

## Metabolic Profiling Strategy of *Caenorhabditis elegans* by Whole-Organism Nuclear Magnetic Resonance

Benjamin J. Blaise,<sup>†,‡</sup> Jean Giacomotto,<sup>†,§</sup> Mohamed N. Triba,<sup>†,‡,§</sup> Pierre Toulhoat,<sup>†,||</sup> Martial Piotto,<sup>⊥,¶</sup> Lyndon Emsley,<sup>‡</sup> Laurent Ségalat,<sup>§</sup> Marc-Emmanuel Dumas,<sup>‡</sup> and Bénédicte Elena<sup>\*,‡</sup>

Université de Lyon, CNRS/ENS Lyon/UCB-Lyon1, Centre de RMN à Très Hauts Champs, 5 rue de la Doua, 69100 Villeurbanne, France, Université de Lyon, CNRS/UCB-Lyon1, CGMC (UMR 5534), Bâtiment Mendel, 43 boulevard du 11 Novembre, 69622 Villeurbanne Cedex, France, INERIS, Parc Technologique Alata, BP2, 60550 Verneuil en Halatte, France, Institut de Chimie, 4 rue Blaise Pascal, Université Louis Pasteur, 67084 Strasbourg, France, and Bruker Biospin, 34 rue de l'industrie, 67166 Wissembourg, France

Received January 7, 2009

In this study, we present a methodology for metabotyping of *C. elegans* using <sup>1</sup>H high resolution magic angle spinning (HRMAS) whole-organism nuclear magnetic resonance (NMR). We demonstrate and characterize the robustness of our metabolic phenotyping method, discriminating wild-type N2 from mutant *sod-1(tm776)* animals, with the latter being an otherwise silent mutation, and we identify and quantify several confounding effects to establish guidelines to ensure optimal quality of the raw data across time and space. We monitor the sample stability under experimental conditions and examine variations arising from effects that can potentially confuse the biological interpretation or prevent the automation of the protocol, including sample culture (breeding of the worms by two biologists), sample preparation (freezing), NMR acquisition (acquisition by different spectroscopists, acquisition in different facilities), and the effect of the age of the animals. When working with intact model organisms, some of these exogenous effects are shown to be significant and therefore require control through experimental design and sample randomization.

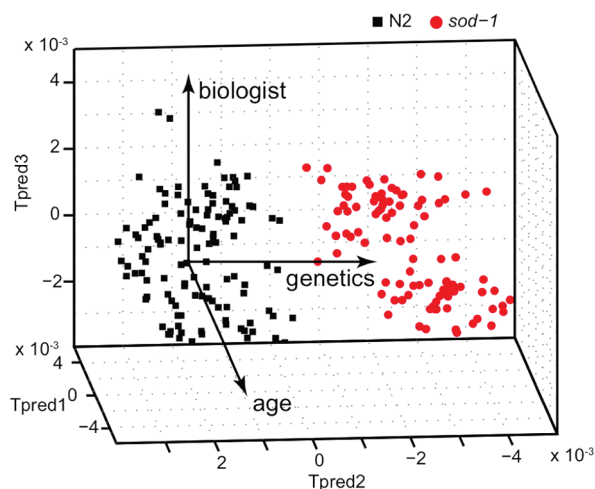
**Keywords:** metabotyping • <sup>1</sup>H HRMAS NMR • *Caenorhabditis elegans* • standardization • metabonomics • metabolomics

### Introduction

Metabonomics/metabolomics is an emerging field of post-genomic sciences, which focuses on identifying and quantifying low molecular weight compounds (metabolites) to obtain information at a molecular level and to understand the response of a living organism to pathophysiological stimuli.<sup>1</sup> This can be achieved by simultaneous quantification of metabolites using spectroscopic methods such as nuclear magnetic resonance (NMR) or mass spectrometry (MS).<sup>2,3</sup> It has been applied to a variety of biological samples to characterize features such as genetics, toxicology, metabolic regulation, and infectious disease.<sup>4–9</sup>

We have recently shown that a whole-organism NMR-based metabonomics strategy can be used to assign distinct metabolotypes to different genetically modified strains of the model organism *C. elegans*<sup>10</sup> but could be extendable to other small model organisms.<sup>11</sup> This provides the rationale for a large-scale functional genomics screening of this animal, or targeted metabolic studies in areas where *C. elegans* has been shown

to be a good platform, for example, in understanding physiological aspects of lifespan and aging, genetic diseases, drug toxicity, and pharmacology studies.<sup>12–16</sup>



**Figure 1.** O2-PLS discrimination model on genetics, age, and biologist. O2-PLS score plot discriminating 228 *C. elegans* (129 N2 and 99 *sod-1(tm776)*) spectra. Major discriminating effects are represented by age ( $Q^2 = 0.73$ ), genetics ( $Q^2 = 0.61$ ), and biologist breeding effect ( $Q^2 = 0.41$ ) axis. Other studied effects, sorted by decreasing significance, are: NMR facilities ( $Q^2 = 0.20$ ), freezing ( $Q^2 = 0.19$ ), and NMR spectroscopists ( $Q^2 = -0.44$ ).

\* To whom correspondence should be addressed. E-mail: benedicte.elena@ens-lyon.fr.

<sup>†</sup> These authors contributed equally to this work.

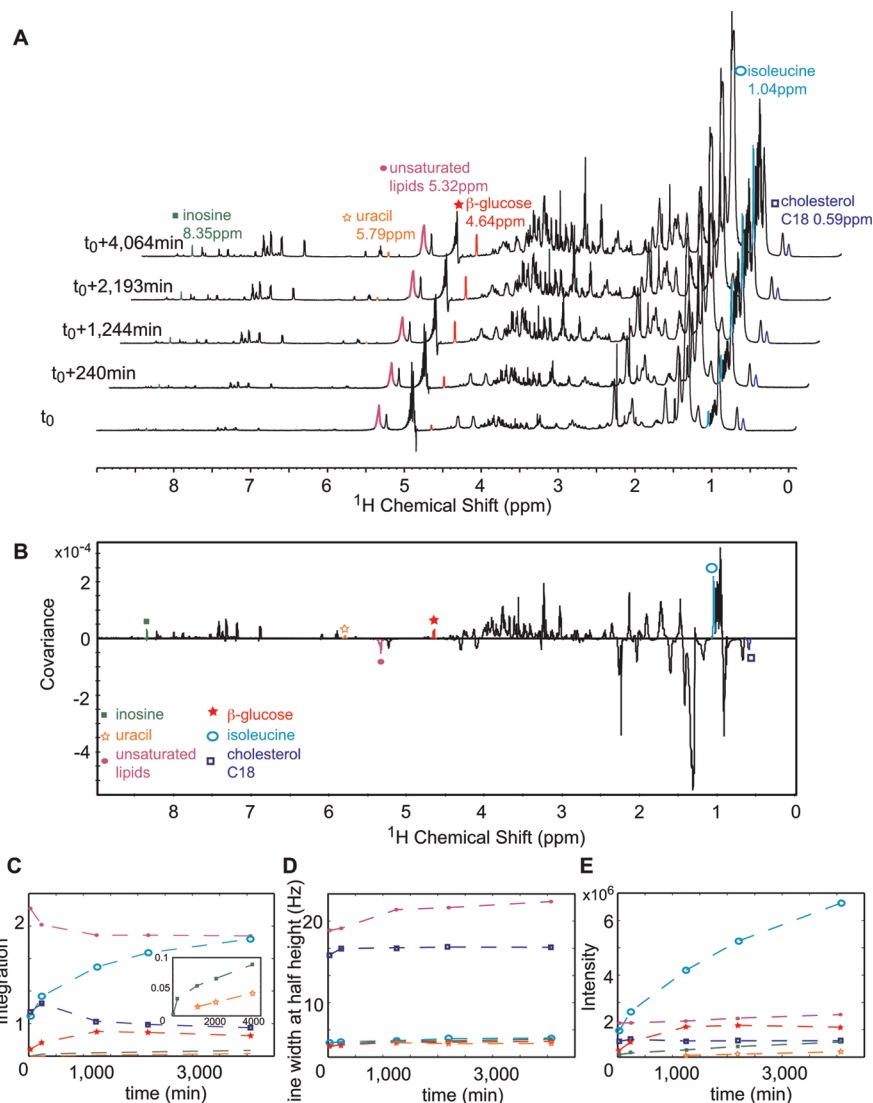
<sup>‡</sup> CNRS/ENS Lyon/UCB-Lyon1.

