Environmental Metabolomics

Daniel W. Bearden

National Institute of Standards and Technology, Hollings Marine Laboratory, Charleston, SC, USA

1	Introduction	1
2	Background	1
3	Rationale	2
4	Techniques	3
5	Current Applications	7
6	The Future	8
7	References	8

1 INTRODUCTION

The application of metabolomics in the field of environmental science or ecology, which has developed based substantially on NMR spectroscopic approaches, is a fast-paced, rapidly developing field which seems to be poised to help reframe the discussion of environmental effects on organisms. Because of the nature of metabolomics, where experiments are based on ensembles of individuals, one is led to observations that are pertinent to population-, community-, and ecosystem-scale issues. This is in contrast to human-health-related metabolomics where one often wishes to observe or diagnose the condition of a single individual from a population. The environmental metabolomics literature is expanding and the field is maturing at a rapid pace, in part not only because of the advances in human-health metabolomics research but also because of the unique insight that this approach brings to an important area with global implications.

This article focuses on the study of environmental factors that impact the health and well-being of "non-model" organisms in the environment, in an effort to demonstrate that the application of NMR-based metabolomics can enhance traditional approaches to environmental science. In this view, non-model organisms represent a unique realm, distinct from the realm of direct human-health-related organisms, although these realms do interact in important ways. This realm includes some members that may serve as early warning sentinels for ecological issues, some which are commercially valuable for tourism, food, or sport, and some which deserve attention because there may be some species which, if negatively impacted by an unanticipated response to pollution, may cause widespread effects on the ecosystem structure, possibly affecting mankind.

One of the opportunities in environmental metabolomics is the number of relevant species about which little specific biochemical information is known. There are species that are relevant based on geography (diatoms in the Antarctic, or plants and animals near point pollution sources, for example) and other species that are seen as the basis for a complex, interconnected food web which may be perturbed owing to environmental change or contamination. Some populations need to be studied because of the need to preserve diversity and conserve protected species. For many of these organisms, little is known at the genomic or proteomic level, and sometimes even basic characteristics such as diet, range, or reproductive patterns are poorly characterized.

Many anthropogenic contaminants have been well characterized in terms of temporospatial distribution and toxicological impacts on relevant species, including humans. While concerns about historical contaminants such as polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs), trace heavy metals, and pesticides have been addressed since the 1960s through the development of robust analytical techniques and systematic toxicological protocols, a rising awareness of less studied pollutants is raising questions that may not be as amenable to the established approaches. New materials for the production of consumer and industrial products or new drugs find their way into the environment through manufacturing processes, usage, and in the waste stream, where, for example, wastewater treatment has not been designed to properly treat the waste stream for the new contaminants.

The need to develop assessments of sublethal stressors such as those related to climate change, personal pharmaceutical products in the waste stream, new generations of pesticides, and new consumer-related chemicals entering the home and workplace continues to grow.[1] Some of these new chemical stressors may act as endocrine mimics, causing subtle effects in the reproductive biology of organisms. Others may have very species-specific interactions that are undetected in the established regulatory processes, which typically have a limited suite of biological tests, resulting in impacts on non-target organisms that only become apparent once the chemicals have been in use and their distribution is widespread.

This onslaught of new chemical and physical stressors, and the awareness of the importance of environmental services linked to non-model organisms, can overwhelm traditional approaches to environmental research. Environmental metabolomics provides new tools to link environmental stressors to specific biological responses, in a discovery mode where the biochemistry of the organisms can be illuminated and also in a quantitative, hypothesis-driven mode where specific questions can be addressed.

2 BACKGROUND

While the role of NMR in general environmental research has been growing, techniques such as chromatography-based light spectroscopy or mass spectrometry continue to be the analytical workhorses in the area of chemical environmental research, where most pollutants are measured at trace levels. NMR spectroscopy has been shown to be useful in a number of important areas of environmental research.[2,3] For example, NMR is a primary technique in the purity assessment of compounds and is key in the identification and quantification of compounds.[4] In the general chemical sciences where, for example, new synthetic products are created, NMR has had a significant role in structure determination, including stereochemical assignments.

Since the explicit proposal that "... the thorough quantitative analysis of body fluids might permit differential diagnosis of many diseases in a more effective way than is possible at the present time" by Linus Pauling *et al*. in 1971,[5] the ability

Encyclopedia of Magnetic Resonance, Online © 2007–2012 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Magnetic Resonance* in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470034590.emrstm1256

to quantitatively assess the complement of small molecule, endogenous metabolites in living organisms has shown practical results for human health through disease research, dietary studies, and numerous other health-related endeavors.

The concept was clarified and expanded beyond just disease diagnosis in 1999 in a seminal work.[6] A systems approach, where the overall variation of metabolite concentrations is considered comprehensively, is fundamentally different from the approach where a few specific metabolites are individually assessed. While the map of the important metabolic pathways was painstakingly developed through skillful classical chemical experimentation on a reaction-by-reaction basis, the concept behind the field of metabolomics is the simultaneous direct observation of as many endogenous metabolites as possible in a "snapshot" of the instantaneous physiological condition of an organism. In the last two or three decades, these ideas have been expanded, especially in relation to human health research, where the utility of NMR-based measurements has been shown to address effectively the concepts of metabolome assessment.

Given the successes in human-health metabolomics research, it is only natural to apply these approaches to environmentally relevant, non-model organisms.[7–13]

3 RATIONALE

NMR is an excellent tool for the assessment of the complex constitution of biomaterials because it is an unbiased detector, absent some well-documented systematic pitfalls, of the organic compounds in multicomponent samples. The NMR signal is a superposition of the spectra of all the components in the sample, although there can be some nonadditive effects because of physical/chemical interactions between compounds which complicate spectral analysis. Samples can be prepared in relatively simple ways, especially for biofluids such as urine, plasma, serum, or cerebrospinal fluid (CSF), often with no need for any chromatographic separation or cleanup, avoiding the quantitative complications associated with chromatography and/or chemical derivatization. The signals of individual chemical constituents are intrinsically proportional to concentration, simplifying quantitative analysis. Most metabolites have spectra that exhibit multiple resonances which allow identification of the compound from simple ^{1}H spectra, and when correlation spectra such as COSY, TOCSY, or heteronuclear single-quantum correlation spectroscopy (HSQC) are used, the identification of compounds becomes even more specific.

Metabolomics has been shown to be very sensitive to external effects on organisms. For chemical exposures, effects of exposure can often be detected at environmentally relevant concentrations, avoiding the difficulties involved in extrapolating from a high-exposure experiment to much lower levels of exposure.[14] However, this sensitivity is a two-edged sword in that experiments must be designed as carefully as possible to eliminate erroneous observations. In a classic example of good practice gone bad,[15] a laboratory exposure involving rats was found to be problematic because of a change in feed between the supplier and the pharmaceutical research laboratory; the rats had not been equilibrated sufficiently on the new diet before the experiments were run. The implications for environmental research are severe, since often relatively uncharacterized organisms are used in such studies. For example, one classic approach to obtain a working population for study involves field collection of organisms and equilibration in the laboratory. However, Hines *et al.*[16] contrasted laboratory-equilibrated mussels with field-frozen mussels and found significant differences in the metabolite profiles. Their recommendation was to only use field-sampled tissues and fluids to avoid increasing the metabolic variability that may mask the effect being studied. In situations that seem very amenable to laboratory studies, for example studies with microorganisms such as bacteria,[17] small changes in sample history can cause apparent metabolomic shifts that may confuse the interpretation of results. The sensitivity of metabolomics to phenotypic variation must be appreciated and controlled as the work in this field advances.

