J Biomol NMR (2011) 49:245–254
DOI 10.1007/s10858-011-9492-6

ARTICLE

^{1}H NMR-based metabolic profiling reveals inherent biological variation in yeast and nematode model systems

Samuel S. W. Szeto · Stacey N. Reinke · Bernard D. Lemire

Received: 17 December 2010 / Accepted: 1 February 2011 / Published online: 25 February 2011

Abstract The application of metabolomics to human and animal model systems is poised to provide great insight into our understanding of disease etiology and the metabolic changes that are associated with these conditions. However, metabolomic studies have also revealed that there is significant, inherent biological variation in human samples and even in samples from animal model systems where the animals are housed under carefully controlled conditions. This inherent biological variability is an important consideration for all metabolomics analyses. In this study, we examined the biological variation in ^{1}H NMR-based metabolic profiling of two model systems, the yeast *Saccharomyces cerevisiae* and the nematode *Caenorhabditis elegans*. Using relative standard deviations (RSD) as a measure of variability, our results reveal that both model systems have significant amounts of biological variation. The *C. elegans* metabolome possesses greater metabolic variance with average RSD values of 29 and 39%, depending on the food source that was used. The *S. cerevisiae* exometabolome RSD values ranged from 8% to 12% for the four strains examined. We also determined whether biological variation occurs between pairs of phenotypically identical yeast strains. Multivariate statistical analysis allowed us to discriminate between pair members based on their metabolic phenotypes. Our results highlight the variability of the metabolome that exists even for less complex model systems cultured under defined conditions. We also highlight the efficacy of metabolic profiling for defining these subtle metabolic alterations.

Keywords Biological variance · Model systems · Metabolic profiling

Samuel S. W. Szeto and Stacey N. Reinke have equally contributed to this work.

S. S. W. Szeto · S. N. Reinke · B. D. Lemire (✉)
Department of Biochemistry, School of Molecular and Systems Medicine, University of Alberta, Edmonton, AB T6G 2H7, Canada
e-mail: bernard.lemire@ualberta.ca

Introduction

Metabolomics is the comprehensive analysis of the metabolites in a biological sample. The metabolic phenotype or metabotype generated from metabolite profiling provides a read-out of the cell's metabolic state. The metabotype is the product of genetic and environmental contributions under a particular set of conditions (Holmes et al. 2008; Nicholson 2006). The metabolome is context-dependent and changes in metabolite concentrations reflect changes in the physiological, developmental or pathological state of the cell, tissue or organism (Raamsdonk et al. 2001). Metabolites are the ultimate end products of gene expression and changes in their levels are thought to be the most sensitive and discriminatory measure of the cellular state (Holmes et al. 2008; Kell et al. 2005). Therefore, metabolic profiling offers insight into how changes in metabolites are related to phenotype (Mapelli et al. 2008). Metabolomics can be used to examine the intricate connection between mutation, metabolism and disease. Several metabolites that can serve as useful diagnostic biomarkers of pathogenesis in humans have been identified (MacIntyre et al. 2010; Lanza et al. 2010; Jansson et al. 2009; Bogdanov et al. 2008; Zira et al. 2010; Carrola et al. 2010). In addition, metabolomics facilitates the generation of specific testable biochemical hypotheses and opens new research directions.

With the application of metabolomics to the study of human disease, it has become increasingly apparent that a

number of intrinsic and extrinsic factors can affect the human metabolome. Factors that contribute to biological variation include, but are not limited to age, genetics, gut microflora, diet, lifestyle and environmental factors (Holmes et al. 2008; Walsh et al. 2008; Crews et al. 2009; Lenz et al. 2003; Solanky et al. 2003; Saude et al. 2007). Significant inter-individual variability of metabolic profiles can complicate the interpretation of data. In a recent study, it was determined that human cerebrospinal fluid and plasma show significant biological variation, with median relative standard deviations (RSD, also referred to as coefficients of variation) of 35 and 46%, respectively (Crews et al. 2009). Urine, a commonly used biofluid in metabolomic studies, is even more variable with RSD values of some metabolites being greater than 200% (Saude et al. 2007).