<sup>§</sup> CNRS/UCB-Lyon1.

<sup>||</sup> INERIS.

<sup>⊥</sup> Université Louis Pasteur.

<sup>¶</sup> Bruker Biospin.



**Figure 2.** Sample degradation. (a) 1D NOESY experiments of *C. elegans* nematodes acquired at different time intervals, 293 K, and 3.5 kHz MAS frequency. (b) O-PLS regression was done with respect to time showing that almost all signals are affected by the spinning. The evolution of six resolved signals belonging to different classes of metabolites (inosine, ■; uracil, ☆; unsaturated lipids, ●;  $\beta$ -glucose, ★; isoleucine, ○; and cholesterol, □) was monitored by measuring (c) peaks integral, (d) line-width at half-height, (e) and maximal intensity.

Whole-organism spectral data used in this type of analysis encapsulates the metabolic information available in a single data set. To interpret this high information density, the use of multivariate statistical analyses (the most popular being PCA, PLS-DA, O-PLS)<sup>17,18</sup> is a necessity to compress the latent information and extract a metabolic signature relevant to a particular question. Our ability to recover latent information is dependent on various exogenous technological factors that can influence the analyses and lead to an overestimation of discriminative capacities.<sup>19</sup> However, for such an approach to become transferable and widespread, long-term, multiuser, multicenter studies require the identification of confounding effects that could interfere with the measurements of biological effects of interest.

This work notably is in line with recent initiatives (Metabolomic Standard Initiative, Standard Metabolic Reporting Structure)<sup>20–22</sup> that have established good laboratory practices for MS and solution NMR-based metabolomics<sup>20–24</sup> and is of current particular relevance as *C. elegans* genetics has recently been the subject of several metabolomic/metabolomic studies.<sup>13,25,26</sup>

## Experimental Section

**Nematode Strains.** *C. elegans* strains used in this study are wild type (N2) and FX776 (*sod-1(tm776)*). They were obtained from the *Caenorhabditis* Genetics Center. These strains were maintained at 23 °C on NGM agar medium with OP50 bacterial lawns as food source, to keep worms well fed at all time.

***C. elegans* Preparation and Conservation.** At the appropriate stage, nematodes were fixed with formaldehyde (3.7% final) in M9 saline buffer during 45 min at room temperature under agitation and then washed four times with ultrapure water. A last wash was performed using D<sub>2</sub>O to provide a field-frequency lock signal for NMR experiments. A population of approximately 1000 worms was then filled into a 4 mm HRMAS rotor with Kel-f inserts restricting the effective sample volume to a 12  $\mu$ L sphere. A speed vacuum engine was used to remove D<sub>2</sub>O surplus. NMR acquisition was performed on the day of sample preparation except for the freezing study. In this last case, worms were prepared between one and two weeks before NMR experiments. After preparation, they were smoothly frozen at

–80 °C for storage and then were thawed at room temperature before transfer into the rotor on the day of NMR acquisition.

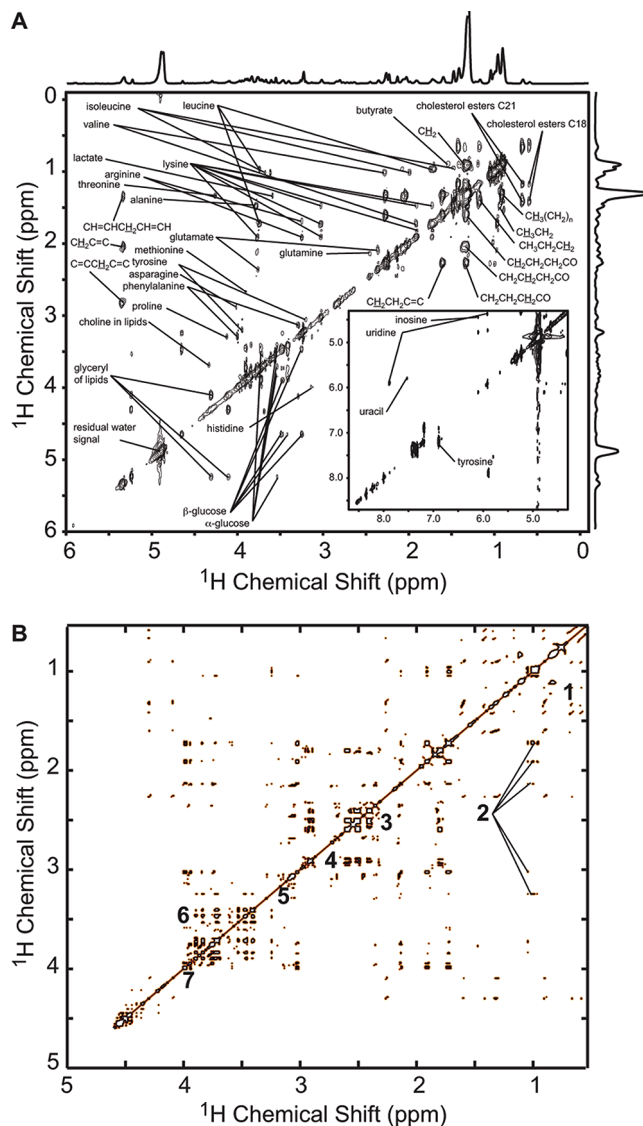
**<sup>1</sup>H HRMAS-NMR Spectroscopy.** All NMR experiments were carried out on Bruker Avance spectrometers operating at 700 MHz, using a standard double resonance (<sup>1</sup>H–<sup>13</sup>C) 4 mm HRMAS probe. Standard HRMAS one-dimensional nOe spectroscopy experiments<sup>27</sup> (recycle delay–90°–*t*–90°–*t<sub>m</sub>*–90°–acquisition) were carried out on each sample for statistical analysis. Water suppression was achieved using low power irradiation of the water resonance during the recycle delay of 1.7 s. The mixing time *t<sub>m</sub>* was set to 100 ms. The 90° pulse length was adjusted to 5.3 μs, and *t* to 4 μs. Sixteen thousand data points with 512 scans were acquired using a spectral width of 8503 Hz, for a total acquisition time of approximately 25 min. The magic angle spinning frequency was set to 3.5 kHz, and the sample temperature was regulated at 293 K. All free induction decays were multiplied by an exponential function equivalent to a 0.3 Hz line-broadening factor before Fourier transformation for a size of real spectrum of 32k data points.