One great advantage of using metabolomics for environmental research is the ability to distinguish different modes of action[18] due to different toxicants in sentinel organisms. Potentially, organisms from the field can be assessed for metabolic fingerprints of the different modes of action of various physical and chemical stressors, so that an effective assessment of community health can be made. These metabolic responses will be time- and dose-dependent, so that in well-modeled systems, a complete dynamic picture of ecosystem health can be developed.

While many of the organisms of interest in environmental science have not been well characterized proteomically or genetically, it is still possible to understand the stress responses from a metabolomics viewpoint. In fact, knowing the idiosyncrasies of the metabolic response may point to areas where genetic/proteomic studies should be pressed. Discovery of disproportionate metabolite signals or hitherto ignored compounds[19] may indicate novel genetic mechanisms which need investigation.

NMR has a dynamic range of several orders of magnitude, which can be increased through longer data acquisition or other approaches that increase the signal-to-noise ratio (SNR). The absolute sensitivity of NMR can be easily exceeded by the use of techniques such as fluorescence or mass spectrometry, but the trade-off in selectivity for these other techniques, coupled with the requisite need for some sort of chromatography, is often compensated by the broad-based, nondestructive, nonselective detection afforded by NMR. In terms of the natures of the compounds detected, the use of NMR affords the widest range of detection of chemical moieties in a single analysis. Sugars, organic acids, lipids, amino acids, and so on are easily detected and quantified in a single sample in a single experiment. While it is important to identify as many metabolites as possible in a sample, the mere nonquantitative detection of a metabolite using mass spectrometry, for example, does not necessarily give insight into the metabolic response to a stressor, especially for subacute responses. Given the extremely complex and correlated nature of the metabolome, one must carefully draw the line between extremely sensitive detection of every compound in a sample, and the need for quantitative or semiquantitative assessment of "important" metabolites that can help with the problem at hand. However, as the field matures, there is a building consensus that the use of multiple modalities of compound detection and quantitation provides significant advantages in understanding the metabolomic system response.

Encyclopedia of Magnetic Resonance, Online © 2007–2012 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Magnetic Resonance* in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470034590.emrstm1256

NMR fits comfortably in the continuum of measurement techniques because of the ability to obtain quantitative metabolite "patterns" while also providing quantitative chemical-specific information for a wide variety of organic compounds. These features give NMR-based metabolomics a role in the discovery of new metabolomic insight and in classical hypothesis-driven investigations that link organism biochemistry to environmental stressors.

4 TECHNIQUES

4.1 Experiment Design

Robust experiment design is key to meaningful metabolomics results. Good experimental design requires careful communication and the ability to work with people from other specializations, developing a dialog with a common vocabulary, perhaps even developing a formal ontology.[20] There is a trend in the literature showing that experiments are improving from an experimental design standpoint.[21] For example, in many published reports, the number of samples analyzed is large enough to develop meaningful statistical inferences, and the need for repeating trials is being recognized. Since many of the practitioners have ties to the environmental research community, the need for standard practices and quality control (QC) is recognized as important in improving the confidence in the reported results. Ideas for robust analytical measurements can be borrowed from the environmental analytical community, such as the use of certified reference materials, project-specific control materials, measures of analytical repeatability,[22] and interlaboratory comparison exercises.[23] In terms of the biological component of the experimental design, husbandry of the organisms must be considered in terms of effects on subsequent NMR experiments. In handling the organisms for sample collection, stress induction must be minimized and rapid quenching of metabolic processes should be of paramount importance, especially in tissues that are metabolically active such as the liver. The effects of feeding, infection (both bacterial and viral), species misidentification, or silent phenotypes must all be considered in the biological design of the experiment. Because of the trueness and precision of NMR experiments, the repeatability of sample preparation and the robustness of the statistical tools used for data analysis, most practitioners end up confronting biological variability as the most challenging aspect of environmental metabolomics. Time spent developing a well-designed biological study will return rewards in high-quality, repeatable results with a significant impact in the field.

4.2 Sample Extraction and Cleanup

The broad appeal of NMR-based metabolomics is that one is able to garner meaningful metabolomics results without the additional complications of chromatographic fractionation. However, sample extraction or cleanup is a critically important factor, because there is no recognized method to isolate quantitatively all the organic metabolites from biological samples; each extraction or cleanup protocol introduces some bias in the quantitative extraction of metabolites. This has not proven to be a major stumbling block, because it is possible to make valid inferences based on well-extracted samples. Perhaps this robustness is due to the fact that the measurements sample a network of metabolites, and as long as the extractions are analytically consistent and reasonably robust, the network responses can be detected in a meaningful manner.

Because of the various organisms and matrices considered in environmental metabolomics, samples are processed in ways that are tissue-dependent, and in most cases, different species require variations in extraction protocols. The most convenient matrix to work with would be the one with the least bench workup required, such as a body fluid. However, the different constituents in fluids such as plasma or serum warrant some effort in validation of the sample workup.[24,25] For tissues, the extraction process can cause a bias if the extraction efficiency is not examined systematically.[26] Depending on the matrix, some extraction schemes are better than others, much as in environmental analytical chemistry. Since no real chromatography is performed, one may view this as a "cleanup" process whereby, for example, large molecular weight molecules such as proteins, DNA/RNA, polysaccharides, and other macromolecules are removed, leaving behind the small molecule (<500 Da) metabolites.

Various extraction schemes have been systematically optimized for some environmentally relevant matrices.[26,27] The optimizations proposed in these systematic studies are selected on the basis of the observed repeatability of the method and some measure of the amount or number of metabolites extracted. A common extraction technique is based on resolving polar and nonpolar metabolites into separate fractions, for example using a modified Bligh and Dyer scheme.[26,28] This is often desirable since some experiments may be rationalized based on more polar components (such as amino acids, TCA cycle metabolites, organic acids, and aromatic compounds) in a polar solvent such as water-based buffer, while some compounds associated with lipids or cholesterol synthetic pathways would be in a nonpolar solvent such as chloroform. For high lipid-content tissues, the spectral simplification after polar/nonpolar resolution is significant, leading to clearer interpretation of the results. Some schemes have been optimized to use size exclusion cleanup based on size exclusion filter technology, either by itself or on various fractions previously separated using other techniques.[29] For blood or hemolymph studies, numerous efforts are reported for optimization of metabolite extractions in plasma and serum, often based on filtering techniques.

Well-executed projects invariably have invested suitable effort in validating the extraction/workup procedures to optimize the metabolite fingerprints and sensitivity of the experiments.

4.3 NMR Data Acquisition

Metabolomics analysis depends greatly on the type and quality of NMR data collected. Most environmental metabolomics projects depend on one-dimensional (1-D) data collection of polar solutes in D_2O. The best spectra have very flat baselines (so that subsequent baseline corrections are easy), good phase characteristics, reasonable resolution, minimal spectral artifacts, and very good SNR. In addition,

Encyclopedia of Magnetic Resonance, Online © 2007–2012 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Magnetic Resonance* in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470034590.emrstm1256

spectrum-to-spectrum consistency is very important because most of the numerical pattern detection techniques will select for features that vary between spectra.

It is good practice to develop a consistent protocol for data acquisition that provides consistently good results. For example, it is more desirable to set up the instrument carefully and then run all the samples in a project in one "session" than to run samples in multiple sessions over a few days. For very large projects, it may be impractical to run all the samples in one session, so the protocols should involve measures that enforce and verify consistency between sessions, such as repeat runs of select samples or measurements of line widths or SNR. Protocols could cover factors such as temperature stability, temperature measurement,[30] shimming protocols, standard parameters for the pulse sequences, pulse width calibration, and standardized processing parameters. While some projects require deviations from standard protocols, having a consistent starting place for making those decisions is good practice.

