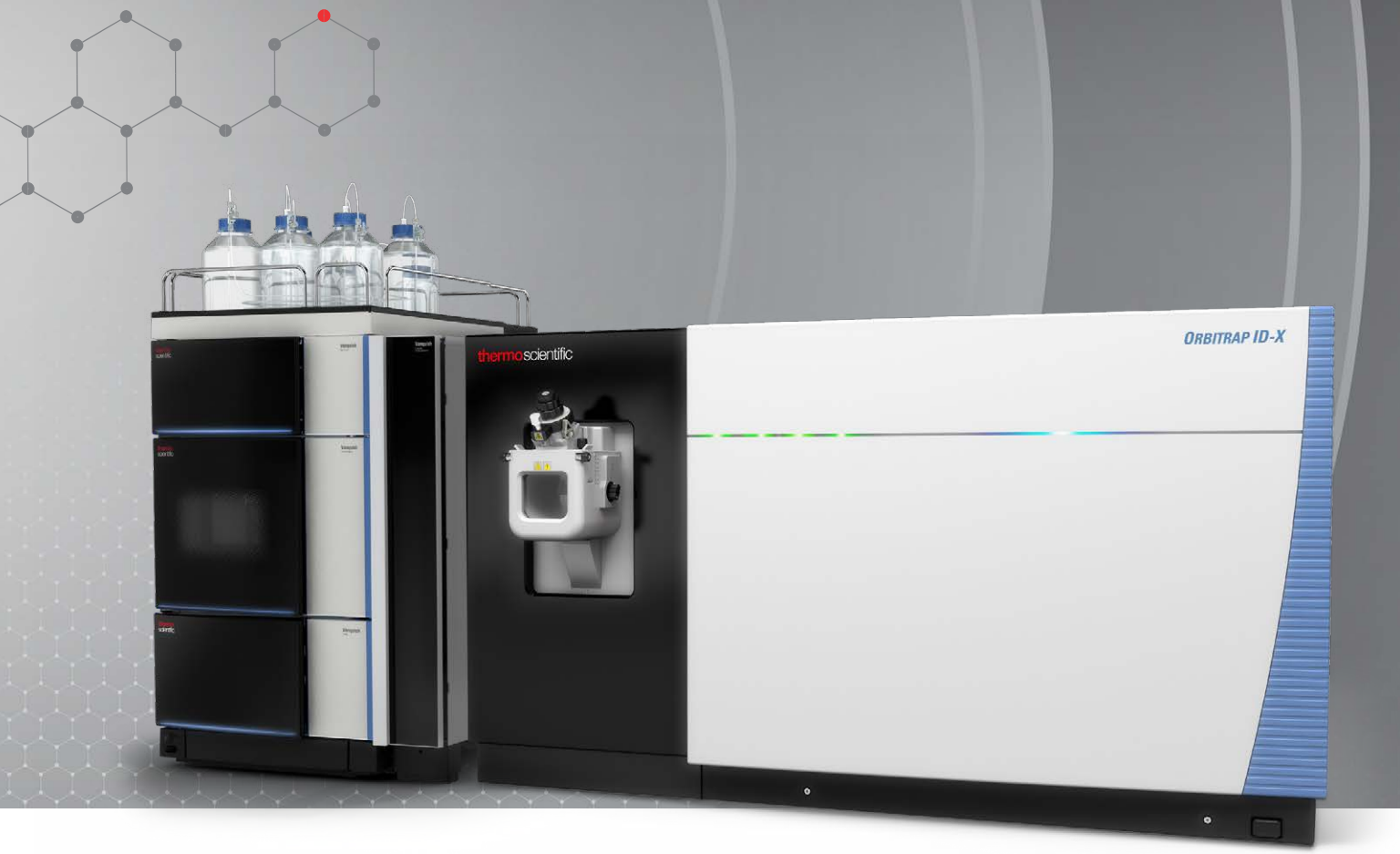
THE JOURNEY FROM FEATURES TO COMPOUND IDENTIFICATION IN METABOLOMICS

WHEN WILL WE GET THERE?

Contributions from the following leading researchers: Dr. Gary Patti (Washington University), Professor Warwick (Rick) Dunn (University of Birmingham), Dr. Darren Creek (Monash University), and Dr. Lloyd W. Sumner (University of Missouri).

thermoscientific

Go beyond
with technology and intelligence



Transform compound identification and characterization

The Thermo Scientific™ Orbitrap ID-X™ Tribrid™ mass spectrometer combines industry-leading mass analyzer technology dedicated for small molecule analysis with the Thermo Scientific™ AcquireX intelligent MS^n data acquisition strategy, to collect more meaningful data. By routinely achieving comprehensive metabolome coverage with confident compound annotations against spectral libraries and de novo structural elucidation of unknown metabolites, the system's intuitive automated workflow for untargeted metabolomics addresses the identification bottleneck to let you go beyond the way metabolomics is done today.