Structural assignment of the metabolic signals observed in the spectra can be achieved by 2D correlation experiments, such as TOCSY, as illustrated in Figure 3a, yielding to correlations within a metabolite (spin systems), or by the measurement of the statistical correlations between NMR variables yielding to correlations between both spin systems and metabolic systems, corresponding to different resonances of the same compound or metabolites present in the same physiological pathway, in a STOCSY spectrum,<sup>28,29</sup> as shown in Figure 3b. The STOCSY can be interpreted as a pseudo 2D NMR spectrum representing the correlation matrix of a data matrix of 1D NMR spectra. The assignment is thus based on the chemical shifts reported for metabolites in the literature.<sup>30</sup>

**Data Import and Pattern Recognition.** <sup>1</sup>H HRMAS-NMR spectra were processed using the Topspin 1.3 interface (Bruker GmbH, Rheinstetten, Germany). They were reduced over the chemical shift range of –0.1 to 9 ppm with exclusion areas around residual water signal (4.6 to 5.10 ppm) using the AMIX software (Bruker GmbH) to 10 000 10<sup>–3</sup> ppm wide regions (buckets), and the signal intensity in each region was integrated. Spectra were scaled to total intensity, and integration was performed with the sum of intensities mode. The corresponding buckets table was then exported to the software Simca-P 11 (Umetrics, Umeå, Sweden) for statistical analysis.

**Multivariate Statistics.** Principal component analyses (PCA)<sup>17</sup> were carried out to check the homogeneity of each subpopulation and eventually exclude outliers. Data were then visualized by score and loading plots. In score plots, each point represents a NMR spectrum and thus a sample. Loadings points stand for NMR spectral regions and show intensity variations sustaining the distinction between subpopulations.

Orthogonal partial least-squares (OPLS)<sup>18</sup> analysis and bi-directional O2-PLS<sup>31</sup> were run to discriminate populations of nematodes by adding a supplementary data matrix *Y*, containing information about genetic, age or technological factors. These methods allow a clearer distinction between populations by canceling orthogonal information to the *Y* matrix, that are of no use for this particular discrimination. As for PCA, the results were visualized by score and loading plots. Model validations were performed by resampling the model 999 times under the null hypothesis, meaning generating models with a randomly permuted *Y* matrix not related to the factors of interest. The decrease in model goodness-of-fit statistics *R*<sup>2</sup> and

