Published in final edited form as:

J Magn Reson. 2013 January ; 226C: 93–99. doi:10.1016/j.jmr.2012.11.008.

MetaboID: A Graphical User Interface Package for Assignment of ^1H NMR Spectra of Bodyfluids and Tissues

Neil MacKinnon^{1,2,*}, Bagganahalli S. Somashekar^{1,2}, Pratima Tripathi^{1,2}, Wencheng Ge², Thekkelnaycke M. Rajendiran³, Arul M. Chinnaiyan^{3,4,5,6}, and Ayyalusamy Ramamoorthy^{1,2,*}

¹Biophysics, University of Michigan, Ann Arbor, MI, USA

²Department of Chemistry, University of Michigan, Ann Arbor, MI, USA

³Michigan Center for Translational Pathology, Department of Pathology, University of Michigan Medical School, Ann Arbor, MI, USA

⁴Department of Urology, University of Michigan Medical School, Ann Arbor, MI, USA

⁵Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, MI, USA

⁶Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, MI, USA

Abstract

Nuclear magnetic resonance based measurements of small molecule mixtures continues to be confronted with the challenge of spectral assignment. While multidimensional experiments are capable of addressing this challenge, the imposed time constraint becomes prohibitive, particularly with the large sample sets commonly encountered in metabolomic studies. Thus, one-dimensional spectral assignment is routinely performed, guided by two-dimensional experiments on a selected sample subset; however, a publicly available graphical interface for aiding in this process is currently unavailable. We have collected spectral information for 360 unique compounds from publicly available databases including chemical shift lists and authentic full resolution spectra, supplemented with spectral information for 25 compounds collected in-house at a proton NMR frequency of 900 MHz. This library serves as the basis for MetaboID, a Matlab-based user interface designed to aid in the one-dimensional spectral assignment process. The tools of MetaboID were built to guide resonance assignment in order of increasing confidence, starting from cursory compound searches based on chemical shift positions to analysis of authentic spike experiments. Together, these tools streamline the often repetitive task of spectral assignment. The overarching goal of the integrated toolbox of MetaboID is to centralize the one dimensional spectral assignment process, from providing access to large chemical shift libraries to providing a straightforward, intuitive means of spectral comparison. Such a toolbox is expected to be attractive

© 2012 Elsevier Inc. All rights reserved.

*Corresponding Authors: Dr. Neil MacKinnon, Biophysics and Department of Chemistry, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109-1055, USA, Phone: (734) 615-0474, Fax: (734) 764-3323, nmackinn@umich.edu. Prof. Ayyalusamy Ramamoorthy, Biophysics and Department of Chemistry, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109-1055, USA, Phone: (734) 647-6572, Fax: (734) 764-3323, ramamoor@umich.edu.

Supplementary material

Supplementary information for this article may be found, in the online version, at [websiteDOI](#). This includes further experimental details (supplemental file 1) and a MetaboID user manual.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

to both experienced and new metabolomic researchers as well as general complex mixture analysts.

Keywords

NMR; metabolomics; assignment; mixture analysis

1. Introduction

Metabolomics is increasingly utilized in biomedical and clinical studies to better understand various diseases. As an analytical tool, nuclear magnetic resonance (NMR) spectroscopy continues to make invaluable contributions to the field of metabolomics/metabonomics. With increasing availability of high magnetic-field spectrometers (e.g., 900 MHz) coupled with advanced data collection and analysis techniques (non-linear sampling [1], covariance [2–4], STOCSY [5]), and extension into selective one-dimensional (1D) and two-dimensional (2D) NMR techniques (e.g. sel-TOCSY [6,7], JRES [8]), NMR-based investigations now rival mass spectrometry metabolomic studies in terms of the large sample cohorts that can be analyzed [9,10]. A significant challenge for NMR-based metabolomic studies remains the assignment of all small molecules contributing to the 1D proton (^1H) spectrum.