Advance quicker at **thermofisher.com/orbitrapID-X**

ThermoFisher SCIENTIFIC

Metabolomics is a relative newcomer in the -omics world that's quickly gaining traction in multiple areas of research. However, compound identification in untargeted metabolomics is still a major bottleneck. The task of annotating an unknown compound from a mass-to-charge ratio measurement comes with twists and turns in the large chemical space of small molecules, but a collective effort from the metabolomics community is helping to propel the field in the right direction.

Leaning In To Learn: Data Reduction Strategies

Contributions by Dr. Gary Patti (Washington University)

Metabolomics is at the bottom of the -omics cascade, after genomics, transcriptomics, and proteomics. If a particular metabolite is changing, it is indicative of a certain pathway being altered. With DNA, RNA, and proteins that's not always the case; for example, a gene or protein might be present or not, but neither scenario has a direct consequence on the phenotype.

As Dr. Gary Patti, the Michael and Tana Powell Associate Professor of Chemistry in Arts & Sciences at Washington University in St. Louis explains, "The nice thing about metabolites is that there's downstream information about [a particular] cascade and that they tend to correlate more with the phenotype, and in some ways, actually are the phenotype."

There are two ways of conducting a metabolomics experiment: using an untargeted approach, where you attempt to measure every metabolite in a sample including "unknowns", or targeted metabolomics, where you already know what metabolites you are looking for.

Something that's always been exciting for Patti about untargeted metabolomics is the possibility of capturing some of these "unknowns", and in this context he's talking about metabolites that are present in the cell that have never been characterized before. "Personally, that's always been a real draw for me. It's so exciting to me to think that you can take a sample of blood and run it through an instrument and there could be a chemical, a molecule in there that's floating around that no one has ever thought about before. For me, from a chemical perspective, that's just extremely fascinating."

And capturing these "unknowns" is becoming easier; advancements over the past several years in instrumentation and metabolite annotation are pushing the field into territories it's never been before. In Patti's opinion, "the most significant advances have been in converting that really cumbersome excel sheet to the beautiful pathways that we see hanging on the lab wall." He recalls a time when years ago as a post doc, after several weeks of manually trying to validate a mass spectrometry signal from a metabolomics experiment, he had a mini celebration with his lab mates after finally elucidating the metabolite's identification. "The resources that were available to try to do those identifications were just extremely limited, there just wasn't a lot out there to convert those signals into metabolites," Patti recollects.

Many steps in metabolite identification have now been automated. However, the problem of going from hundreds – or sometimes thousands – of signals to real potential compounds has not disappeared completely.

Patti explains that one of the knee-jerk reactions in metabolomics experiments is to assume that because something doesn't exist in any of the metabolomics databases, it must be a molecule that's never been characterized before. In reality, there are just a couple of thousand metabolites that have been described. But the number of signals in a metabolomics experiment is an order of magnitude more than that. As Patti explains, "I think one of the challenges has been why is there this big disparity? We have to be careful [to consider] that there are other things that are contributing to these signals."

Over the last five to 10 years, a new perspective on "unknown" signals has emerged within the metabolomics community, and a number of groups have contributed to this new way of thinking. Patti explains: "I think what's happened over the years is that people are starting to discover that a lot of these molecules that we thought were unknowns have been found to be artifacts in some ways – sometimes they might be plasticizers, or fragments of molecules for example (...) or some product of something else that isn't listed in a database."

Patti's lab has found that the actual number of metabolites in their average untargeted metabolomics experiments may be less than one-tenth of the total number of signals detected. By defining the "noise" signals coming from contaminants, metabolite degeneracy, and other artifacts, Patti's team reduced an E. coli dataset containing more than 25,000 unique features to less than 1,000 unique metabolites.[1]

When asked why unknown signals cannot simply be identified and "noise" signals quickly filtered out in a routine metabolomics experiment, Patti explains that, unlike in a proteomics experiment, where you can start to put together the proteome from the genome, the metabolome doesn't really have any kind of "blueprint". But, as Patti says, "That's what makes metabolomics so fun, from a chemical perspective. There's no parts list for the metabolome."

For references see page 10

Singling Out Signals: Guidelines for Compound Identification

Contributions by Professor Warwick (Rick) Dunn (University of Birmingham)

In 2007, Sumner et al. published "Proposed minimum reporting standards for chemical analysis",[1] which was one of a series of papers where The Metabolomics Standards Initiative was looking to develop the reporting standards that the metabolomics community could apply to datasets.

These reporting standards proposed four levels of confidence for metabolite identification, ranging from level 1, confidently identified compounds, to level 4, unknown compounds.

A problem that's arisen in the community, which makes it more difficult to compare datasets between labs, is that most of the reporting standards aren't used routinely. There are many reasons why they aren't being used consistently but one of the main reasons is that there's been "more of a carrot rather than a stick approach," explains Professor Warwick (Rick) Dunn, Professor of Analytical and Clinical Metabolomics, School of Biosciences and Phenome Centre Birmingham, University of Birmingham, UK.

What he means by that is that journals and data repositories aren't insisting that the standards be followed, so the impetus isn't always there for labs to follow the reporting standards.