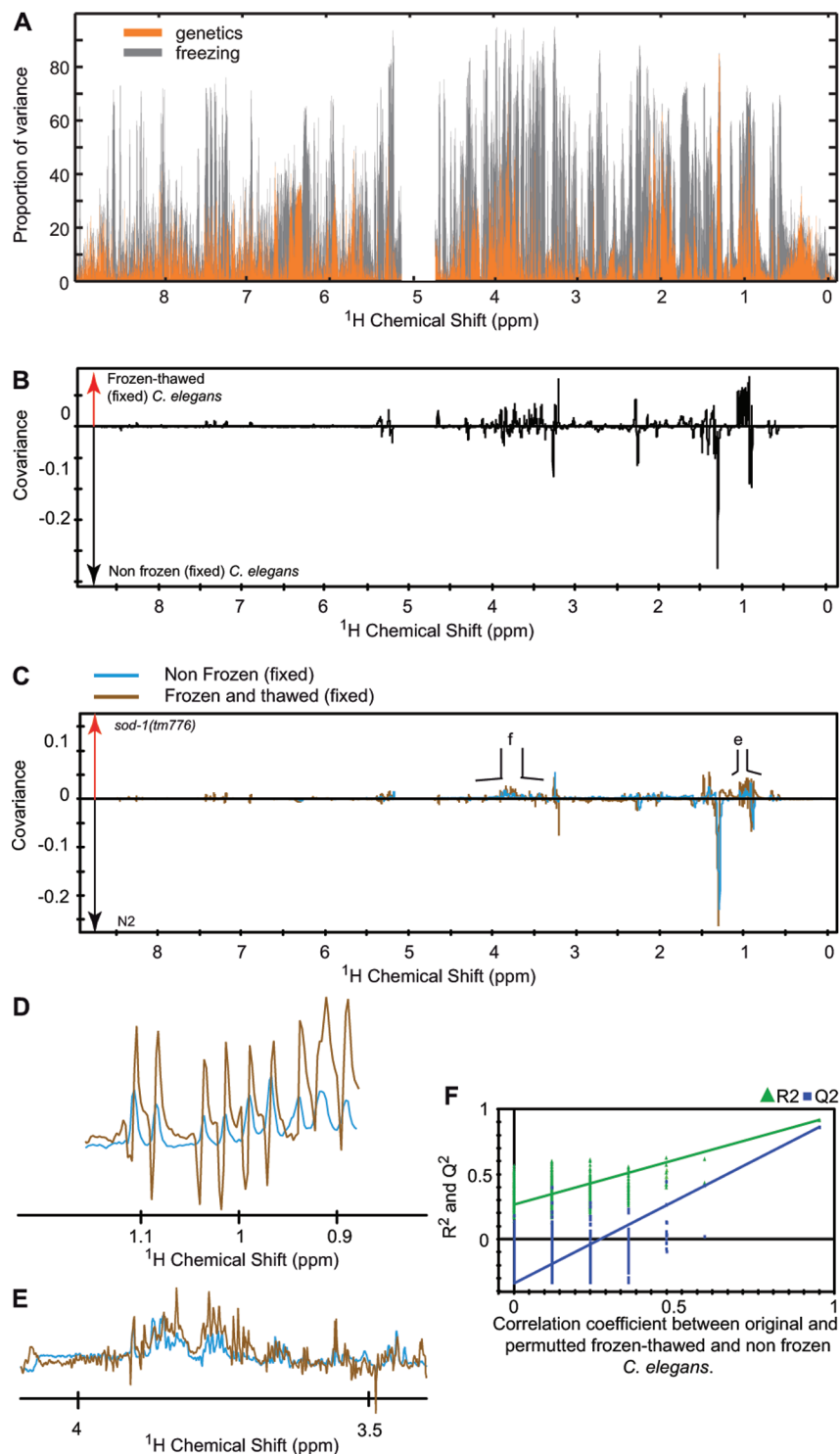


**Figure 3.** Structural assignment. (a) 2D <sup>1</sup>H TOCSY spectrum of *sod-1(tm776)* *C. elegans* mutants acquired at 700 MHz, and assigned according to the literature.<sup>30</sup> 2048 *t<sub>2</sub>* data points with 256 scans per increment and 128 *t<sub>1</sub>* data points were acquired. The spectral width in both dimensions was 12 ppm, for a total acquisition time of about 16 h. 2D NMR data is processed with a squared sine apodization of the signal corresponding to a line broadening of 2 Hz in *F<sub>2</sub>* and 1 Hz in *F<sub>1</sub>*. (b) STOCSY pseudo 2D spectrum constructed with a correlation cutoff of 0.92 from a data set of 147 1D NOESY spectra from N2 and *sod-1(tm776)* mutants. The intensity of the correlation of the NMR variables is reinforced by a color code from yellow to dark. Some correlation systems are identified: (1) lipids, (2) lysine, (3) glutamine/glutamate, (4) asparagine, (5) tyrosine/phenylalanine, (6) glucose, (7) serine/histidine.

*Q*<sup>2</sup> as a function of the correlation between the permuted and the original *Y* matrix indicates the quality of the model.

## Results and Discussion

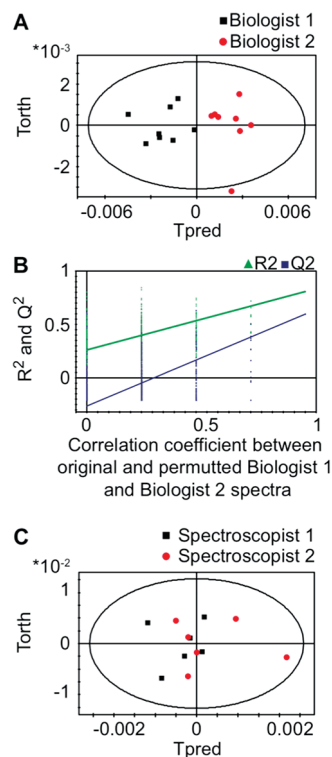
The present analysis focuses on identifying the different key steps that are necessary to extend the metabolic profiling (metabotyping) of selected mutations to a broader strategy for functional genomics of *C. elegans*, targeting the study of a large collection of mutants. To this end, the goal is to ensure that



**Figure 4.** Sample preparation: effect of genetics and sample freezing are presented on a component variance model. (b) Loading plot quantifies the effect of sample freezing for a mixture of fixed N2 and *sod-1(tm776)* *C. elegans* samples. (c) Metabotypes discriminating *sod-1(tm776)* from N2 with frozen-thawed (brown) or non frozen (blue) nematodes are reasonably similar with (d) area of important or (e) small variations. (f) Model describing the effect of freezing was validated by resampling under the null hypothesis.

the variability in the NMR data linked to genetic differences is larger than any variability connected with any other possible confounding effects leading to intersample or interindividual variability. In the current approach, interindividual effects are essentially removed by using NMR samples for analysis consisting of several hundred individuals.

Figure 1 demonstrates that, indeed, using the standardized method developed here, the discrimination induced by genetic factors is not confounded by any dispersion induced by potentially disrupting factors. Two-hundred twenty-eight  $^1\text{H}$  HRMAS NMR spectra of *C. elegans* from wild-type N2 and *sod-1(tm776)* mutants are represented in an orthogonal bidirec-



**Figure 5.** Operator effects. (a) O-PLS analysis score plot constructed from analysis of 2 sets of 8 wild-type (N2) *C. elegans* samples prepared by two different biologists with  $R^2 = 0.807$  and a  $Q^2 = 0.566$ . (b) Model resampling was performed under the null hypothesis showing a clear decrease of  $R^2$  and  $Q^2$  in random models. (c) O-PLS score plots for equivalent analysis in the case of two different NMR spectroscopists performing acquisition after shimming with two matrices of NMR shims using a rotor filled with 10% chloroform in acetone-D6. The shim quality target was set to a  $^1\text{H}$  line-width of 7.5 Hz at the  $^{13}\text{C}$ -satellites height. Resolution of 6.48 and 7.30 Hz was obtained respectively by the two spectroscopists. The poor quality ( $R^2 = 0.195$ ) and prediction ( $Q^2 = -1.32$ ) factors for this model confirm the impossibility of rejecting the null hypothesis, that is, the inability to distinguish the two NMR users.

tional partial least-squares regression (O2-PLS)<sup>31</sup> score plot. In the following, we explore the dispersion related to several different factors such as age and biological breeding (represented by arrows together with genetics in Figure 1). Notably, we show that many factors such as batch, sample preparation, NMR settings, HRMAS analysis, or NMR facilities do not have a predictive value in the model. We conclude that despite the variability induced by exogenous technological factors (some of them being here amplified on purpose as concerns intrinsic effects such as age), functional genomics of *C. elegans* by  $^1\text{H}$  HRMAS NMR-based metabotyping is robust enough to monitor or cancel out these effects in large-scale studies as long as standards, which we establish in this robust metabotyping method, are respected.

**Monitoring Sample Stability.**  $^1\text{H}$  HRMAS NMR allows the acquisition of resolved spectra from a semisolid sample<sup>32</sup> such as an ensemble of nematodes. This NMR technique involves rotating the sample at relatively high spinning frequencies of a few kilohertz.