One-dimensional ^1H NMR spectra are by far the most common data routinely collected on biological fluids, cell extracts, and tissues due to the high NMR sensitivity of the hydrogen nucleus. While several tools for processing and preparing raw NMR data for multivariate statistical analysis are publicly available (NMRLAB [11], ProMetab [12], matNMR [13], MetaboAnalyst [14], MetaboLab [15], rNMR [16]), the challenge remains resonance assignment which is inhibited by the severe signal overlap (or spectral crowding). The traditional solution to the overlap problem is to extend the NMR experiment into higher frequency dimensions, either with homonuclear (e.g. JRES, COSY, and TOCSY) or heteronuclear (e.g. HSQC and HMBC) experiments. This strategy is limited due to the large time requirements of the 2D experiment, which may be as much as 10 – 40 times longer than that of the 1D ^1H experiment. The collection of a full suite of 2D experiments required for full spectral assignment of every sample is clearly a daunting task, particularly when considering the large sample cohorts typically utilized in metabolomics studies. However, it is not the case that every sample is completely unique within a cohort; often there is significant correspondence between the samples and thus selecting one or several representative samples to undergo the full battery of 2D experimental measurements is all that is necessary. Thus, the final analysis still requires resonance assignment of all the peaks in 1D ^1H spectra, with the knowledge of the 2D spectra of selected samples guiding the process.

The challenge in assigning 1D ^1H spectra lies in streamlining the method and attempting to minimize the repetitive procedure of candidate resonance assignment. While public metabolite databases offer search functions based on NMR chemical shift lists (HMDB [17,18] and BMRB [19]), which is often the first step in the resonance assignment process, a convenient method to directly compare an experimental spectrum with an authentic compound spectrum is unavailable. Excellent commercial products such as Chenomx and Bruker BioSpin's AMIX are available; however, to the best of our knowledge only recently have publicly available tools designed for the assignment problem been proposed [20]. We present a graphical interface designed to aid in 1D ^1H spectral assignment, which is written within the technical programming environment of MatlabTM (R2010b, Mathworks, Natick, MA). The interface, named MetaboID, is built upon a library of 360 unique small

molecule ^1H NMR spectra compiled from public metabolomic databases (HMDB [17,18] and BMRB [19]), and offers a convenient workspace for comparing experimental spectra with the spectra of authentic reference compounds to guide the resonance assignment process. Importantly, generating user-specific libraries (e.g. magnetic field specific, solvent specific, etc.) is straightforward and thus MetaboID is fully customizable to the particular user's requirements. Overall, the central goal of MetaboID is to provide a user friendly platform for the chemical assignment of metabolites by providing an interface to a large NMR spectral library of authentic compounds (360 compounds) with a seamless ability to plot individual authentic spectra while interacting with experimental NMR data of biofluids.

2. Implementation

MetaboID is a collection of three primary user interfaces designed to aid in the various aspects of 1D ^1H NMR spectral assignment. The design is such that assignment of each spectrum may be considered as an individual task with associated task-specific files (managed by a File Manager) including an authentic compound peak-list library, an experimental spectrum (an *.xls(x) file), compound sub-lists, and experimental spectra collected after authentic compound spiking. The overall envisioned workflow was divided into the following tasks (Scheme 1):

1. Preliminary authentic compound searching via chemical shift peak lists obtained, for example, directly from the 1D spectrum or from 2D homonuclear correlation spectra (e.g., COSY, TOCSY, NOESY).
2. Visualization of the authentic spectra of potential compounds (also *.xls(x) files), either as individual spectra or overlaid with an experimental spectrum.
3. Assignment confirmation by comparison of an experimental spectrum before and after spiking the sample with an authentic compound.
4. Generation of task-specific compound lists which may include compounds that are safely excluded, candidate assignments, or targeted profiling compounds. These lists can then be used to edit the main compound list and reduce redundant or repeated searches.

In order for such a tool to be useful, a comprehensive NMR spectral library of authentic compounds should be available. MetaboID has been built around a library of 360 unique compounds to date, including most common metabolites encountered in metabolomic studies.