There's now a task group within The Metabolomics Society called The Metabolite Identification Task Group. They are currently working on 1) revising the reporting standards, and 2) educating the metabolomics community on how to apply them in more detail.

Professor Dunn, who is a co-chair of the new task group explains that there will be a new set of reporting standards defined this year (2018) and that there are also a number of journals and data repositories that will take them on board, so any metabolomics data that's submitted will have to follow these new reporting standards. "We're moving more towards a stick rather than a carrot approach," says Dunn.

Once the reporting standards have been defined the guidelines will go out to the metabolomics community for comment before they are published in a peer-reviewed journal.

The new reporting guidelines will help to standardize data reporting within metabolomics, however, Dunn cautions that it won't completely fix the issue with how metabolites are annotated: "There are guidelines to report what you've done, which is what the reporting standards are about, and then there are guidelines about what you should do to be able to annotate your metabolites. So, I think it will push us forward for the reporting standards, but the area of metabolite annotation and identification is still really developing within the metabolomics community."

In the future however, Dunn thinks there are various areas where we will see the most advancement within metabolite identification. He predicts that in the next two to three years we will be seeing a lot more of "semi-targeted" metabolomics, which is a combination of targeted and untargeted metabolomics where, for example, the known metabolites can be reported with high confidence, but the "unknown" metabolites can also be identified as well using a combination of different approaches.

He also predicts that we will also see huge improvements in the transfer of data to knowledge, where a researcher will be able to jump directly from obtaining data into deriving a conclusion.

However, the most exciting advancements may come from a field that his own research group calls "metabolome annotation". For example, the human metabolome database is an excellent database for studying humans, but it's largely derived from the scientific literature and there are a whole range of metabolites in it that have been predicted but not necessarily ever detected. As Dunn explains, "So there's a very large database of what could be present [in a sample], but actually what we detect is a much smaller component of that, and that provides difficulties in identifying metabolites."

Being able to characterize different metabolomes that we analyze and the metabolites that we detect, using more sensitive instrumentation along with semi-targeted approaches, removes the identification barrier, and enables researchers in metabolomics to define the "parts list" of everything they'd expect to detect. Dunn says, "I think that's a really important thing we'll want to do in the future, to actually experimentally characterize the metabolomes that we're measuring."

One thing that's clear is that the metabolomics community is banding together, and the future looks full of promise when it comes to standardization in compound identification.

For references see page 10

Advancement of Databases and Libraries

Contributions by Dr. Darren Creek (Monash University), and Dr. Gary Patti (Washington University)

Over the past several years, a wealth of libraries and databases containing information relating to spectra from tandem mass spectrometry (MS/MS), gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS) and nuclear magnetic resonance (NMR) analyses, as well as metabolic pathways, statistical models, compound structures, and metabolites have been developed. Databases contain chemical information, whereas libraries are databases that also have searchable fragment spectral data.

Researchers use these tools to identify MS and NMR spectra peaks, assist in peak alignment, exclude false positive, and other manipulations to analyze data gathered from metabolomics experiments by comparing them to values present in databases or libraries. Databases and libraries are built using information from the literature and authentic standards and/or via computational prediction methods.

Some of these databases and libraries have seen huge expansion over the past decade. For example, The Human Metabolome Database (HMDB) is a freely available web resource containing detailed information about the human metabolome and was one of the first comprehensive, organism-specific metabolomic databases. It was first released in 2007 with information pertaining to 2,180 human metabolites. Since then, it has undergone four major upgrades; HMDB 4.0 now contains data for upwards of 114,000 metabolites.[1] The new database mzCloud has also seen fast expansion;[2] it comprises freely searchable curated databases of high and low resolution multi-stage mass spectrometry (MS^n) spectra acquired under a number of experimental conditions, and as of June 25th 2018 had 8,081 compounds analyzed in different conditions, totaling 2,763,141 spectra arranged in 12,243 tree structures.

As described by Wishart et al.,[1] "This expansion has been motivated by an emerging crisis in metabolomics – a severe bottleneck in metabolite identification."

Typically, less than 2% of MS peaks from untargeted metabolomics MS-based experiments can be identified, suggesting that "existing chemical compound data sets and existing experimental spectral data sets for metabolomics are far too inadequate for comprehensive metabolite identification or quantification."[1]

In silico approaches to overcome this conundrum have been pivotal in the expansion of databases like the HMDB. Without them, using only traditional curation techniques, it's predicted[1] that databases would take decades to expand enough to achieve the realistic metabolome predictions that are required in untargeted metabolomics.

Metabolite repositories have also advanced by leaps and bounds in the last decade. There are two main types of repository for metabolites: the first focused on metabolite information regardless of source, and the second more specific type, which allows contextualization of submitted experiments by including experimental metadata.[3]

Popular computational tools for metabolite and spectral prediction include the Mass Spectrometry Interactive Virtual Environment (MassIVE),[4] Competitive Fragmentation Modeling Identification (CFM-ID),[5] the databases PubChem,[6] ChemSpider,[7] MetaboLights,[8,9] and the library METLIN,[10] to name just a few.