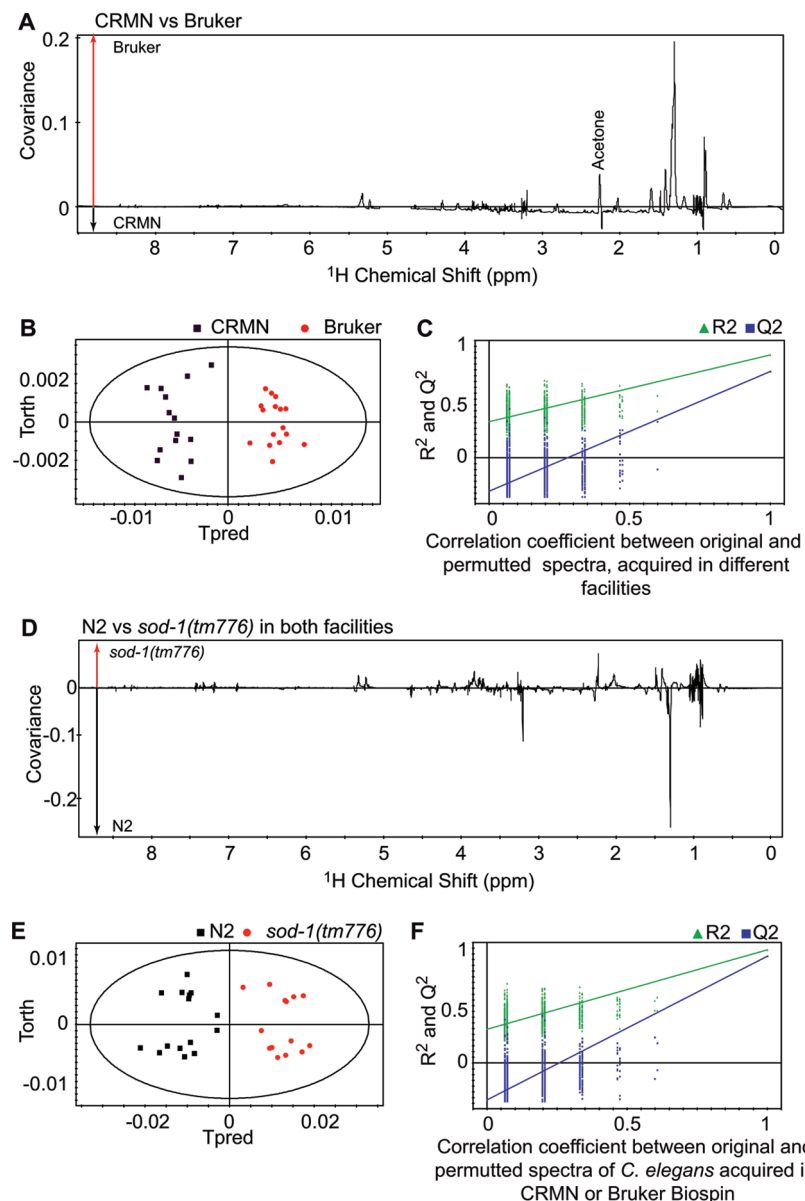
The sample, under magic angle spinning (MAS), may be subject to physical degradation over time leading to bias in the analysis. We acquired 1D spectra of a pellet of fixed *sod-*

*1(tm776)* mutants, filled into a rotor (NMR sample container), at different time intervals to monitor the evolution of the sample under rotation (Figure 2a). We have already shown that an experiment carried out at 293 K over 25 min does not noticeably alter the spectrum, as it is not possible to distinguish by O-PLS two NMR spectra acquired on the same sample.<sup>10</sup> However, large changes are detected after 240 min of sample spinning at 3500 Hz. A regression of the spectral data with respect to time (Figure 2b) results in a variation of almost all NMR signals; moreover, lipids seem to be highly affected. We plotted peak integrals, intensity, and line-width at half-height for six resolved compounds belonging to different metabolite classes (cholesterol, isoleucine,  $\beta$ -glucose, unsaturated lipids, uracil, and inosine) as a function of time in Figure 2c, d, and e. NMR signals were assigned using  $^1\text{H}$  TOCSY and STOCSY spectra (Figure 3a and b). An overall increase (with the exception of lipids in the integral) was observed. The absence of significant change in the line-width indicates a higher concentration of metabolites in the unrestricted liquid phase. This probably results from the destruction of the cellular environment under the effect of MAS and temperature.

In summary, *C. elegans* spectra for a quantitative metabolism analysis should be recorded within one hour of inserting the sample in the NMR instrument, in order to be relevantly used in further statistical analysis. Note that 2D spectra for assignment and identification, where the absolute concentration is not so critical, can be recorded over longer time (16 h for  $^1\text{H}$ - $^1\text{H}$  TOCSY shown in Figure 3a and 40 h for  $^1\text{H}$ - $^{13}\text{C}$  HSQC).

**Sample Preparation.** *C. elegans* metabolic phenotypes reflect the interaction of the genome and the environment.<sup>33</sup> Therefore we investigated the impact of freezing *C. elegans* samples at  $-80^\circ\text{C}$ , as compared to worms pellets prepared on the day of the NMR experiments, in order to decouple the breeding stage from the NMR analysis, thereby improving considerably the flexibility of the protocol.

We compared the metabolic signature between N2 and *sod-1(tm776)* postfixation in two different experimental settings: NMR analysis of non frozen (Protocol A) and frozen nematodes (Protocol B). Half of the *C. elegans* samples were placed in a freezer at  $-80^\circ\text{C}$ , directly after breeding and fixation, and frozen within one hour. They were left approximately two hours at  $-80^\circ\text{C}$  then thawed at room temperature over ten minutes before NMR analysis. NMR acquisition was performed directly after fixation for the other half of the samples. The results of a variance component model presented in Figure 4a show that overall genetics has a slightly greater effect on the spectra than freezing. More importantly, an O-PLS model (Figure 4b and f) shows explicitly the effect of freezing on the complete spectrum. Sample freezing has a strong effect on the spectra, showing an increase in concentration of water-soluble metabolites and a decrease in lipids and fatty acids, which has previously been observed by Middleton and co-workers who characterized the effect of sample freezing in HRMAS NMR spectra of biological tissues.<sup>34</sup> We also observed an additional variation of the chemical shift in the 0.9–1.1 ppm area (Figure 4d). Thus, sample freezing seems to be responsible for a concerted chemical shift variation affecting the amino acid methyl resonances (Figure 4d) and an increase of the intensity of the loadings, as shown in the 3–4 ppm region (Figure 4e). Although the metabolites associated to the discrimination between N2 and *sod-1(tm776)* are not fully identical (Figure 4c), we found a correlation of 0.67 between the two loading plots and the signals in the loadings remain generally the same