A library of authentic compound data is critical in undertaking complex mixture analyses. The library provided with MetaboID is composed of publicly available compound data (HMDB and BMRB) including compound names and common alternative names, CAS number, KEGG compound identification [21], non-standard NMR sample conditions (standard conditions were considered to be aqueous at pH 7), chemical shift peak lists, authentic 1D ^1H NMR spectra, and compound structures. From this database, the user has the ability to search for particular chemical shift values, groups of chemical shift values, or compound names with optional display of chemical shift positions, authentic 1D ^1H spectra, and/or the chemical structure. Editing or creating a user-specific library to be used with MetaboID is straightforward, so long as the structure of the main compound chemical shift Excel file is maintained (see the User Manual, supplemental file 2).

Assignment generally begins with library searches based on selected signal chemical shift values, followed by successive elimination of potential candidate compounds through authentic and experimental resonance comparisons (i.e. chemical shift, multiplicity and relative intensity). Using MetaboID, a coarse search may begin with a particular chemical

shift value using the Peak Searching interface that is designed to efficiently identify candidate compounds at low computational cost. This is accomplished via searching and displaying chemical shift values instead of the complete high-resolution ^1H NMR spectra, which is the computationally expensive alternative. Refinement of the candidate compound list may be done by examination of the chemical shift positions or by loading the full resolution authentic spectra. Once a full resolution spectrum is imported into MetaboID, normalization to a total intensity of 1 is performed and the intensity within the region between 4.5 – 6.0 ppm is set to 0 (water and urea region) to eliminate the dominant solvent signal. The solvent region may be defined by the user through the File Management interface (see User Manual, supplemental file 2).

Using the user-friendly Authentic Spectral Overlay interface, confidence is then gained by overlaying the authentic spectra with the experimental spectrum. At this stage of the assignment process, a short-list of candidate compounds for a given unknown resonance will be generated and a select series of authentic spike-in experiments can be performed to aid in the assignment verification. For convenience, an Authentic Spike Analysis interface has been included for simple comparison of experimental spectra pre- and post-authentic compound spiking. Alternatively, this interface could be utilized for a comparison of any two experimental spectra, for example an original and a selective-TOCSY spectrum. Together, the integrated toolbox of MetaboID is meant to centralize the 1D ^1H resonance assignment process by providing access to large chemical shift libraries to enable a straightforward, intuitive means of spectral comparison.

3. Results and Discussion

3.1 General Overview

For demonstration of the envisioned workflow in assigning a 1D ^1H NMR spectrum using MetaboID, a ^1H NMR spectrum of the aqueous extract from a prostate cancer cell line (LnCAP, available from the ATCC) was collected. A 900 MHz NMR spectrometer (Bruker Biospin, Rheinstetten, Germany) was utilized to measure the ^1H NMR spectrum of LnCAP cell extract and spectra of 25 individual authentic compounds (further experimental details are available supplemental file 1). The 25 authentic compound spectral data generated in-house was incorporated into the compound library, supplementing the spectra of 360 authentic compounds obtained from public databases (HMDB [17,18] and BMRB [19]). The full ^1H NMR spectrum of the aqueous cell extract is given in Fig. 1, and for demonstration of MetaboID functionality the region between 0.92 and 1.05 ppm was subjected to preliminary assignment processing (Fig. 1, inset).

3.2 Peak Searching Interface

The Peak Searching interface was designed as the first step in the assignment processing, offering a computationally inexpensive method for identifying candidate compounds. The user has the ability to search the compound library by several criteria, including a single chemical shift, multiple chemical shifts (assumed to be from the same molecule), a chemical shift range, or by name. For demonstration, a search was performed using the chemical shift region option, selecting the 0.92 to 1.05 ppm range. As a result, 24 compounds within the chemical shift library contained at least one resonance satisfying the search criteria. The compounds are then sorted according to their highest frequency chemical shift and subsequently displayed along with all chemical shift positions (Fig. 2). At this stage, the user has the option to display all or a selection of the identified compound 1D ^1H NMR spectra in order to aid in narrowing potential candidates, which is available through the Authentic Spectrum Plotting option (Fig. 2). In addition, a tool for displaying the chemical structure of each of the candidate compounds is provided which can aid in predicting expected J -

coupling patterns and chemical shifts to further narrow the candidate compound list. Alternatively, the Authentic Spectral Overlay interface may be initialized for a clear visual comparison of the experimental spectrum with candidate compound spectra.