In fact, countless databases are now available, and although some consolidation of databases might be welcome in the future, the metabolomics community still sees the need for multiple databases to exist. As Dr. Darren Creek, NHMRC Career Development Fellow, and Metabolomics Director of the Monash Proteomics and Metabolomics Facility explains, "they all serve different purposes." Different databases are appropriate for different organisms, instruments, and applications. But, there is room for improvement: "There is some subjectivity in their use," Dr. Creek explains. Many different parameters can be used to classify a spectra as a match to something in a database. "And if you have just one person or 10 people designing a database, they will do it in different ways. But the key [to improving them] is better annotation." The challenge, he says, is working out who will pay for and maintain them.

Dr. Gary Patti, the Michael and Tana Powell Associate Professor of Chemistry in Arts & Sciences at Washington University in St. Louis stresses that the metabolomics community as a whole could also benefit from more people adding to currently available databases: "the number of people uploading their data in shared repositories is still not what one would expect."

Database and repository advancements allow the metabolomics community to advance new discoveries in biology.[11] Updates to metabolomic databases and repositories have been essential in, for example, the discovery of oncometabolites,[12] damaged metabolites,[13] and microbiota metabolites.[14] However, *in silico* approaches are not always the answer: Matching metabolite spectra against huge databases can give rise to false-positive identifications, databases sometimes contain erroneous spectra, and instrument settings are not always included.[15]

So, database and repository improvements are still necessary. Metabolomics tools suffer from added complications that aren't present in other -omics due to the inherent complexities of metabolites in comparison to genes and proteins for example. Their "building blocks" cannot be reduced to four or five nucleotide bases in the same way that gene sequences can, and their molecular structures are extraordinarily varied.

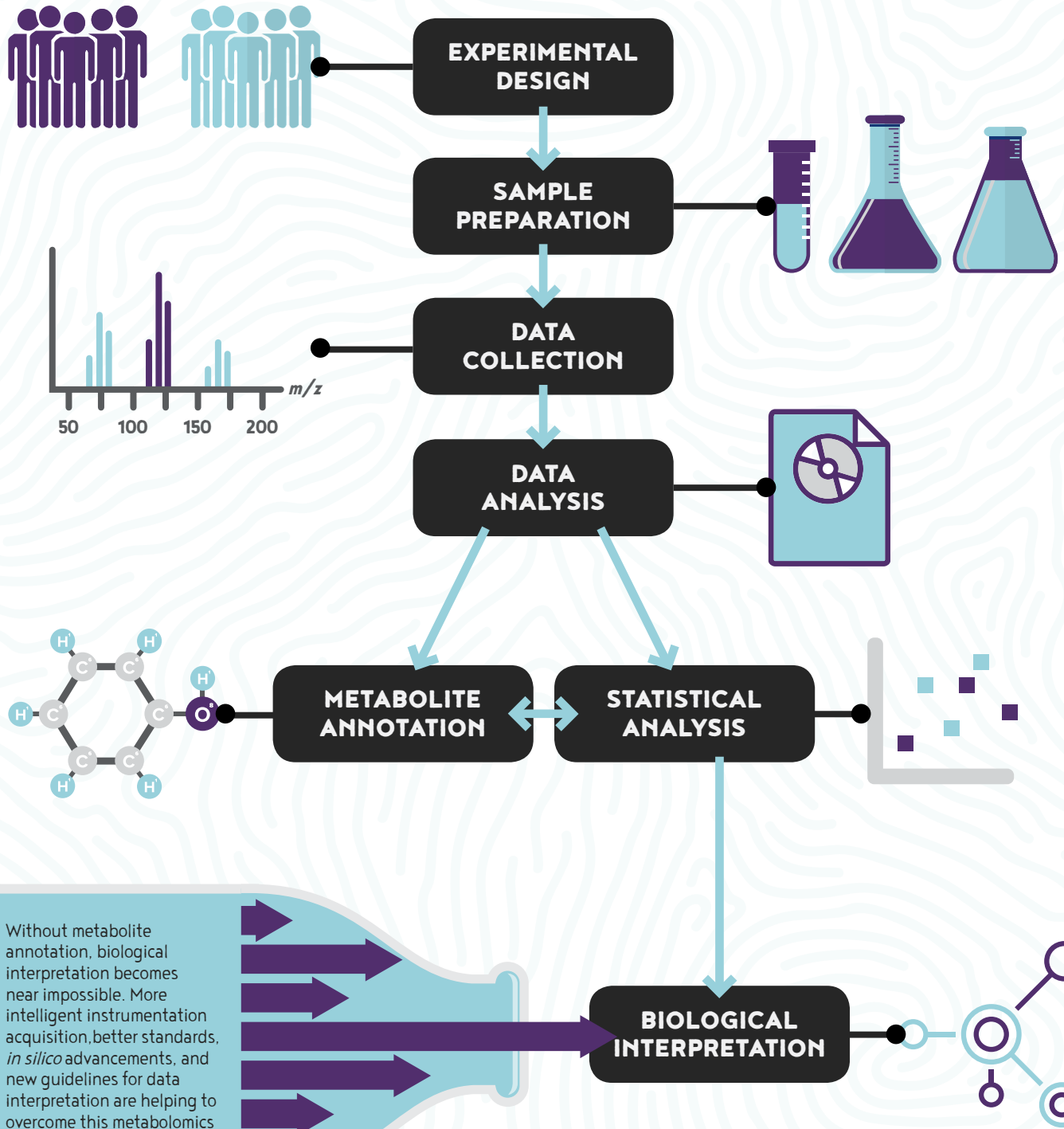
Furthermore, in MS analysis, metabolites may be present as multiple adducts with different mass to charge ratios and different instruments and preparation techniques in metabolomics analyses introduce further variability.