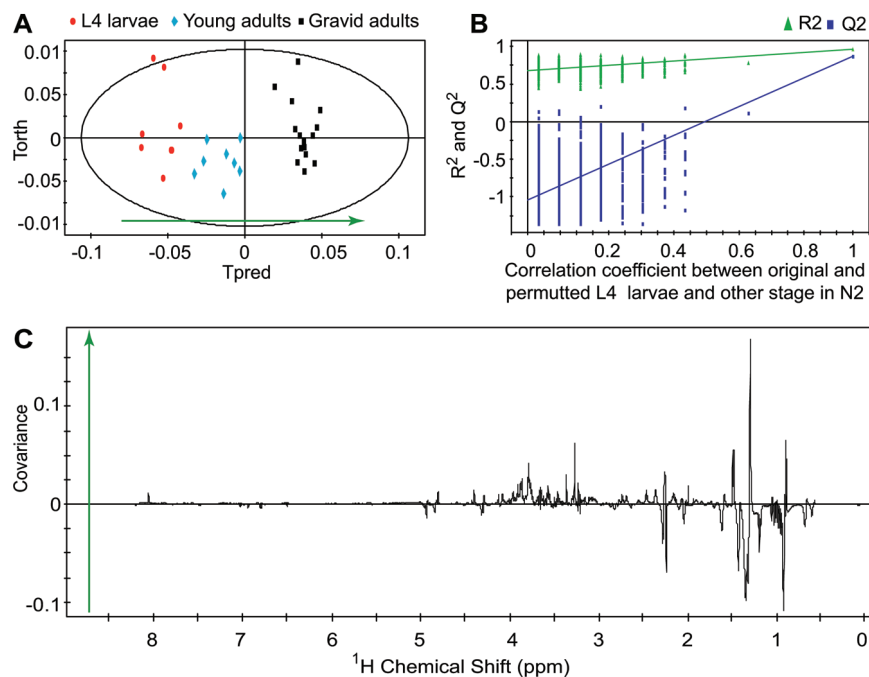
**Figure 6.** NMR center effect. The same experimental design was used to compare N2 and *sod-1* nematodes in different facilities. The (a) O-PLS model loadings, (b) score plots, and (c) quality factors discriminate the set of N2 and *sod-1* spectra on the basis of the NMR facilities (CRMN or Bruker), with a loading plot showing a baseline variation. Model quality coefficients are  $R^2 = 0.926$  and  $Q^2 = 0.874$ . (d–f) O-PLS model discriminating these spectra on the basis of the genetic strain. Model quality coefficients are  $R^2 = 0.876$  and  $Q^2 = 0.748$ . The correlation coefficient between the *sod-1* vs N2 metabolic phenotype shown in (d) and the one established at a single facility (Supplementary Figure 2e, Supporting Information) is 0.93.

(except for the changes in the methyl region discussed above), indicating that the change of the experimental protocol does not affect the results of the biological metabolic analysis. We consider two metabolites to be reasonably similar when the same peaks are identified with a significant covariance (i.e.,  $> 0.02$ ) on both loadings. The complete analysis on the effect of freezing is detailed in Supplementary Figure 1, Supporting Information. Freezing is hence validated as part of an experimental protocol that allows the separation of the two main experimental steps of the method, that is, worm breeding and NMR acquisition, which is a practical necessity for large-scale metabolomics studies for functional genomics of *C. elegans*.

**Operator-Related Effects.** Another type of confounding factor arises from the use of one protocol (even though standardized) for nematode breeding, sample preparation or

NMR acquisition, by different operators sharing their results. We studied nematode paired-breeding and NMR paired-acquisition to assess and quantify these operator-related effects.

Worm breeding variation was thus investigated by the comparison of N2 nematodes bred by two different investigators that filled a total of 16 (8 per investigator) HRMAS rotors. O-PLS analysis reveals a discrimination of the samples based on culture variation (Figure 5a). Metabolic differences between the two cultures affect a large range in the spectrum (data not shown). Validation of the statistical model was performed by resampling the model under the null hypothesis (Figure 5b), showing no random models outperforming the initial one in terms of explanation ( $R^2 = 0.807$ ) or prediction ( $Q^2 = 0.566$ ). Several factors could be at work here. Since the biologist carries out several steps to breed and prepare the sample, the



**Figure 7.** Developmental metabonomics in N2 *C. elegans*. (a) O-PLS model score plot discerning the effect of developmental stage in a population of wild-type N2 worms. (b) Model validation was performed by resampling the model under the null hypothesis ( $R^2 = 0.949$ ;  $Q^2 = 0.849$ ). (c) Corresponding loading plot, that is, metabotype.

experiment can be affected by the time needed to wash the Petri dishes (10 dishes per rotor) and to transfer worms to the tube where they will be fixed (technological effect). Depending on the experimentalist, though following the same standard protocol, this step can be more or less time-consuming, exposing worms to variable levels of stress. This points out that biological preparation is a significant confounding factor that must be taken with great care to allow the extraction of useful information. Randomization in sample preparation is thus essential to minimize this effect.

Similarly, the spectroscopist carries out a succession of steps to obtain an NMR spectrum. The quality of an NMR spectrum depends on many parameters. Among them, the NMR shims set by the operator to achieve high resolution of the proton resonance lines will play an important role, as well as tuning and matching of the probe, or the setup of an efficient water presaturation. A set of 6 wild-type nematode samples was prepared, and NMR spectra were recorded for each sample, in a random order, by two different spectroscopists, setting up independently their NMR acquisition conditions. A standard protocol was followed by both spectroscopists: shims were set on a 10% chloroform sample in acetone-D<sub>6</sub> and adjusted to reach a proton line width at the height of the <sup>13</sup>C-satellites inferior to 7.5 Hz. Respectively, a resolution of 6.48 and 7.30 Hz was obtained by our two operators. Each NMR spectroscopist used his or her own shim matrix and performed final calibrations on each rotor (tuning and matching, calibration of the water presaturation, and additional shims adjustment). An O-PLS model was unable to discriminate populations according to the NMR operator (Figure 5c). Quality of the model in terms of quality  $R^2 = 0.195$  or prediction  $Q^2 = -1.32$  is low, and resampling under the null-hypothesis shows that random models can outperform the capacity of the initial model in terms of quality or prediction. This indicates the impossibility to distinguish two spectroscopists by rejecting the

null hypothesis. An NMR operator, carefully following the standardized acquisition protocol, is thus not a confounding factor.

We also investigated the effect of multiple acquisition platforms in geographically distinct locations. Two sets of N2 and *sod-1(tm776)* samples were prepared and frozen, with one data set acquired at CRMN (Lyon), and the other shipped to the application laboratory of Bruker Biospin in Wissembourg. Experiments were performed on Bruker spectrometers operating at 700 MHz but with different electronic console configurations (respectively Avance I and Avance III models with differences in the electronics and digital filters). We thus compare 2 batches of samples acquired on spectrometers with different electronics in two locations. It is therefore clearly possible here to discriminate samples, under supervised analysis, with respect to the facility where the corresponding NMR spectra were acquired (Figure 6a–c) in a high quality model ( $R^2 = 0.926$  and  $Q^2 = 0.874$ ). However, when orienting the supervised analysis to the parameter of interest in our metabotyping study, that is, the discrimination between mutants and wild type nematodes, Figure 6e shows that, when gathering the two groups of data sets, we can still discriminate the two strains, with a metabotype reasonably similar to the one determined during the previous study carried out at the CRMN (Figure 6d–f,  $R^2 = 0.876$  and  $Q^2 = 0.748$ ). Note that the data acquired at Bruker Biospin (Wissembourg) had a supplementary exclusion area around a singlet at 2.23 ppm corresponding to residual acetone after drying in our washing procedure. This result allows defining an instrument independent metabolic signature (correlation of 0.93 between the metabotypes obtained at CRMN and the one obtained by assembling data acquired at CRMN and Bruker Biospin). We show that it is also possible to detect at both facilities a discrimination between N2 and *sod-1(tm776)* with close observed metabotypes and thus to define the metabolic signature of the mutation (Supplemen-

**Metabolic Profiling Strategy of *Caenorhabditis elegans***

tary Figure 2, Supporting Information). A correlation of 0.62 between the two loadings was determined. The results described above therefore clearly indicates that, though there is a need for the harmonization of spectrometer facilities involved in a multicenter study, a relevant discrimination between given *C. elegans* strains metabolotypes can be achieved when data acquisition is randomly distributed between different NMR centers.