3.3 Authentic Spectral Overlay Interface

Further confidence in spectral assignment is gained by comparison between authentic and experimental spectra. Thus, the Authentic Spectral Overlay interface was designed to address this step in the assignment workflow. Various options are available for a visual comparison of one or several authentic compound spectra with the experimental spectrum, in addition to an integrated search capability to assist in candidate compound identification. For demonstration, a text search was performed for compounds that were collected in-house, followed by selection of valine and leucine. The spectral region between 0.92 and 1.05 ppm was selected, and all authentic spectra were overlaid after interactive adjustment of their individual intensities (Fig. 3). The excellent match of the chemical shift and *J*-coupled splitting patterns lends confidence to the candidate assignments.

It is important to note the ‘assignment guidance’ aspect of MetaboID, particularly in instances where the experimental spectrum is measured at a proton resonance frequency that differs from the authentic compound spectrum. This is highlighted in Fig. 4, where the experimental spectrum measured at 900 MHz is compared with two valine authentic spectra, one collected at 900 MHz and another at 500 MHz (500 MHz data obtained from the BMRB [19]). While the average chemical shift of the two valine methyl groups are nearly identical at both magnetic field strengths, the *J*-coupling will not be the same in terms of ppm units (they will be the same in units of Hertz) and thus assignment ambiguity may arise when using data collected at different field strengths. This emphasizes the importance of ensuring all data is collected at the same magnetic field strength for robust assignments, and is the reason why the MetaboID chemical shift library is designed to be easily edited for user-specific requirements.

3.4 Authentic Spiking Interface

A common experimental technique utilized to confirm a resonance assignment is to add a small volume of a known authentic compound concentration into the experimental sample. Slight variations in sample conditions (i.e. pH and salt concentration) may result in slight chemical shift mismatches after overlaying authentic and experimental spectra. Thus, measurement of the experimental mixture spectrum before and after spiking, where the sample conditions are identical, should eliminate this ambiguity and significantly increase the confidence in the assignment. MetaboID includes an interface for simple analysis of such spike-in experiments (Fig. 5). For demonstration, leucine was spiked to the cell extract sample and the 900 MHz ¹H NMR spectrum was measured. Overlaying the original and spiked spectra using the Authentic Spike Analysis tool indicates an intensity increase of the signal at 0.95 ppm, which was assigned to leucine using the Authentic Spectral Overlay tool (Fig. 5). Plotting the difference between the original and spiked spectra clearly reproduces the authentic leucine spectrum.

It is worth noting that the Authentic Spectral Overlay interface may be used for a comparison of two spectra before and after any type of selective compound enhancement. While initially designed for an enhancement of the concentration of a candidate compound, an NMR experiment such as selective total correlation spectroscopy (sel-TOCSY) could be performed with the same intentions. Selective excitation of a candidate resonance followed by a TOCSY mixing scheme results in a 1D spectrum of the *J*-coupled spin network associated with the primary resonance. Examination of the original and sel-TOCSY 1D

spectra would report on which resonances arise from the same compound, further aiding in the assignment process.

3.5 Future Development

Assignment of 1D ^1H NMR spectra in an efficient fashion while ensuring a large degree of user control continues to be a challenging problem. As a first generation model, MetaboID aims to address this problem by providing a collection of tools useful for analyses common to the assignment task in a user-friendly environment. As an open source toolbox, it is envisioned that user-specific modification will force the evolution of MetaboID into exciting new functionalities.