For water-based samples, most laboratories use water suppression pulse sequences. Depending on the sample preparation protocols, samples may be in 90% H_2O or neat D_2O or somewhere in between, so the exact water suppression technique must be optimized for that class of sample and for the particular instrument being used. Because of the spectral artifacts that can be introduced and the need for high-quality semiquantitative spectra, the optimization of the water suppression technique is critical. Various suppression schemes have been optimized for water suppression in systematic studies,[31–33] and sequences based on a three-pulse NOESY-type sequence[32] are often used on samples prepared in D_2O. However, more rigorous water suppression techniques are not uncommon. Optimization of suppression often considers baseline distortions, intensity perturbations near the water resonance at the theoretical lobes of the suppression sequence, and difficulty of calibration and setup.[34,35] Trade-offs between these factors often come into play, and local optimization of water suppression is crucial to meaningful, consistent results.

For samples that contain residual proteins or high molecular weight lipids, such as plasma or serum, the use of spin-echo Carr–Purcell–Meiboom–Gill sequences can act as T_2-weighting filters that reduce the contribution of broad-line signals from high molecular weight species. The optimization of these sequences balances the duration of the effective spin-echo delay against the phase distortion due to homonuclear couplings and the loss of intensity due to pulse imperfections and relaxation effects. Reports of cumulative spin-echo delays in the range of 100 ms have given satisfactory results. Combining the spin-echo sequence with strong water suppression in high H_2O content samples can also be challenging,[36–38] and must be weighed against the improvement gained by processing the samples to remove high molecular weight components.

Some reports show that 2-D spectroscopy can lead to superior results in pattern recognition and compound identification. One 2-D experiment that seems useful is 2-D J-resolved spectroscopy (2D-JRES).[39] The tilt-corrected data can be used for identification of compounds, and a skyline projection along the direct dimension results in a homonuclear decoupled spectrum that significantly reduces the spectral complexity by collapsing the homonuclear multiplets. This projection can then be used in pattern recognition approaches in the very same manner as direct 1-D spectra. Other 2-D homonuclear experiments such as COSY and TOCSY can also be useful, especially for compound identification. For more reliable compound identification, heteronuclear experiments such as ^{13}C-HSQC and ^{13}C-HMBC provide nearly unambiguous compound identification in natural abundance samples, although at a somewhat higher cost than the 1-D experiments because of the need for longer acquisition.

4.4 Data Processing and Analysis

There are many algorithmic approaches to discerning the systematic variation in the spectra from a metabolomics experiment, including many of the tools developed for fields such as functional genomics. Once patterns are detected, it is important to carefully evaluate whether the patterns are statistically significant or a result of systematic error. In some experiment designs, the number of samples is too small in an experimental group, and a determination must be made on whether the data is representative of the populations or whether outliers have a significant influence. This assessment is no different from metabolomics in other fields. Interpretation of the results depends on the robustness of the experimental design, where phenotypic variability of the sample pool was assessed, for example, and should be evaluated in the light of the results on QC samples, which help quantify the variation due to sample preparation or NMR spectral quality.

Once there is confidence that a true pattern exists, the compounds that contribute to the separation in the pattern must be identified so that linkages to metabolic pathways can be established. Therefore, pattern recognition techniques should be chosen based on their ability to provide both pattern detection and chemically relevant compound identification. In the simplest of cases, a simple univariate approach based on comparison of group-averaged spectra, for example, may lead to an understanding of the biochemical basis for differences between the treatment groups.[40–42] In these cases, there are probably very small numbers of compounds that are significantly different between treatment groups. However, metabolic response to a stressor may be more subtle, and spread over a wide range of metabolic compounds or pathways, since the whole metabolome responds to the stressors in the experiment. In this case, multivariate analysis techniques that are sensitive to coherent variation in numerous chemical signals simultaneously are most informative for detecting this coherent variation. Because of the nature of NMR spectra, where an individual compound often has numerous peaks, there is considerable correlation in the data, so that each spectral point or bin does not necessarily represent a fully independent variable. This results in a reduction of rank, and numerical methods that are robust to this nonindependence should be the most trustworthy.

The preprocessing of the data also affects the pattern recognition process. For example, baseline correction is necessary because of imperfections in instrumentation, including background signals from the probe or preamplifiers.[43,44] Sometimes, there may be effects from the receiver/digitizer system that cause rolling baselines. Baselines are corrected with any of a myriad of techniques, ranging from simple polynomial subtractions to more sophisticated algorithms. It is best to

Encyclopedia of Magnetic Resonance, Online © 2007–2012 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Magnetic Resonance* in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470034590.emrstm1256

avoid severe baseline corrections by having a well-designed and maintained NMR instrument that produces flat baselines on good test samples. Good laboratory practice involves careful setup, testing and execution of experiments, in a way that detects spectral quality issues as soon as possible.

Most tools for data analysis organize the individual spectra into the rows of a matrix in which the columns then represent the chemical shift of the spectra. The number of columns of the matrix may be reduced by "binning" the spectra in a systematic manner as part of the data pretreatment.[45,46] The simplest form of binning involves dividing the spectrum into a fixed number of fixed width bins, summing the individual points that fall into a bin. The selection of the appropriate bin width is often dependent on the experience of the analyst, but using a bin width that is too large means the selectivity of the analysis is reduced, since multiple compounds may contribute to a bin, while selecting a bin width that is too small means the results may be overly sensitive to spectral features, such as the line width, so that the effect of shimming, for example, gets exaggerated. Other considerations include the possibility of inappropriate alignment of bin edges on spectral features, separation of spectral multiplets into individual bins, or peaks that shift slightly from sample to sample due to pH effects or ionic strength issues causing the spectral feature to jump between bins in different rows of the matrix. These considerations led to the development of more sophisticated binning algorithms and to the development of more robust spectral alignment tools.[47] This is an active area of development across all of NMR-based metabolomics.[48]

Parts of the NMR spectrum, such as the water region, can be excluded a priori from the analysis, and this is often done by simply deleting the data columns associated with certain chemical shift regions. For example, in a chemical dosing experiment, the toxicant may show up in the samples, and including this in the subsequent analysis might be inappropriate. Another example would be that residual solvent or inadvertent contamination from the sample extraction process remains in the sample and would contribute to the variance in the data set.

Since the effect of spectral noise on some pattern recognition techniques is not well determined, it is best to collect data in a way that keeps the SNR consistent throughout data acquisition. In cases where sample concentration cannot be controlled, due to the sampling techniques for example, experiments may need to be run with an appropriate adjustment of the number of spectral scans. Most often, spectra are normalized so that the total spectral area is constant, but there may be a reason to use a single metabolite or spectral region for data normalization because of the nature of the samples. For example, creatinine is often used in urine-based experiments because of the historical clinical practice of normalization to the creatinine level.[49]

In some experiments, there is a relatively small number of spectral peaks that dominate the spectrum, and unless those compounds are the particular ones of interest, these peaks can be scaled so that the variance of less intense peaks is detected. In extreme cases, the samples may need to be treated differently during sample preparation in order to mitigate the intense signals that may actually obscure smaller signals that convey the important information.[21] Alternatively, bins may be normalized to the variance of the data in each column, and this gives equal importance to each column variable. Other schemes, such as Pareto weighting, where each column is normalized by the square root of the variance, can reduce the influence of large peaks while keeping scale information for the other spectral regions with less overall variation. Some more complex transforms are also possible, such as log transformations.[50] Since metabolomics is evolving rapidly, numerous variance stabilizing transforms are being proposed and tested.