**Developmental Metabonomics in N2 *C. elegans*.** Other metabolic variations of *C. elegans* populations obviously arise from physiological processes. Metabolic changes doubtlessly occur during the growth of the nematodes. In order to probe this factor and investigate the impact of developmental stages on the metabolome, we compare N2 nematodes at three stages: L4 larvae, young adults (adults which do not yet have eggs) and gravid adults. An O-PLS model shows a clear discrimination between these three sample categories and the definition of a metabolic phenotype with respect to developmental stages (Figure 7), with quality and prediction parameters  $R^2 = 0.949$  and  $Q^2 = 0.849$ . We observed no overlap between subpopulations based on the predicted age with respect to the observed one, fully validating the model. Furthermore, the model was resampled under the null hypothesis and validated. We thus clearly isolate an effect linked to the developmental stage in N2, which must be controlled by standardization of the developmental stage for genetic analysis in order to identify relevant metabolic variations induced by mutations.

In summary, the effects studied here can be classed into three groups: (i) insignificant (NMR spectroscopist), (ii) significant but manageable (NMR facility, hardware), and (iii) significant requiring control through experimental design (age, fixation, freezing, biologist) and sample randomization (age, biologist, etc.). The particular effects cited here in each category correspond to our experience with the protocol, through implementation of thorough quality control criteria. In studies with smaller numbers of samples, exogenous biological effects must be carefully controlled by a homogenization of the population in terms of both the developmental stage and by randomization of the sample breeding if different biologists are involved, whereas it appears that in a larger scale study (i.e., with a large number of samples per genetic strain), even the age and biologist can be meaningfully isolated from the genetic effect, but do not hamper the genetic analysis.

**Conclusion**

This study is part of the effort observed in metabonomics for the standardization of procedures (sample culture and preparation, spectral acquisition, data treatment and interpretation, report) to prevent analysis being misled by exogenous factors. This work shows that genetic metabolotypes can be reliably established using the method introduced here without being perturbed by confounding factors. Here we identified and assessed potential confounding factors for  $^1\text{H}$  HRMAS whole-organism NMR and established guidelines to control them. In a large-scale metabolotyping context, respecting these guidelines will certainly improve robustness of the NMR analysis and paves the way for large-scale metabolotyping in functional genomics,<sup>35</sup> chemical genomic, or pharmacometabonomics<sup>6</sup> screening efforts for *C. elegans*. The method is by no means limited to *C. elegans* but could obviously be adapted to other small animal models suitable for whole-organism NMR.

**Acknowledgment.** We thank Bruker Biospin for its financial support and helpful discussions (Drs Hartmut

Schaffer, Manfred Spraul, and Alain Belguise). We declare no competing financial interests. We acknowledge financial support from the Agence Nationale de la Recherche (ANR-07-JCJC-0042-01).

**Supporting Information Available:** Supplementary Figures 1 and 2. This material is available free of charge via the Internet at <http://pubs.acs.org>.

**References**

- (1) Nicholson, J. K.; Lindon, J. C.; Holmes, E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* **1999**, *29* (11), 1181–1189.
- (2) Nicholson, J. K.; Connelly, J.; Lindon, J. C.; Holmes, E. Metabonomics: a platform for studying drug toxicity and gene function. *Nat. Rev. Drug Discov.* **2002**, *1* (2), 153–161.
- (3) Oldiges, M.; Lutz, S.; Pflug, S.; Schroer, K.; Stein, N.; Wiendahl, C. Metabolomics: current state and evolving methodologies and tools. *Appl. Microbiol. Biotechnol.* **2007**, *76* (3), 495–511.
- (4) Fiehn, O.; Kopka, J.; Dormann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L. Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* **2000**, *18* (11), 1157–1161.
- (5) Dumas, M. E.; Wilder, S. P.; Bihoreau, M. T.; Barton, R. H.; Fearnside, J. F.; Argoud, K.; D'Amato, L.; Wallis, R. H.; Blancher, C.; Keun, H. C.; Baunsgaard, D.; Scott, J.; Sidemann, U. G.; Nicholson, J. K.; Gauguier, D. Direct quantitative trait locus mapping of mammalian metabolic phenotypes in diabetic and normoglycemic rat models. *Nat. Genet.* **2007**, *39*, 666–672.
- (6) Clayton, T. A.; Lindon, J. C.; Cloarec, O.; Antti, H.; Charuel, C.; Hanton, G.; Provost, J. P.; Le Net, J. L.; Baker, D.; Walley, R. J.; Everett, J. R.; Nicholson, J. K. Pharmacometabonomic phenotyping and personalized drug treatment. *Nature* **2006**, *440* (7087), 1073–1077.
- (7) Nicholson, J. K.; Holmes, E.; Wilson, I. D. Gut microorganisms, mammalian metabolism and personalized health care. *Nat. Rev. Microbiol.* **2005**, *3* (5), 431–438.
- (8) Harrison, R.; Papp, B.; Pal, C.; Oliver, S. G.; Delneri, D. Plasticity of genetic interactions in metabolic networks of yeast. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (7), 2307–2312.
- (9) Villas-Boas, S. G.; Moon, C. D.; Noel, S.; Hussein, H.; Kelly, W. J.; Cao, M.; Lane, G. A.; Cookson, A. L.; Attwood, G. T. Phenotypic characterization of transposon-inserted mutants of *Clostridium proteoclasticum* B316(T) using extracellular metabolomics. *J. Biotechnol.* **2008**, *134* (1–2), 55–63.
- (10) Blaise, B. J.; Giacomotto, J.; Elena, B.; Dumas, M. E.; Toullhoat, P.; Segalat, L.; Emsley, L. Metabolotyping of *Caenorhabditis elegans* reveals latent phenotypes. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 19808–19812.
- (11) Kameh, M. A.; Hobani, Y.; Dow, J. A. T.; Watson, D. G. Metabolomic profiling of *Drosophila* using liquid chromatography Fourier transform mass spectrometry. *FEBS Lett.* **2008**, *582* (19), 2916–2922.
- (12) Johnson, T. E. *Caenorhabditis elegans* 2007: The premier model for the study of aging. *Exp. Gerontol.* **2008**, *43* (1), 1–4.
- (13) Atherton, H. J.; Jones, O. A. H.; Malik, S.; Miska, E. A.; Griffin, J. L. A comparative metabolomic study of NHR-49 in *Caenorhabditis elegans* and PPAR- $\alpha$  in the mouse. *FEBS Lett.* **2008**, *582* (12), 1661–6.
- (14) Tampakakis, E.; Okoli, I.; Mylonakis, E. A. *C. elegans*-based, whole animal, in vivo screen for the identification of antifungal compounds. *Nat. Protoc.* **2008**, *3* (12), 1925–31.
- (15) Leung, M. C. K.; Williams, P. L.; Benedetto, A.; Au, C.; Helmcke, K. J.; Aschner, M.; Meyer, J. N. *Caenorhabditis elegans*: An emerging model in biomedical and environmental toxicology. *Toxicol. Sci.* **2008**, *106* (1), 5–28.
- (16) Silverman, G. A.; Luke, C. J.; Bhatia, S. R.; Long, O. S.; Vetica, A. C.; Perlmutter, D. H.; Pak, S. C. Modeling Molecular and Cellular Aspects of Human Disease Using the Nematode *Caenorhabditis elegans*. *Pediatr. Res.* **2009**, *65* (1), 10–18.
- (17) Wold, S.; Esbensen, K.; Geladi, P. Principal Component Analysis. *Chemometrics Intellig. Lab. Syst.* **1987**, *2* (1–3), 37–52.
- (18) Trygg, J.; Wold, S. Orthogonal projections to latent structures (O-PLS). *J. Chemometrics* **2002**, *16* (3), 119–128.
- (19) Teahan, O.; Gamble, S.; Holmes, E.; Waxman, J.; Nicholson, J. K.; Bevan, C.; Keun, H. C. Impact of analytical bias in metabonomic studies of human blood serum and plasma. *Anal. Chem.* **2006**, *78* (13), 4307–4318.