The ultimate goal for metabolomics databases and libraries is to make metabolomics analysis quicker, cheaper, and easier. Improvements are still needed to organize metabolites and their metadata to facilitate inter-lab, inter-instrument, and inter-experiment comparisons. However, advances in the last decade have been numerous and major, and this trend looks set to continue.

For references see page 10

PROFILING THE PHENOTYPE

Untargeted Metabolomics Workflow



Without metabolite annotation, biological interpretation becomes near impossible. More intelligent instrumentation acquisition, better standards, *in silico* advancements, and new guidelines for data interpretation are helping to overcome this metabolomics bottleneck.

Predicament in Progress: Problems with Standards

Contributions by Professor Warwick (Rick) Dunn (University of Birmingham), Dr. Darren Creek (Monash University) and Dr. Gary Patti (Washington University)

In 2009, at the American Society of Mass Spectrometry Workshop, Oliver Fiehns' research group surveyed the metabolomics community and found that the two perceived biggest bottlenecks were the identification of metabolites (35%) and the assignment of biological interest (22%).[1] Professor Warwick (Rick) Dunn, Professor of Analytical and Clinical Metabolomics, School of Biosciences and Phenome Centre Birmingham, University of Birmingham, UK, predicts that if you asked that same question now, around 20% of people would still say that metabolite identification is the biggest bottleneck in metabolomics research.

However, "we are certainly much further along the path to easier and higher quality annotations," Dunn explains. "We have better databases, better and larger mass spectral libraries, and there's more experimental data out there which is available."

One of the biggest issues that affects compound annotation is that there just aren't yet the chemical standards available to allow metabolites to be annotated with high confidence. As Dunn says, "That's one of the difficulties. (…) looking from outside in, you can see that if it's not communicated appropriately, it looks like this fishing trip where we don't necessarily know what we're measuring. Within the [metabolomics] community we know that there is a problem, and a lot of different groups are working on different solutions to be able to overcome it."

Dr. Darren Creek, NHMRC Career Development Fellow, and Metabolomics Director of the Monash Proteomics and Metabolomics Facility highlights the importance of standards in metabolomics: "Reference standards I think are really important in metabolomics. Our lab routinely runs hundreds of reference standards every week to make sure that we're getting accurate retention times and that everything is performing well. In doing that we have much more confidence in our data."

He describes that increased uses of reference standards has really only happened in the last couple of years, thanks to more companies offering them in larger quantities and in groupings: "It's [now] much easier to access pure metabolite standards for most of the metabolites from the central metabolic pathways." The, bottleneck, Creek explains, is really when you go outside of those central metabolic pathways; "and you're looking at secondary pathways with lots of unique metabolites, especially for plants or exotic organisms."

But as Dunn explains, not having the authentic chemical standards doesn't mean that you can't get a relatively high confidence. Thanks to *in silico* advancements within the community there are many programs that can be used to work out the mass retention times and other properties, "so that adds an extra level of confidence," he says.

He explains that we're never going to get to a point where we have standards for everything and have that highest level of identification and confidence for everything. But, "we're part way along a route of solving that bottleneck and improving what we can do in that bottleneck."

For example, for the human metabolome, less than 10% of known human metabolites have authentic, experimentally collected nuclear magnetic resonance (NMR), gas chromatography-mass spectrometry (GC-MS), or tandem mass spectrometry (MS/MS) spectra.[2]

Experimental solutions to this problem are predicted to take decades of research, and to cost billions, and therefore *in silico* approaches will help to solve this conundrum.[3] However, improvements to experimental approaches are also making a difference: "We're getting better at calling the knowns 'knowns', so there are certain things that we know we should detect," Dunn explains. For example, "glucose in blood. If we don't [detect it] then there's something wrong with the experiment or the person."

Kits containing authentic standards have also progressed rapidly in the past decade, furthering metabolite identification particularly in targeted metabolomics. These kits are not only useful in standardizing metabolomics experiments within a lab, but also help to standardize experiments between labs, so that data collected on different instruments for example may in the future be more easily compared.