4.5 Principal Components Analysis

The workhorse of multivariate analysis in NMR-based metabolomics is principal components analysis (PCA), where the preprocessed data matrix is resolved into a "scores" matrix, which represents each sample spectrum as a point in a high-dimensional space, and an accompanying loadings matrix, which describes the optimal axes for this new space in terms of the spectral bins. These new axes are determined based on the criteria of maximizing the explained variance (EV) along each orthogonal axis. The scores are sorted by decreasing eigenvalue, since the smaller eigenvalues correspond to less explanatory power, and the overall dimensionality is reduced by considering the first few components corresponding to the largest eigenvalues. There are numerous ways to decide how many principal components (PCs) are sufficient to model the data, but seldom are more than two or three considered. Systematic investigation of higher PCs, however, is good practice. The decrease in EV for successive PCs is an indicator of the quality of the model, and often there is a significant explanatory power in the first few PCs. A very gradual increase in the cumulative EV with PC number may be an indication of a less definitive model. Auxiliary information from the PCA analysis, for example, plots of Hotelling's T^2, can be used to identify potential outliers in the data set.

Practically, the scores plots are examined for grouping or trends according to treatment group. If there is apparent grouping in scores plots, univariate testing can be done on the score values to determine the significance of the separation, even if there is significant scatter in the individual treatment groups.

Loadings plots contain information about which bins contribute to the EV for the corresponding PCs. Loadings plots are also often plotted as two-dimensional plots corresponding to the PCs in the scores plot. A loadings plot shows which compounds are correlated or anti-correlated to separations in the corresponding scores plots. In a 2-D loadings plot, there is one data point for each variable or bin in the data matrix.

Although PCA scores and loadings are a powerful, unbiased way to examine the data, the interpretation of PCA scores and loadings plots is somewhat difficult because there is no constraint on the algorithm to present a linear combination of pure NMR spectra. Sometimes, there is a strong effect in the data set which is tied to a few compounds and these spectral features dominate the loadings plots, making it straightforward to identify these important compounds and progress to a biological interpretation. Unfortunately, it is difficult for people trained in the thought processes of "single response, single variable" to conceptualize a system-wide response vector. There is a natural tendency to revert to univariate thinking in discussing results, and seldom are the data treated as a multicomponent, coherent effect. If there are a few strong signals in a loadings plot, libraries of spectra or peak data

Encyclopedia of Magnetic Resonance, Online © 2007–2012 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Magnetic Resonance* in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470034590.emrstm1256

tables can be used to identify the relevant compounds. In some cases, there is no small number of intense peaks, so one is faced with a much more difficult interpretation of the loadings vectors. The interpretation of loadings vectors is often more difficult, as well, when higher PCs are examined. Also, translating a coherent change in metabolite levels to a metabolic pathway interpretation is difficult, especially given the dearth of specific knowledge of metabolic pathways in non-model organisms. Metabolomics results presented in a "network" topology, often correlated with established metabolic pathways, can be useful for conveying the multivariate response that is observed in metabolomics experiments.[51] The classification capability of the analysis is often assessed using receiver operating characteristic (ROC) curves, and the associated area under the curve (AUC) parameter.[48,52]

4.6 PLS and PLS-DA

Often, experiments are designed to have an independent variable or classification of the treatments as the basis for determining an effect. This information is most often brought into the multivariate analysis through techniques based on partial least squares (PLS) projections, especially when the class separation is not as apparent in a PCA analysis. The classification information is incorporated into the analysis through a vector or matrix relating the experimental variable for each sample to the sample spectrum. If the experimental variable is not a continuous variable such as temperature, length, or pH, for example, but a discrete parameter such as male/female, this "Y" matrix is constructed as a discriminant matrix where the class is assigned a numerical value, leading to PLS with discriminant analysis (PLS-DA). For cases where there are multiple discriminators, the Y matrix is constructed with one column for each discriminant value and a numerical value, such as 1 or 0, is used to denote class membership. The difficulty with this approach is that the algorithm can blindly find correlations in the variables that satisfy the constraints, even if those variables are really just incidentally correlated noise. Therefore, one must very carefully test the results of PLS analysis for accidental correlations and bias, and numerous robust techniques have been developed to assess the "trueness" of detected correlations.[53–56]

PLS analysis also leads to scores and loading plots that can be used to tie the systematic variation between treatment groups to the specific chemical variation that distinguishes the groups. These identified compounds can then be linked to metabolic pathways, indicating the systematic response to the treatment variables.

4.7 Other Pattern Recognition Techniques

Numerous other pattern recognition tools can be and have been used for metabolomics studies. These range from the previously mentioned significant difference spectra (SDS) analysis[42] to artificial neural networks (ANN)[18] and support vector machines (SVM).[57] Each of these techniques has particular strengths, but none has found as widespread applicability in environmental metabolomics as PCA and PLS techniques.

4.8 Spectral Libraries

Compound identification is key to tying the spectral information to biochemical pathways and biological interpretation. In most experiments, reports indicate the assignment of approximately 50 and up to 100 compounds from NMR data. By carefully matching sample spectra to spectra of pure compounds in libraries, collected under similar sample and experimental conditions, confidence in identification and quantitation of the peaks in the mixture spectra grows. Many spectral libraries are freely publicly available[58–61] and some are commercially available. Matching can be accomplished through manual peak enumeration and comparison of chemical shift tables, or through interactive library searches or interactive peak alignment. Care must be exercised in matching because there are several chemical shift standards in use, and peaks may not match well if the chemical shift standard is not unambiguously identified.[62] For the non-model organisms of interest in environmental metabolomics, one has the potential problem that peaks from existing databases may not include metabolites of importance,[19,63–65] even though most of the libraries contain several hundred compounds. Most of the libraries are focused on more polar compounds; however, it is often desirable to consider the metabolomics of nonpolar compounds, and these libraries are less developed at this time. Under optimum conditions, the quantification of metabolites can be accomplished based on libraries, and these quantified metabolites can then be used for subsequent data analysis, rather than binned spectra. This process has been named "targeted profiling".[32] As always, ambiguities can always be resolved using analytical techniques such as authentic compound standard additions or chromatography-based purification and structure elucidation.

4.9 Quality Control

Analysis of individual spectra for quality parameters such as lineshape, baseline distortions, and instrumental artifacts (quadrature images, spurious radiofrequency signals) is essential for generating quality metabolomics results. Statistical analysis of QC samples that were processed with the experimental samples[22] can be used to classify the reliability of the overall experiment, although this is rather rare in the published literature at this time.

Intercomparison exercises, where participants analyze identical samples according to a specific protocol and the results are compared for consistency, have shown that at least the technical analysis of samples using NMR spectroscopy can have a high level of consistency across laboratories.[23] Even with different magnetic field strengths, instruments from different vendors, and analysis with different software packages, substantial agreement is feasible in NMR-based metabolomics. This contrasts sharply with mass-spectrometry-based metabolomics and with efforts in other -omics fields to show analytical consistency.[66–69] Because of the complexity of the biological models developed and the need for larger studies that may involve instrumental analysis across different laboratories, the ability to consistently, quantitatively analyze metabolomics samples with a high degree of interlaboratory reproducibility is crucial.

Encyclopedia of Magnetic Resonance, Online © 2007–2012 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Magnetic Resonance* in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470034590.emrstm1256

4.10 Data and Reporting Standards

The advancement of the field of environmental metabolomics depends on laboratory exchanges of data and consistent descriptions of data treatment and interpretation. Efforts to standardize the mechanism of data exchange are ongoing and can piggyback on efforts such as the metabolomics standards initiative (MSI).[70] For NMR data, the actual exchange and storage of raw spectral data is straightforward; however, the accompanying metadata, which describes the experimental design, the sample collection and handling, NMR data collection and processing, and subsequent multivariate analysis, is still evolving.[71,72] Once practical data standards are established and put into widespread use, data can be placed into repositories in meaningful ways. Data repositories will prove useful for future analysis with new algorithms, for long-term environmental studies, and for development of species-specific "stressor libraries" compiled from numerous independent research efforts.