- (20) Fiehn, O.; Robertson, D.; Griffin, J.; van der Werf, M.; Nikolau, B.; Morrison, N.; Sumner, L. W.; Goodacre, R.; Hardy, N. W.; Taylor, C.; Fostel, J.; Kristal, B.; Kaddurah-Daouk, R.; Mendes, P.; van Ommen, B.; Lindon, J. C.; Sansone, S. A. The metabolomics standards initiative (MSI). *Metabolomics* **2007**, *3*, 175–178.
- (21) Lindon, J. C.; Nicholson, J. K.; Holmes, E.; Keun, H. C.; Craig, A.; Pearce, J. T. M.; Bruce, S. J.; Hardy, N.; Sansone, S. A.; Antti, H.; Jonsson, P.; Daykin, C.; Navarange, M.; Beger, R. D.; Verheij, E. R.; Amberg, A.; Baunsgaard, D.; Cantor, G. H.; Lehman-McKeeman, L.; Earll, M.; Wold, S.; Johansson, E.; Haselden, J. N.; Kramer, K.; Thomas, C.; Lindberg, J.; Schuppe-Koistinen, I.; Wilson, I. D.; Reilly, M. D.; Robertson, D. G.; Senn, H.; Krotzky, A.; Kochhar, S.; Powell, J.; van der Ouderaa, F.; Plumb, R.; Schaefer, H.; Spraul, M. Standard Metab Reporting Struct, w., Summary recommendations for standardization and reporting of metabolic analyses. *Nat. Biotechnol.* **2005**, *23* (7), 833–838.
- (22) Griffin, J. L.; Nicholls, A. W.; Daykin, C. A.; Heald, S.; Keun, H. C.; Schuppe-Koistinen, I.; Griffiths, J. R.; Cheng, L. L.; Rocca-Serra, P.; Rubtsov, D. V.; Robertson, D. Standard reporting requirements for biological samples in metabolomics experiments: mammalian/ in vivo experiments. *Metabolomics* **2007**, *3*, 179–188.
- (23) Beckonert, O.; Keun, H. C.; Ebbels, T. M. D.; Bundy, J.; Holmes, E.; Lindon, J. C.; Nicholson, J. K. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protocols* **2007**, *2*, 2692–2703.
- (24) Broadhurst, D. I.; Kell, D. B. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* **2006**, *2* (4), 171–196.
- (25) Falk, M. J.; Zhang, Z.; Rosenjack, J. R.; Nissim, I.; Daikhin, E.; Sedensky, M. M.; Yudkoff, M.; Morgan, P. G., Jr. Metabolic pathway profiling of mitochondrial respiratory chain mutants in *C. elegans*. *Mol. Genet. Metab.* **2008**, *93* (4), 388–397.
- (26) Kell, D. B.; Brown, M.; Davey, H. M.; Dunn, W. B.; Spasic, I.; Oliver, S. G. Metabolic footprinting and systems biology: The medium is the message. *Nat. Rev. Microbiol.* **2005**, *3* (7), 557–565.
- (27) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. Investigation of Exchange Processes by 2-Dimensional NMR-Spectroscopy. *J. Chem. Phys.* **1979**, *71* (11), 4546–4553.
- (28) Zhang, F. L.; Bruschweiler, R. Indirect covariance NMR spectroscopy. *J. Am. Chem. Soc.* **2004**, *126* (41), 13180–13181.
- (29) Cloarec, O.; Dumas, M. E.; Craig, A.; Barton, R. H.; Trygg, J.; Hudson, J.; Blancher, C.; Gauguier, D.; Lindon, J. C.; Holmes, E.; Nicholson, J. Statistical total correlation spectroscopy: An exploratory approach for latent biomarker identification from metabolic H-1 NMR data sets. *Anal. Chem.* **2005**, *77* (5), 1282–1289.
- (30) Nicholson, J. K.; Foxall, P. J. D.; Spraul, M.; Farrant, R. D.; Lindon, J. C. 750-Mhz H-1 and H-1-C-13 NMR-Spectroscopy of Human Blood-Plasma. *Anal. Chem.* **1995**, *67* (5), 793–811.
- (31) Trygg, J. O2-PLS for qualitative and quantitative analysis in multivariate calibration. *J. Chemometrics* **2002**, *16* (6), 283–293.
- (32) Fitch, W. L.; Detre, G.; Holmes, C. P.; Shoolery, J. N.; Keifer, P. A. High-Resolution H-1-NMR in Solid-Phase Organic-Synthesis. *J. Org. Chem.* **1994**, *59* (26), 7955–7956.
- (33) Bollard, M. E.; Stanley, E. G.; Lindon, J. C.; Nicholson, J. K.; Holmes, E. NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition. *NMR Biomed.* **2005**, *18* (3), 143–162.
- (34) Middleton, D. A.; Bradley, D. P.; Connor, S. C.; Mullins, P. G.; Reid, D. G. The effect of sample freezing on proton magic-angle spinning NMR spectra of biological tissue. *Magn. Reson. Med.* **1998**, *40* (1), 166–169.
- (35) Granger, L.; Martin, E.; Segalat, L. Mos as a tool for genome-wide insertional mutagenesis in *Caenorhabditis elegans*: results of a pilot study. *Nucleic Acids Res.* **2004**, *32* (14), e117.

PR900012D