Isotopic approaches to compound identification are also helping to crack the metabolite identification bottleneck. Stable isotopes (*e.g.*, compounds containing ^{13}C or ^{15}N) can be introduced at various stages of a metabolomics experiment and can be used to quantify metabolites relative to their isotopically labeled equivalents. This is useful for the simultaneous analysis of multiple metabolites.[4]

In fact, stable isotopes are useful in a multitude of ways in metabolomics. As Dr. Gary Patti, the Michael and Tana Powell Associate Professor of Chemistry in Arts & Sciences at Washington University in St. Louis describes, "our ability to track these isotopes, both from an instrumentation perspective and a biochemical perspective, means we can see that the same metabolite sometimes comes from multiple metabolic pathways, each constituting a fraction of the total metabolite pool." In effect, the unknown "arrows" on those pathway wall charts that are scattered on every lab wall can be deciphered, as well as the structures. "What's really exciting to me is that we can see not just unknown molecules, but unknown pathways," Patti says.

The metabolomics community is banding together when it comes to the metabolite identification bottleneck. With recent improvements in experimental approaches, increased numbers of authentic standards, and huge advancements in *in silico* approaches, the identification bottleneck may soon be a problem of the past.

For references see page 10

In Fine Detail: Structural Elucidation Strategies

Contributions by Dr. Lloyd W. Sumner (University of Missouri)

Mass spectrometry (MS) is the most common method for performing metabolomics experiments, with nuclear magnetic resonance (NMR) coming a close second.

Mass spectrometers can be coupled to a variety of other analytical instruments to perform metabolomics: liquid chromatography (LC-MS), gas chromatography (GC-MS), and tandem MS (MS/MS) are common setups.

When it comes to mass spectrometry and metabolomics (which is and always will be a quantitative science), the higher the resolving power of the mass analyzer, the more accurate peak assignment becomes. And huge advancements in mass analyzers over the past decade mean that MS-based metabolomics is becoming easier, more accurate, and faster.

The resolving power of a mass spectrometer differs between mass analyzers; Ion trap analyzers have unit mass resolution, Time-of-Flight (TOF) based analyzers have a resolving power of up to 60k, orbital ion trapping analyzers can reach one million, and Fourier-transform ion cyclotron resonance-based (FT-ICR) spectrometers can reach as high as twenty million. It's this difference in resolving power that puts hybrid orbital ion trapping analyzers and FT-ICR based-mass analyzers into the high-resolution, accurate mass (HRAM) category.

HRAM-MS is capable of detecting minute concentrations of metabolites, aiding in metabolite annotation and quantification. High-resolution instruments have become more popular in metabolomics research because of these properties, but are they ready to become mainstream in the metabolomics community?

Dr. Lloyd W. Sumner, Professor of Biochemistry and Director of the University of Missouri Metabolomics Center states that HRAM MS instruments really are already mainstream in the community, "and we as scientists appreciate that and are trying to use them."

HRAM instruments are essential for untargeted metabolomics experiments. There are three standards that are typically used to identify metabolites: "First we try to purify them, then we use accurate mass MS or tandem MS/MS to get molecular formulas and / or structures, and then ultimately we try to put them into three-dimensional space by using other types of NMR or spectroscopy. So accurate mass MS is fundamental to that process," Sumner explains.

Various groups have shown the benefits of using high resolution accurate mass techniques.[1-6] For example, in an evaluation performed to show the differences between a unit mass resolution (UMR) ion-trap mass spectrometer in comparison with an HRAM MS, ion trap mass analyzer instrument, the HRAM data stream detected 118 additional known compounds in human serum, leading to the identification of more than 500 identified compounds.[7]

Dr. Sumner describes other instrumentation improvements that have been seen over the past decade: "One of the exciting things that we're seeing is coupling of MS to other technologies, for example ion mobility MS. We look forward to that being another multi-dimensional tool that provides orthogonal data to help [the metabolomics community] with metabolite identification."

Another area where great improvements are being made is in the coupling of MS with NMR: "We are actually developing LC-MS coupled to solid phase extraction for not only comparative profiling but also automated purification of targets for NMR spectral data acquisition," Dr. Sumner explains. With these types of instrument ensembles, comparative analysis (i.e., healthy vs. disease analysis) is becoming easier, and target identification to really understand the biology is also becoming more straightforward. "We can actually get to that identification in a much quicker and more confident manner using these [ensemble] approaches," says Sumner.

Given the metabolite annotation bottleneck and the highly diverse nature of metabolites, highly accurate mass analyzers are a helpful tool in metabolomics. They are particularly useful in metabolomics experiments involving stable isotopes, as they have the resolution power to distinguish more easily between "native" and isotopically labelled metabolites.

And instrumentation for metabolomics is only getting better. As Sumner states, "Other types of technologies will hopefully evolve as well to couple multiple technologies into some type of innovative platform. And as we develop new tools, we will continue to push forward the scientific promise and outcomes of metabolomics." Multi-stage mass spectrometry (MS^n) trees show promise in this regard. Fragmentation trees aid in sub-structure identification, and mass spectral trees delineate the dependencies in multi-stage MS of collision-induced dissociations.[8]

With increased accessibility and availability, HRAM mass analyzers look set to become an increasingly common site in metabolomics labs across the globe.