Several open issues remain to be addressed which are currently underway. We anticipate that the ability to generate commonly encountered compound lists will be useful as search filters. For example, generating a list of compounds based on the proton resonance frequency, the sample conditions, or the chemical class (e.g. amino acids, sugars, etc.) could be easily done using the Compound List Manager tool. Maintenance of the chemical shift library is critical, and a tool for automatically generating the peak list from an authentic spectrum and adding the data to a user-specified library will be developed. In order to reduce assignment ambiguities arising from J – coupling mismatch due to spectra collected at various magnetic field strengths, a calculator for predicting an authentic spectrum at different field strengths will be incorporated. The ability to highlight resonances that have tentative assignments is an additional feature currently being explored. We are investigating the ability to incorporate the chemical shift assignments into the chemical structure image files displayed in the Peak Searching interface. Finally, integration of 2D NMR experimental data into MetaboID will be expected to greatly enhance resonance assignment. For example, authentic compound peak lists obtained from 2D NMR experiments, including homonuclear (e.g. ^1H - ^1H COSY, TOCSY) and heteronuclear (e.g. ^1H - $^{13}\text{C}/^{15}\text{N}$ HSQC, HMBC) techniques could be incorporated into a new assignment tool for 2D NMR analyses.

3.6 Availability and Requirements

MetaboID is available at no cost for academic research. To obtain a copy of the user interface and library, please visit <http://rams.biop.lsa.umich.edu/cancer-metabolomics.aspx> for further information. No part of MetaboID may be reproduced for commercial benefit. The current version has been developed for Windows-based systems; however, a Linux-compatible version will be made available. MetaboID is coded in the technical computing environment of MatlabTM, and thus the user must have a standard MatlabTM installation.

4. Conclusions

Extracting the rich information contained within a typical ^1H NMR spectrum obtained in a metabolomics study continues to be challenging. While multivariate techniques can efficiently identify important signals related to the biological question at hand, characterizing the molecule responsible for the signal can be difficult. MetaboID was designed as a user-friendly, customizable interface for accessing the vast amount of authentic metabolite NMR data in an intuitive, centralized fashion. In this way, the resonance assignment task becomes streamlined and the repetitive compound querying utilizing publicly available online databases is significantly reduced. Thus, we believe MetaboID will be useful not just for the experienced NMR metabolomic expert, but to new researches exploring the metabolomics field and to researchers in fields in which NMR-based complex mixture analyses are undertaken. In particular, when coupled with data processing and preparation tools such as NMRLAB, matNMR, MetaboLab, MetaboAnalyst,

ProMetab, MetaboHunter, and rNMR [11–16, 20], a powerful toolkit is realized for complex mixture analysis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Jeffrey Brender and Ms. Stéphanie Le Clair for their careful review and insightful discussions. This study is partly supported by funds from NIH (to A.R.).

List of abbreviations used

1D/2D	One-Dimensional/Two-Dimensional
BMRB	Biological Magnetic Resonance Bank
HMBC	Heteronuclear Multiple Bond Correlation
HMDB	Human Metabolome Database
HSQC	Heteronuclear Single Quantum Coherence
JRES	J-coupling Resolved
NMR	Nuclear Magnetic Resonance
SeITOCYSY	Selective Total Correlation Spectroscopy
STOCSY	Statistical Total Correlation Spectroscopy
TOCSY	Total Correlation Spectroscopy