5 CURRENT APPLICATIONS

As mentioned in the introduction, the application of NMR-based metabolomics can enhance traditional approaches to environmental science, and can address environmental factors that impact the health and well-being of "non-model" organisms in the environment. These organisms are important as functioning members of the ecosystem, forming the basis of the food web, providing important ecological services and providing us with sustenance, besides having important societal and cultural value.

5.1 Laboratory Exposures/Treatments

An essential element of environmental toxicology is the laboratory-based experiment. In these experiments, an organism is maintained in an artificial environment where conditions such as temperature, water conditions, or nutrition are under control. In well-designed experiments with appropriate control organisms, the response to chemical toxicants or physical stressors is measured in a way that should allow extrapolation to a "real-world" exposure or stress. However, as useful and essential as these experiments are, the laboratory environment often does not mimic every factor that may be found in the field.

The problem of linking field-collected samples to laboratory studies has not been addressed in general. The equilibration of field-collected organisms to the laboratory may cause a bias in the results and predictions. For example, the organism selection process (capture, transport, shock survival, etc.) can lead to a bias based on organism survival, biased phenotype selection, or limited gene pools. In microorganism culture experiments, only a subset of the population may be cultivable in the laboratory, so that only a small part of the representative organisms can survive to the laboratory environment. For organisms with gut or symbiotic microorganisms, change from a wild environment to laboratory environment may cause alterations in the microflora, impacting the metabolome in important ways. In addition, other factors such as adaption to consistent feeding, lack of predation, lack of temperature or physical variability, or lack of multispecies signaling may lead to confounding factors which impact the applicability of laboratory-based assessments to field observations. Factors such as full-spectrum sunlight, diurnal cycles, tidal cycles, predation, and competition for food are difficult to replicate in a consistent manner. These effects are often observed in metabolomics experiments, while for other measurement modalities, these effects may not be considered important in interpreting the experimental outcome, perhaps hiding major contributing factors to the experiment.

Naturally, laboratory-based metabolomics measurements do have some advantages. Laboratory exposures make diet, temperature, and other environmental factors controllable so that experiments can be done with reasonable sample sizes, keeping the logistics manageable. Also, single captive organisms may be followed over time as the individual responds to treatment. Often the protocols, while not "perfect," are well defined, and therefore replicable to a great degree in other laboratories. A consistent protocol allows at least for a systematic framework for comparing the toxicity of widely disparate chemical exposures and stressors.

5.2 Field Collections

To connect laboratory data to true environmental problems or monitoring, one has to move toward analysis of field-collected samples. Ecological aspects are probably best answered by field collections. However, in terms of interpreting environmental metabolomics results, one must consider the problem of uncontrolled variables such as diet, temperature, predators and factors such as pollution; these may have to be explored through careful laboratory exposures. This also implies that single individual sampling for a "quick" environmental assessment will be problematic. Again, just because metabolomics is influenced by the effects of uncontrolled variables does not mean that the technique is flawed; it probably means that more of the factors influencing the results can come into play, resulting in a more robust population-level analysis.

5.3 Case Studies

Many environmental problems are being addressed through "case studies" where an organism is subjected to a relevant stressor and the metabolomic response is rationalized. These types of studies are important because they are the building blocks that can be used to design more comprehensive studies, better understand the biology of the non-model organisms, and develop expertise in understanding multistressor, multiorganism ecological models based on the biochemical response to stressors. Two groups of organisms, earthworms and bivalves, have been the objects of numerous studies and the expanding body of knowledge may prove to be very valuable.

As an integrated assessor of environmental processes in soil, worms are garnering a lot of attention.[21,73–88] Studies based on several species of field-collected and laboratory-dosed worms indicate that the organism has a robust metabolic response to soil contamination, including organic compounds and heavy metals. These studies show a splendid progression of interest and report on a number of sample preparation schemes, exposure routes, and elucidation of different modes

Encyclopedia of Magnetic Resonance, Online © 2007–2012 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Magnetic Resonance* in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470034590.emrstm1256

of action. As an Organization for Economic Cooperation and Development (OECD) recommended test species,[89] *Eisenia fetida* is the focus of a majority of the studies, although other ecologically important species are being studied.

Mussels and clams have also been investigated using metabolomics.[7,16,90–100] Given the stationary nature of these mollusks and their aquatic environment, they may prove to be an important monitoring organism as an early warning sentinel for incipient pollution issues. Studies involving organic and inorganic pollutants have shown that bivalves are metabolically sensitive and different modes of action are apparent in their metabolic fingerprints. In one study,[92] the sex of mussels was determined using NMR-based metabolomics, and, while not as accurate as reverse transcriptase polymerase chain reaction (RT-PCR) for sex determination, metabolomics was a better indicator of functional reproductive status in both ripe and spent mussels.

Numerous case studies, which may be part of longer term investigations, have been reported since two reviews of the environmental metabolomics field were published.[10,101] Research involved NMR-based metabolomics studies of coral-associated bacteria,[17] Atlantic blue crabs,[102] and fish.[14,103–108] Environmental stressors varied from hypoxia, to microbial challenge, to temperature, to oil exposure, to heavy metals contamination.

5.4 Comprehensive Approaches

A recent report demonstrated the potential of environmental metabolomics to address the full range of linkages in environmental assessment from ecosystem-scale measurements to specific modes of action from environmental stressors.[95] The study illustrated the linkage of metabolomic biomarkers to an accepted assessment of organismal health based on the scope for growth (SFG), a well-defined biological index of the fitness of an organism for growth, reproduction, and survival. In a review of this work,[94] Robertson stated "The elegance in the work... is that they not only generated the models in the laboratory environment but they further field tested them in a real-world application" and indicated that the field of environmental metabolomics had reached another level of expectation and performance. Not all current studies are as comprehensive at this point, but the mark has been set and the potential is tremendous.

6 THE FUTURE

The opportunities for impacting the field of environmental research seem to be growing, based on the increasing number of publications and increasing scope of study. Perhaps the future holds exciting biological discoveries as more non-model organisms come under the "NMR-metaboscope" and more specific metabolic pathway maps are refined. The possibility of linking established biological indexes to metabolomic information means that there may be effective ways to assess environmental impacts and set public policy based on specific biochemical interactions, leading to better science-based management decisions.

Improving the tools of NMR-based metabolomics means that more consistent analysis protocols and metabolome mining techniques will appear, either based on specific needs of the analysis of non-model organisms or as an adaptation of approaches in human-health metabolomics.[109,110] Consistent ways of reporting, archiving, and sharing data are emerging which will allow groups to confidently leverage existing data and analysis.

As progress continues, there will be examples of long-term, regional monitoring which may help provide an early warning of encroaching environmental issues. NMR-based environmental metabolomics can develop systematic descriptions of mode of action responses to specific stressors in sensitive organisms, leading to new biological insights and better toxicological understanding of the complex, multispecies, multistressor environment in which we and our fellow creatures live.