For references see page 10

Article 1 - Leaning In To Learn: Data Reduction Strategies

References

1. N.G. Mahieu and G.J. Patti, "Systems-Level Annotation of a Metabolomics Data Set Reduces 25,000 Features to Fewer than 1000 Unique Metabolites," *Anal Chem* 89(19): 10397-10406, 2017.

Article 2 - Singling Out Signals: Guidelines for Compound Identification

References

1. L.W. Sumner et al., "Proposed Minimum Reporting Standards for Chemical Analysis," *Metabolomics* 3(3): 211-221, 2007.

Article 3 - Advancement of Databases and Libraries

References

1. D.S. Wishart, et al., "HMDB 4.0: The Human Metabolome Database for 2018," *Nucleic Acids Res* 46(D1):D608-D617, 2018.
2. "m/z cloud, Advanced Mass Spectral Database," www.mzcloud.org. Accessed on June 20, 2018.
3. K. Longnecker et al., "Environmental Metabolomics: Databases and Tools for Data Analysis," *Marine Chemistry* 177:366-373, 2015.
4. "Mass Spectrometry Interactive Virtual Environment (MassIVE)," massive.ucsd.edu. Accessed on June 20, 2018.
5. F. Allen et al., "CFM-ID: A Web Server for Annotation, Spectrum Prediction and Metabolite Identification from Tandem Mass Spectra," *Nucleic Acids Res* 42:W94-W99, 2014.
6. E.E. Bolton et al., "PubChem: Integrated Platform of Small Molecules and Biological Activities," *Annual Reports in Computational Chemistry*, 2008.
7. H.E. Pence and A. Williams, "ChemSpider: An Online Chemical Information Resource," *J Chem Educ* 87:1123-1124, 2010.
8. K. Haug et al., "MetaboLights-An Open-Access General-Purpose Repository for Metabolomics Studies and Associated Meta-Data," *Nucleic Acids Res* 41:D781-D786, 2013.
9. C. Steinbeck et al., "MetaboLights: Towards a New COSMOS of Metabolomics Data Management," *Metabolomics* 8:757-760, 2012.
10. C. Guijas et al., "METLIN: A Technology Platform for Identifying Knowns and Unknowns," *Anal Chem* 90(5):3156-3164, 2018.
11. H. Tsugawa, "Advances in Computational Metabolomics and Databases Deepen the Understanding of Metabolisms," *Curr Opin Biotechnol* 54:10-17, 2018.
12. J.W. Locasale, "Serine, Glycine and One-Carbon Units: Cancer Metabolism in Full Circle," *Nat Rev Cancer*, 13:572-583, 2013.
13. C.L. Linster, et al., "Metabolite Damage and its Repair or Pre-emption," *Nat Chem Biol* 9:72-80, 2013.
14. M.G. Rooks and W.S. Garrett, "Gut Microbiota, Metabolites and Host Immunity," *Nat Rev Immunol*, 16:341-352, 2016.
15. I. Aretz and D. Meierhofer, "Advantages and Pitfalls of Mass Spectrometry Based Metabolome Profiling in Systems Biology," *Int J Mol Sci* 17:632-645, 2016.