References

1. Pontoizeau C, Herrmann T, Toulhoat P, Elena-Herrmann B, Emsley L. Targeted projection NMR spectroscopy for unambiguous metabolic profiling of complex mixtures. *Magn Reson Chem.* 2010; 48:727–733. [PubMed: 20648569]
2. Brüschweiler R, Zhang F. Covariance nuclear magnetic resonance spectroscopy. *J Chem Phys.* 2004; 120:5253–5260. [PubMed: 15267396]
3. Brüschweiler R. Theory of covariance nuclear magnetic resonance spectroscopy. *J Chem Phys.* 2004; 121:409–414. [PubMed: 15260561]
4. Zhang F, Brüschweiler R, Zhang F. Indirect Covariance NMR Spectroscopy. *J Am Chem Soc.* 2004; 126:13180–13181. [PubMed: 15479045]
5. Cloarec O, Dumas ME, Craig A, Barton RH, Trygg J, Hudson J, Blancher C, Gauguier D, Lindon JC, Holmes E, Nicholson J. Statistical Total Correlation Spectroscopy: An Exploratory Approach for Latent Biomarker Identification from Metabolic ¹H NMR Data Sets. *Anal Chem.* 2005; 77:1282–1289. [PubMed: 15732908]
6. Sandusky P, Raftery D. Use of Selective TOCSY NMR Experiments for Quantifying Minor Components in Complex Mixtures: Application to the Metabonomics of Amino Acids in Honey. *Anal Chem.* 2005; 77:2455–2463. [PubMed: 15828781]
7. Sandusky P, Raftery D. Use of Semiselective TOCSY and the Pearson Correlation for the Metabonomic Analysis of Biofluid Mixtures: Application to Urine. *Anal Chem.* 2005; 77:7717–7723. [PubMed: 16316181]
8. Ludwig C, Viant MR. Two-dimensional *J*-resolved NMR Spectroscopy: Review of a Key Methodology in the Metabolomics Toolbox. *Phytochem Anal.* 2010; 21:22–32. [PubMed: 19904730]

9. Giskeødegård GF, Grinde MT, Sitter B, Axelson DE, Lundgren S, Fjøsne HE, Dahl S, Gribbestad IS, Bathen TF. Multivariate Modeling and Prediction of Breast Cancer Prognostic Factors Using MR Metabolomics. *J Proteome Res.* 2010; 9:972–979. [PubMed: 19994911]
10. Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J, Laxman B, Mehra R, Lonigro RJ, Li Y, Nyati MK, Ahsan A, Kalyana-Sundaram S, Han B, Cao X, Byun J, Omenn GS, Ghosh D, Pennathur S, Alexander DC, Berger A, Shuster JR, Wei JT, Varambally S, Beecher C, Chinnaiyan AM. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature.* 2009; 457:910–915. [PubMed: 19212411]
11. Günther UL, Ludwig C, Rüterjans H. NMRLAB – Advanced NMR Data Processing in Matlab. *J Magn Reson.* 2000; 145:201–208. [PubMed: 10910688]
12. Viant MR. Improved methods for the acquisition and interpretation of NMR metabolomic data. *Biochem Biophys Res Co.* 2003; 310:943–948.
13. Van Beek JD. matNMR: A flexible toolbox for processing, analyzing and visualizing magnetic resonance data in Matlab®. *J Magn Res.* 2007; 187:19–26.
14. Xia J, Psychogios N, Young N, Wishart DS. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* 2009; 37(Web Server Issue):W652–W660. [PubMed: 19429898]
15. Ludwig C, Günther UL. MetaboLab – advanced NMR data processing and analysis for metabolomics. *BMC Bioinformatics.* 2011; 12:366. [PubMed: 21914187]
16. Lewis IA, Schommer SC, Markley JL. rNMR: open source software for identifying and quantifying metabolites in NMR spectra. *Magn Reson Chem.* 2009; 47:S123–S126. [PubMed: 19821464]
17. Wishart DS, Knox C, Guo AC, Eisner R, Young N, Bautam B, Hau DD, Psychogios N, Dong E, Bouatra S, Mandal R, Sinelnikov I, Xia J, Jia L, Crux JA, Lim E, Sobsey CA, Shrivastava S, Huang P, Liu P, Fang L, Peng J, Fradette R, Cheng D, Tzur D, Clements M, Lewis A, De Souza A, Zuniga A, Dawe M, et al. HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res.* 2009; 37(Database Issue):D603–D610. [PubMed: 18953024]
18. Wishart DS, Tzur D, Knox C, Eisner R, Guo AC, Young N, Cheng D, Kewell K, Arndt D, Sawhney S, Fung C, Nikolai L, Lewis M, Coutouly MA, Forsythe I, Tang P, Shrivastava S, Jeroncic K, Stothard P, Amegbey G, Block D, Hau DD, Wagner J, Miniaci J, Clements M, Gebremedhin M, Guo N, Zhang Y, Duggan GE, MacInnis GD, et al. HMDB: the Human Metabolome Database. *Nucleic Acids Res.* 2007; 35(Database Issue):D521–D526. [PubMed: 17202168]
19. Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z, Nakatani E, Schulte CF, Tolmie DE, Wenger RK, Yao H, Markley JL. BioMagResBank. *Nucleic Acids Res.* 2008; 36(Database Issue):D402–D408. [PubMed: 17984079]
20. Tulpan D, Léger S, Belliveau L, Culf A, Šupljiković M. MetaboHunter: an automatic approach for identification of metabolites from ¹H-NMR spectra of complex mixtures. *BMC Bioinformatics.* 2011; 12:400. [PubMed: 21999117]
21. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 2000; 28:27–30. [PubMed: 10592173]