7 REFERENCES

1. S. D. Richardson and T. A. Ternes, *Anal. Chem.*, 2011, **83**, 4614–4648.
2. A. J. Simpson, D. J. McNally, and M. J. Simpson, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2011, **58**, 97–175.
3. L. A. Cardoza, A. K. Korir, W. H. Otto, C. J. Wurrey, and C. K. Larive, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2004, **45**, 209–238.
4. F. Malz and H. Jancke, *J. Pharm. Biomed. Anal.*, 2005, **38**, 813–823.
5. L. Pauling, A. B. Robinson, R. Teranishi, and P. Cary, *Proc. Natl. Acad. Sci. U.S.A.*, 1971, **68**, 2374–2376.
6. J. K. Nicholson, J. C. Lindon, and E. Holmes, *Xenobiotica*, 1999, **29**, 1181–1189.
7. M. R. Viant, J. H. Walton, P. L. TenBrook, and R. S. Tjeerdema, *Aquat. Toxicol.*, 2002, **57**, 139–151.
8. M. R. Viant, E. S. Rosenblum, and R. S. Tieerdema, *Environ. Sci. Technol.*, 2003, **37**, 4982–4989.
9. M. R. Viant, *Mar. Ecol.: Prog. Ser.*, 2007, **332**, 301–306.
10. J. Bundy, M. Davey, and M. Viant, *Metabolomics*, 2009, **5**, 3–21.
11. M. Viant, *Metabolomics*, 2009, **5**, 1–2.
12. C. Y. Lin, M. R. Viant, and R. S. Tjeerdema, *J. Pestic. Sci.*, 2006, **31**, 245–251.
13. M. Macel, N. M. Van Dam, and J. J. Keurentjes, *Mol. Ecol. Resour.*, 2010, **10**, 583–593.
14. A. D. Southam, A. Lange, A. Hines, E. M. Hill, Y. Katsu, T. Iguchi, C. R. Tyler, and M. R. Viant, *Environ. Sci. Technol.*, 2011, **45**, 3759–3767.
15. D. G. Robertson, *Chem. Res. Toxicol.*, 2008, **21**, 1917–1922.
16. A. Hines, G. S. Oladiran, J. P. Bignell, G. D. Stentiford, and M. R. Viant, *Environ. Sci. Technol.*, 2007, **41**, 3375–3381.
17. A. F. Boroujerdi, M. I. Vizcaino, A. Meyers, E. C. Pollock, S. L. Huynh, T. B. Schock, P. J. Morris, and D. W. Bearden, *Environ. Sci. Technol.*, 2009, **43**, 7658–7664.
18. N. Aranibar, B. K. Singh, G. W. Stockton, and K. H. Ott, *Biochem. Biophys. Res. Commun.*, 2001, **286**, 150–155.
19. J. G. Bundy, E. M. Lenz, N. J. Bailey, C. L. Gavaghan, C. Svendsen, D. Spurgeon, P. K. Hankard, D. Osborn, J. M. Weeks, S. A. Trauger, P. Speir, I. Sanders, J. C. Lindon, J. K. Nicholson, and H. Tang, *Environ. Toxicol. Chem.*, 2002, **21**, 1966–1972.
20. S.-A. Sansone, D. Schober, H. Atherton, O. Fiehn, H. Jenkins, P. Rocca-Serra, D. Rubtsov, I. Spasic, L. Soldatova, C. Taylor, A. Tseng, and M. Viant, Ontology Working Group M, *Metabolomics* 2007, **3**, 249–256.
21. Q. Guo, J. Sidhu, T. Ebbels, F. Rana, D. Spurgeon, C. Svendsen, S. Stürzenbaum, P. Kille, A. Morgan, and J. Bundy, *Metabolomics*, 2009, **5**, 72–83.

Encyclopedia of Magnetic Resonance, Online © 2007–2012 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Magnetic Resonance* in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470034590.emrstm1256