Article 4 - Predicament in Progress: Problems with Standards

References

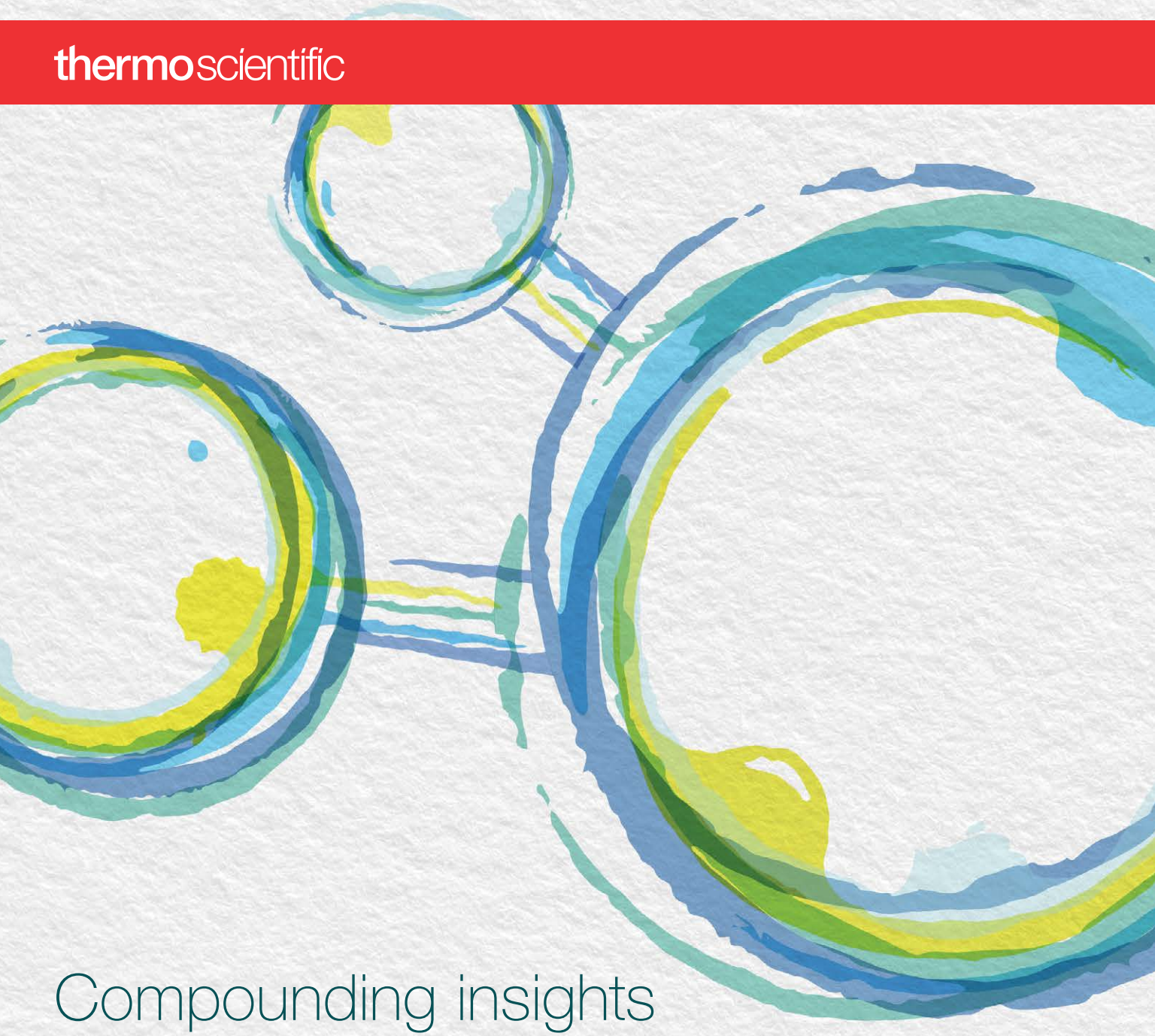
1. American Society for Mass Spectrometry, 2009.
2. D.S. Wishart, et al., "HMDB 3.0-The Human Metabolome Database in 2013," *Nucleic Acids Res* 41(Database issue):D801-D807, 2013.
3. D.A. Dias et al., "Current and Future Perspectives on the Structural Identification of Small Molecules in Biological Systems," *Metabolites* 6(E46), 2016.
4. F.A. de Jong and C. Beecher, "Addressing the current bottlenecks of metabolomics: Isotopic Ratio Outlier Analysis™, an isotopic-labeling technique for accurate biochemical profiling," *Bioanalysis* 4(18):2303-2314, 2013.

Article 5- In Fine Detail: Structural Elucidation Strategies

References

1. L. Najdekr et al., "Influence of Mass Resolving Power in Orbital Ion-Trap Mass Spectrometry-Based Metabolomics," *Anal Chem* 6;88(23):11429-11435, 2016.
2. J. McHale et al., "Metabolomics of Hermaphroditic C. elegans via Isotopic Ratio Outlier Analysis using High-Resolution Accurate Mass LC/MS/MS," conference paper, ASMS 2013.
3. B. Rochat, "Quantitative/Qualitative Analysis Using LC-HRMS: The Fundamental Step Forward for Clinical Laboratories and Clinical Practice," *Bioanalysis* 4(14):1709-1711, 2012.
4. N.R. Zhang et al., "Quantitation of Small Molecules Using High-Resolution Accurate Mass Spectrometers – A Different Approach for Analysis of Biological Samples," *Rapid Commun Mass Spectrom* 23:1085-1094, 2009.
5. A. Kaufman et al., "Quantitative and Confirmative Performance of Liquid Chromatography Coupled to High-Resolution Mass Spectrometry Compared with Tandem Mass Spectrometry," *Rapid Commun Mass Spectrom* 25:979-992, 2011.
6. R. Ramanathan et al., "It is Time for a Paradigm Shift in Drug Discovery Bioanalysis: From SRM to HRMS," *J Mass Spectrom* 46:595-601, 2011.
7. A.M. Evans et al., "High Resolution Mass Spectrometry Improves Data Quantity and Quality as Compared to Unit Mass Resolution Spectrometry in High-Throughput Profiling Metabolomics," *Metabolomics* 4:132, 2014.
8. A. Vaniya and O. Fiehn, "Using fragmentation trees and mass spectral trees for identifying unknown compounds in metabolomics," *Trends Analyt Chem*, 69:52-61, 2015.

thermoscientific



Compounding insights

Integrated, complete, toolset solves untargeted metabolomics analysis challenges

Thermo Scientific™ Orbitrap™ mass spectrometers produce information-rich metabolomics data. By integrating multiple tools to support unknown compound identification, Thermo Scientific™ Compound Discoverer™ software enables you to annotate more of your unknowns—without leaving the software application. With a comprehensive, integrated set of libraries, databases, and statistical analysis tools linked in customizable workflows, the software streamlines unknown identification, determination of real differences between samples, and elucidation of biological pathways.

Find out more at **thermofisher.com/compounddiscoverer**

ThermoFisher SCIENTIFIC