Highlights

- An intuitive user interface for assignment of mixture ^1H NMR spectra is presented
- Spectral data for 360 unique compounds are provided as a base library
- Ease of new compound data addition allows MetaboID to be highly personalized
- The modular design allows for straightforward addition of further processing tools

Fig. 1. 900 MHz ^1H NMR spectrum of the aqueous extract from an LnCAP cell line. For demonstration of MetaboID, the region between 0.92 and 1.05 ppm will be assigned (red box, inset).

Fig. 2.

An example of the Peak Search interface after performing a compound search. The Peak Search interface is meant to be a convenient first approach tool for assignment. A search of compounds with at least one resonance between 0.92 and 1.05 ppm was performed, and 24 compounds were identified with their respective chemical shift positions displayed as colored dots. The search region is highlighted within the display by the red dashed vertical lines. Also shown is the structure for valine.

Fig. 3. An example demonstrating the Authentic Spectral Overlay interface. The Authentic Spectral Overlay interface aids in the refinement of the assignment process. A search for compounds whose name contains 'auth' (i.e. collected in-house at 900 MHz) was performed with valine and leucine selected from the list. After adjustment of the intensities of the individual authentic spectra, all were overlaid with the experimental spectrum with excellent resonance matching. The experimental spectrum prior to (A) and post (B) authentic spectra overlay are shown.

Fig. 4. Demonstration of the importance of noting magnetic field dependent spectral differences. A comparison of the authentic spectra of valine, obtained at 500 MHz (blue, from the BMRB) and 900 MHz (green, collected in-house) with the 900 MHz experimental spectrum. Assignment ambiguity may arise in cases where experimental spectra are collected at a magnetic field strength that differs from the authentic spectrum, and emphasizes the assignment guidance aspect of MetaboID. Ideally, a full compound library obtained at the magnetic field strength of the experimental spectra should be assembled.

Fig. 5.

Demonstration of the Authentic Spike Analysis interface. Comparison of the aqueous cell line sample pre- (blue) and post-addition (green) of authentic leucine (spike) visualized in the Authentic Spike Analysis interface. The difference spectrum (red) clearly results in the authentic leucine spectrum and further solidifies the resonance assignment.

Scheme 1.

Resonance assignment workflow through the MetaboID interfaces. Coarse searching occurs within the Peak Searching interface for candidate compound nomination, followed by fine searching within the Authentic Spectra Overlay interface. Assignment confidence increases as the processing approaches spike-in (or related) experimental confirmation.

Table 1

Summary of compound data included in MetaboID library.

Property ^a	# of Compounds
900 MHz	25
600 MHz	87
500 MHz	176
400 MHz	97
Non-standard sample conditions	18
Total Compounds	392
Unique Compounds	360

^a Compounds with spectra collected at the given ¹H resonance frequency (MHz); compounds with spectra collected under non-standard (i.e. standard considered aqueous, pH 7) conditions (e.g. CD₃OD, CDCl₃, pH 7); unique compounds include all compounds minus those collected at 900 MHz, thus 25 compounds have data at two magnetic field strengths.