22. H. M. Parsons, D. R. Ekman, T. W. Collette, and M. R. Viant, *Analyst*, 2009, **134**, 478–485.
23. M. R. Viant, D. W. Bearden, J. G. Bundy, I. W. Burton, T. W. Collette, D. R. Ekman, V. Ezernieks, T. K. Karakach, C. Y. Lin, S. Rochfort, J. Sd. Ropp, Q. Teng, R. S. Tjeerdema, J. A. Walter, and H. Wu, *Environ. Sci. Technol.*, 2009, **43**, 219–225.
24. S. Tiziani, A. H. Emwas, A. Lodi, C. Ludwig, C. M. Bunce, M. R. Viant, and U. L. Gunther, *Anal. Biochem.*, 2008, **377**, 16–23.
25. P. Bernini, I. Bertini, C. Luchinat, P. Nincheri, S. Staderini, and P. Turano, *J. Biomol. NMR*, 2011, **49**, 231–243.
26. C. Y. Lin, H. Wu, R. S. Tjeerdema, and M. R. Viant, *Metabolomics*, 2007, **3**, 55–67.
27. S. A. E. Brown, A. J. Simpson, and M. J. Simpson, *Environ. Toxicol. Chem.*, 2008, **27**, 828–836.
28. E. G. Bligh and W. J. Dyer, *Can. J. Biochem. Physiol.*, 1959, **37**, 911–917.
29. S. Stolzenburg, M. B. Lauridsen, H. Toft, P. A. Zalloua, and D. Baunsgaard, *Metabolomics*, 2010, **7**, 270–277.
30. M. Findeisen, T. Brand, and S. Berger, *Magn. Reson. Chem.*, 2007, **45**, 175–178.
31. E. Saude, C. Slupsky, and B. Sykes, *Metabolomics*, 2006, **2**, 113–123.
32. A. M. Weljie, J. Newton, P. Mercier, E. Carlson, and C. M. Slupsky, *Anal. Chem.*, 2006, **78**, 4430–4442.
33. A. J. Simpson and S. A. Brown, *J. Magn. Reson.*, 2005, **175**, 340–346.
34. N. Aranibar, K. H. Ott, V. Roongta, and L. Mueller, *Anal. Biochem.*, 2006, **355**, 62–70.
35. G. Zheng and W. S. Price, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2010, **56**, 267–288.
36. Q. N. Van, G. N. Chmurny, and T. D. Veenstra, *Biochem. Biophys. Res. Commun.*, 2003, **301**, 952–959.
37. M. Kriat, S. Confort-Gouny, J. Vion-Dury, P. Viout, and P. J. Cozzone, *Biochimie*, 1992, **74**, 913–918.
38. L. H. Lucas, C. K. Larive, P. S. Wilkinson, and S. Huhn, *J. Pharm. Biomed. Anal.*, 2005, **39**, 156–163.
39. C. Ludwig and M. R. Viant, *Phytochem. Anal.*, 2009, **21**, 22–32.
40. D. R. Ekman, Q. Teng, D. L. Villeneuve, M. D. Kahl, K. M. Jensen, E. J. Durhan, G. T. Ankley, and T. W. Collette, *Environ. Sci. Technol.*, 2008, **42**, 4188–4194.
41. I. Katsiadaki, T. D. Williams, J. S. Ball, T. P. Bean, M. B. Sanders, H. Wu, E. M. Santos, M. M. Brown, P. Baker, F. Ortega, F. Falciani, J. A. Craft, C. R. Tyler, M. R. Viant, and J. K. Chipman, *Aquat. Toxicol.*, 2010, **97**, 174–187.
42. T. B. Schock, S. Newton, K. Brenkert, J. Leffler, and D. W. Bearden, An NMR-based metabolomic assessment of cultured cobia health in response to dietary manipulation', *Food Chemistry*, 2012, http://dw.doi.org/10.1016/j.foodchem.2011.12.077, in press.
43. Y. Xi and D. M. Rocke, *BMC Bioinformatics*, 2008, **9**, 324. http://dw.doi.org/10.1186/1471-2105-9-324.
44. J. C. Cobas, M. A. Bernstein, M. Martin-Pastor, and P. G. Tahoces, *J. Magn. Reson.*, 2006, **183**, 145–151.
45. E. Holmes, P. J. Foxall, J. K. Nicholson, G. H. Neild, S. M. Brown, C. R. Beddell, B. C. Sweatman, E. Rahr, J. C. Lindon, M. Spraul, and P. Neidig, *Anal. Biochem.*, 1994, **220**, 284–296.
46. M. Spraul, P. Neidig, U. Klauck, P. Kessler, E. Holmes, J. K. Nicholson, B. C. Sweatman, S. R. Salman, R. D. Farrant, E. Rahr, C. R. Beddell, and J. C. Lindon, *J. Pharm. Biomed. Anal.*, 1994, **12**, 1215–1225.
47. P. E. Anderson, D. A. Mahle, T. E. Doom, N. V. Reo, N. J. DelRaso, and M. L. Raymer, *Metabolomics*, 2010, **7**, 179–190.
48. S. M. Kohl, M. S. Klein, J. Hochrein, P. J. Oefner, R. Spang, and W. Gronwald, *Metabolomics*, 2011. http://dw.doi.org/ 10.1007/s11306-011-0350-z.
49. S. Zhang, L. Liu, D. Steffen, T. Ye, and D. Raftery, *Metabolomics*, 2011. http://dx.doi.org/10.1007/s11306-011-0315-2.
50. P. V. Purohit, D. M. Rocke, M. R. Viant, and D. L. Woodruff, *Omics-a J. Integr. Biol.*, 2004, **8**, 118–130.
51. A. D. Southam, J. M. Easton, G. D. Stentiford, C. Ludwig, T. N. Arvanitis, and M. R. Viant, *J. Proteome Res.*, 2008, **7**, 5277–5285.
52. G. X. Xie, T. L. Chen, Y. P. Qiu, P. Shi, X. J. Zheng, M. M. Su, A. H. Zhao, Z. T. Zhou, and W. Jia, *Metabolomics*, 2012, http://dx.doi.org/10.1007/s11306-011-0302-7, in press.
53. E. Szymañska, E. Saccenti, A. K. Smilde, and J. A. Westerhuis, *Metabolomics*, 2011. http://dx.doi.org/10.1007/s11306-011-0330-3.
54. J. Westerhuis, H. Hoefsloot, S. Smit, D. Vis, A. Smilde, E. van Velzen, J. van Duijnhoven, and F. van Dorsten, *Metabolomics*, 2008, **4**, 81–89.
55. S. Wiklund, D. Nilsson, L. Eriksson, M. Sjöström, S. Wold, and K. Faber, *J. Chemom.*, 2007, **21**, 427–439.
56. L. Eriksson, J. Trygg, and S. Wold, *J. Chemom.*, 2008, **22**, 594–600.
57. S. Mahadevan, S. L. Shah, T. J. Marrie, and C. M. Slupsky, *Anal. Chem.*, 2008, **80**, 7562–7570.
58. Q. Cui, I. A. Lewis, A. D. Hegeman, M. E. Anderson, J. Li, C. F. Schulte, W. M. Westler, H. R. Eghbalnia, M. R. Sussman, and J. L. Markley, *Nat. Biotechnol.*, 2008, **26**, 162–164.
59. D. S. Wishart, D. Tzur, C. Knox, R. Eisner, A. C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M. A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D. D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G. E. Duggan, G. D. Macinnis, A. M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B. D. Sykes, H. J. Vogel, and L. Querengesser, *Nucleic Acids Res.*, 2007, **35**, D521–D526.
60. D. S. Wishart, C. Knox, A. C. Guo, R. Eisner, N. Young, B. Gautam, D. D. Hau, N. Psychogios, E. Dong, S. Bouatra, R. Mandal, I. Sinelnikov, J. Xia, L. Jia, J. A. Cruz, E. Lim, C. A. Sobsey, S. Shrivastava, P. Huang, P. Liu, L. Fang, J. Peng, R. Fradette, D. Cheng, D. Tzur, M. Clements, A. Lewis, A. De Souza, A. Zuniga, M. Dawe, Y. Xiong, D. Clive, R. Greiner, A. Nazyrova, R. Shaykhutdinov, L. Li, H. J. Vogel, and I. Forsythe, *Nucleic Acids Res.*, 2009, **37**, D603–D610.
61. C. Ludwig, J. M. Easton, A. Lodi, S. Tiziani, S. E. Manzoor, A. D. Southam, J. J. Byrne, L. M. Bishop, S. He, T. N. Arvanitis, U. L. Günther, and M. R. Viant, *Metabolomics*, 2012. **8**, 8–18.
62. M. Alum, P. Shaw, B. Sweatman, B. Ubhi, J. Haselden, and S. Connor, *Metabolomics*, 2008, **4**, 122–127.
63. R. A. Kleps, T. C. Myers, R. N. Lipcius, and T. O. Henderson, *PLoS ONE*, 2007, **2**, e780.
64. S. Moura, A. Ultramari Mde, D. M. de Paula, M. Yonamine, and E. Pinto, *Toxicon*, 2009, **53**, 578–583.
65. P. Polychronopoulos, P. Magiatis, A. L. Skaltsounis, F. Tillequin, E. Vardala-Theodorou, and A. Tsarbopoulos, *Nat. Prod. Lett.*, 2001, **15**, 411–418.
66. J. J. Chen, H. M. Hsueh, R. R. Delongchamp, C. J. Lin, and C. A. Tsai, *BMC Bioinformatics*, 2007, **8**, 412. http://dw.doi.org/10.1186/1471-2105-8-412.
67. A. W. Bell, E. W. Deutsch, C. E. Au, R. E. Kearney, R. Beavis, S. Sechi, T. Nilsson, J. J. M. Bergeron, and H. T. S. W. Group, *Nat. Methods*, 2009, **6**, 423–430.

Encyclopedia of Magnetic Resonance, Online © 2007–2012 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Magnetic Resonance* in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470034590.emrstm1256

68. J. P. A. Ioannidis, D. B. Allison, C. A. Ball, I. Coulibaly, C. Xiangqin, An. C. Culhane, M. Falchi, C. Furlanello, L. Game, G. Jurman, J. Mangion, T. Mehta, M. Nitzberg, G. P. Page, E. Petretto, and V. van Noort, *Nat. Genet.*, 2009, **41**, 149–155.
69. M. Mann, *Nat. Methods*, 2009, **6**, 717–719.
70. O. Fiehn, D. Robertson, J. Griffin, M. Werf, B. Nikolau, N. Morrison, L. W. Sumner, R. Goodacre, N. W. Hardy, C. Taylor, J. Fostel, B. Kristal, R. Kaddurah-Daouk, P. Mendes, B. Ommen, J. C. Lindon, and S.-A. Sansone, *Metabolomics*, 2007, **3**, 175–178.
71. D. V. Rubtsov, H. Jenkins, C. Ludwig, J. Easton, M. R. Viant, U. Günther, J. L. Griffin, and N. Hardy, *Metabolomics*, 2007, **3**, 223–229.
72. N. Morrison, D. Bearden, J. G. Bundy, T. Collette, F. Currie, M. P. Davey, N. S. Haigh, D. Hancock, O. A. H. Jones, S. Rochfort, S.-A. Sansone, D. Štys, Q. Teng, D. Field, and M. R. Viant, *Metabolomics*, 2007, **3**, 203–210.
73. M. A. Warne, E. M. Lenz, D. Osborn, J. M. Weeks, and J. K. Nicholson, *Biomarkers*, 2000, **5**, 56–72.
74. J. G. Bundy, J. C. Osborn, J. M. Weeks, J. C. Lindon, and J. K. Nicholson, *FEBS Lett.*, 2001, **500**, 31–35.
75. J. G. Bundy, D. J. Spurgeon, C. Svendsen, P. K. Hankard, D. Osborn, J. C. Lindon, and J. K. Nicholson, *FEBS Lett.*, 2002, **521**, 115–120.
76. J. G. Bundy, H. Ramlov, and M. Holmstrup, *Cryo Lett.*, 2003, **24**, 347–358.
77. J. G. Bundy, H. C. Keun, J. K. Sidhu, D. J. Spurgeon, C. Svendsen, P. Kille, and A. J. Morgan, *Environ. Sci. Technol.*, 2007, **41**, 4458–4464.
78. J. G. Bundy, J. K. Sidhu, F. Rana, D. J. Spurgeon, C. Svendsen, J. F. Wren, S. R. Sturzenbaum, A. J. Morgan, and P. Kille, *BMC Biol.*, 2008, **6**, 1–25.
79. S. L. Hughes, J. G. Bundy, E. J. Want, P. Kille, and S. R. Sturzenbaum, *J. Proteome Res.*, 2009, **8**, 3512–3519.
80. S. Rochfort, V. Ezernieks, and A. Yen, *Metabolomics*, 2009, **5**, 95–107.
81. M. J. Simpson and J. R. McKelvie, *Anal. Bioanal. Chem.*, 2009. **394**, 137–149, http://dw.doi.org/10.1007/s00216-009-2612-4.
82. J. McKelvie, J. Yuk, Y. Xu, A. Simpson, and M. Simpson, *Metabolomics*, 2009, **5**, 84–94.
83. J. R. McKelvie, D. M. Wolfe, M. Celejewski, A. J. Simpson, and M. J. Simpson, *Environ. Pollut.*, 2010, **158**, 2150–2157.
84. S. A. E. Brown, J. R. McKelvie, A. J. Simpson, and M. J. Simpson, *Environ. Pollut.*, 2010, **158**, 2117–2123.
85. M. L. Whitfield Aslund, A. J. Simpson, and M. J. Simpson, *Ecotoxicology*, 2011. **20**, 836–846. http://dx.doi.org/10.1007/s10646-011-0638-9.
86. B. P. Lankadurai, D. M. Wolfe, A. J. Simpson, and M. J. Simpson, *Environ. Pollut.*, 2011. **159** (10), 2845–2851. http://dx.doi.org/10.1016/j.envpol.2011.04.044.
87. J. R. McKelvie, D. M. Wolfe, M. A. Celejewski, M. Alaee, A. J. Simpson, and M. J. Simpson, *Environ. Pollut.*, 2011. **159** (12), 3620–3626. http://dx.doi.org/10.1016/j.envpol.2011.08.002.
88. M. Whitfield Åslund, M. Celejewski, B. P. Lankadurai, A. J. Simpson, and M. J. Simpson, *Chemosphere*, 2011, **83**, 1096–1101.
89. OECD, ''OECD Guideline for Testing of Chemicals, OECD Guideline 207'', Organisation for Economic Co-operation and Development (OECD): Paris, France, 1984.
90. D. R. Livingstone, J. K. Chipman, D. M. Lowe, C. Minier, C. L. Mitchelmore, M. N. Moore, L. D. Peters, and R. K. Pipe, *Int. J. Environ. Pollut.*, 2000, **13**, 56–91.
91. M. R. Viant, J. H. Walton, and R. S. Tjeerdema, *Pestic. Biochem. Physiol.*, 2001, **71**, 40–47.
92. A. Hines, W. H. Yeung, J. Craft, M. Brown, J. Kennedy, J. Bignell, G. D. Stentiford, and M. R. Viant, *Anal. Biochem.*, 2007, **369**, 175–186.
93. W. Tuffnail, G. Mills, P. Cary, and R. Greenwood, *Metabolomics*, 2009, **5**, 33–43.
94. D. G. Robertson, *Toxicol. Sci.*, 2010, **115**, 305–306.
95. A. Hines, F. J. Staff, J. Widdows, R. M. Compton, F. Falciani, and M. R. Viant, *Toxicol. Sci.*, 2010, **115**, 369–378.
96. H. Wu and W. X. Wang, *Aquat. Toxicol.*, 2010, **100**, 339–345.
97. H. Wu and W. X. Wang, *Environ. Toxicol. Chem.*, 2010, **30**, 806–812.
98. L. Zhang, X. Liu, L. You, D. Zhou, H. Wu, L. Li, J. Zhao, J. Feng, and J. Yu, *Mar. Environ. Res.*, 2011, **72**, 33–39.
99. L. Zhang, X. Liu, L. You, D. Zhou, Q. Wang, F. Li, M. Cong, L. Li, J. Zhao, D. Liu, J. Yu, and H. Wu, *Environ. Toxicol. Pharmacol.*, 2011, **32**, 218–225.
100. X. Liu, L. Zhang, L. You, M. Cong, J. Zhao, H. Wu, C. Li, D. Liu, and J. Yu, *Environ. Toxicol. Pharmacol.*, 2011, **31**, 323–332.
101. M. R. Viant, *Mol. Biosyst.*, 2008, **4**, 980–986.
102. T. B. Schock, D. A. Stancyk, L. Thibodeaux, K. G. Burnett, L. E. Burnett, A. F. B. Boroujerdi, and D. W. Bearden, *Metabolomics*, 2010, **6**, 250–262.
103. C. Y. Lin, B. S. Anderson, B. M. Phillips, A. C. Peng, S. Clark, J. Voorhees, H. D. Wu, M. J. Martin, J. McCall, C. R. Todd, F. Hsieh, D. Crane, M. R. Viant, M. L. Sowby, and R. S. Tjeerdema, *Aquat. Toxicol.*, 2009, **95**, 230–238.
104. A. R. Van Scoy, C. Yu Lin, B. S. Anderson, B. M. Philips, M. J. Martin, J. McCall, C. R. Todd, D. Crane, M. L. Sowby, M. R. Viant, and R. S. Tjeerdema, *Ecotoxicol. Environ. Saf.*, 2010, **73**, 710–717.
105. E. Kokushi, S. Uno, T. Harada, and J. Koyama, *Environ. Toxicol.*, 2010. http://dx.doi.org/10.1002/tox.20653.
106. L. M. Samuelsson, B. Bjorlenius, L. Forlin, and D. G. Larsson, *Environ. Sci. Technol.*, 2011, **45**, 1703–1710.
107. K. H. Chiu, S. Ding, Y. W. Chen, C. H. Lee, and H. K. Mok, *Fish Physiol. Biochem.*, 2011, **37**, 701–707.
108. S. Hayashi, M. Yoshida, T. Fujiwara, S. Maegawa, and E. Fukusaki, *Z. Naturforsch., C*, 2011, **66**, 191–198.
109. G. Gibson, C. Gieger, L. Geistlinger, E. Altmaier, M. Hrabé de Angelis, F. Kronenberg, T. Meitinger, H.-W. Mewes, H. E. Wichmann, K. M. Weinberger, J. Adamski, T. Illig, and K. Suhre, *PLoS Genet.*, 2008, **4**, e1000282.
110. K. Suhre, H. Wallaschofski, J. Raffler, N. Friedrich, R. Haring, K. Michael, C. Wasner, A. Krebs, F. Kronenberg, D. Chang, C. Meisinger, H. E. Wichmann, W. Hoffmann, H. Volzke, U. Volker, A. Teumer, R. Biffar, T. Kocher, S. B. Felix, T. Illig, H. K. Kroemer, C. Gieger, W. Romisch-Margl, and M. Nauck, *Nat. Genet.*, 2011, **43**, 565–569.

Biographical Sketches

Daniel W. Bearden. *b*. 1957. BS, 1980, Electrical Engineering, Southern Methodist University, Dallas, TX; BS, 1980, Physics, Southern Methodist University, Dallas, TX; MS, 1983, Physics, Rice University, Houston, TX; PhD, 1986, Physics, Rice University, Houston, TX; Postdoctoral Fellow, Research School of Chemistry, Australian National University 1987–1989; Lecturer, 1990–1993, Department of Chemistry, Clemson University, Clemson, SC; Research Chemist, 1993–2007, National Oceanographic and Atmospheric Administration; Research Chemist, 2007–present, Analytical Chemistry Division, National Institute of Standards and Technology, Charleston, SC. Current research interests: NMR-based marine environmental metabolomics, quantitative NMR spectroscopy.

Encyclopedia of Magnetic Resonance, Online © 2007–2012 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Magnetic Resonance* in 2012 by John Wiley & Sons, Ltd.
DOI: 10.1002/9780470034590.emrstm1256