Metabolomic studies have been applied to a variety of non-human model systems in an effort to minimize variation caused by extrinsic factors. In model system studies, environmental, behavioral and genetic variation can be minimized or controlled (Crews et al. 2009). In one study, urine metabolic profiles of animals housed in controlled conditions and consuming the same diet showed a smaller degree of biological variation when compared to humans (Bollard et al. 2005). However, in another study on guinea pigs, the profiles showed similar biological variability for some metabolites but greater variability for others when compared to a human control population (Saude et al. 2007). These data demonstrate that animal model systems will inherently display a significant degree of biological variation and this has the potential of obscuring metabolic differences between wild-type and disease states. Biological variation must be carefully considered before any conclusions are drawn. Metabolic analysis of other model organisms, such as *Arabidopsis thaliana* reveals significant biological variation in metabolite concentrations, with an average RSD of ~40% (Fiehn et al. 2000). In a recent study, Parsons et al. determined the RSD values for a number of metabolomic datasets, spanning a variety of sample types from mammals, fish, invertebrates and a cultured cell line. Inter-individual biological variation ranged from a RSD of 7.2% for rat tissue extracts to 58.4% for fish plasma (Parsons et al. 2009). In several of the previously mentioned studies, analytical reproducibility was also examined as a potential source of the variability and shown not to be a major contributing factor (Parsons et al. 2009; Fiehn et al. 2000; Crews et al. 2009).

We have previously used simple model systems, such as the yeast *Saccharomyces cerevisiae* and the nematode *Caenorhabditis elegans* for metabolomic studies focusing on understanding mitochondrial diseases, particularly defects of the mitochondrial respiratory chain (Reinke et al. 2010; Szeto et al. 2007, 2010). In *C. elegans*, we investigated how the metabolome was affected by complex I (NADH:ubiquinone oxidoreductase) dysfunction. In *S. cerevisiae*, we examined how the extracellular metabolome or exometabolome was affected by succinate dehydrogenase (SDH, complex II) dysfunction. Changes in the exometabolome have been shown to be a direct reflection of intracellular metabolic activity (Mapelli et al. 2008; Shaham et al. 2010; Kell et al. 2005; Allen et al. 2003; Dowlatabadi et al. 2009). Monitoring the exometabolome is also efficient, noninvasive and not subject to the technical challenges related to isolating intracellular metabolites (Kell et al. 2005; Allen et al. 2003). In light of the reports of biological variation seen in other model systems, we examined the variability present in our yeast and worm model system studies. In this study, we determined the RSD values for the metabolites characterized and report that an appreciable amount of biological variation exists. This biological variation is also apparent when examining genetically comparable yeast strains. Using multivariate statistical analysis, we are able to discriminate between the metabotypes of two SDH knockout strains and the two knockouts complemented with a plasmid carrying wild-type SDH genes; yet, these strains form two phenotypically identical pairs. Our results demonstrate that even "simple" model systems cultured under controlled conditions exhibit significant biological variation in their metabolomes.

Experimental procedures

Worm strains and culture conditions

The *C. elegans* N2 (Bristol) wild-type strain was used. Worms were cultured on either nematode growth media (NGM) plates seeded with *E. coli* OP50 or on NGM containing 1 mM IPTG and 25 μg ml^{-1} carbenicillin and seeded with *E. coli* HT115 (DE3) transformed with the L4440 vector (Lewis and Fleming 1995; Fraser et al. 2000). Cultures were inoculated with synchronized L1 stage worms and maintained at 20°C as described (Reinke et al. 2010). Worms were harvested, incubated in sterile water for 30 min to allow digestion of gut bacteria and washed three times (Grad et al. 2007). Worm pellets were suspended in 900 μl of sterile water and three 10 μl aliquots were removed to quantify total protein. Seven biological replicates were collected for worms fed *E. coli* OP50 and six were collected for worms fed *E. coli* HT115 L4440.

Yeast strains, media and culture conditions

The *S. cerevisiae* knockout strains sdh3W1 (MH125, *sdh3::TRP1*) and sdh4W2 (MH125, *sdh4::TRP1*) have been described previously (Oyedotun and Lemire 1997, 1999; Szeto et al. 2007). The yeast knockout strains were

transformed with either empty low-copy plasmids or plasmids containing the corresponding wild-type *SDH3* or *SDH4* genes by lithium acetate-mediated transformation (Gietz et al. 1992). Strains were plated on solid SD medium containing casamino acids (0.5% w/v) without tryptophan for plasmid retention. Cells were grown for 36 h at 30°C in 2 ml of YP medium containing 0.25% glucose. This is a modified version of a medium previously used to examine acid secretion (Szeto et al. 2007; de Kok et al. 1975). To eliminate potential variations in the metabolite profiles arising from the growth medium composition, a single batch of medium was used to culture all strains in this study. Following completion of the incubation period, the optical densities (OD) of the cultures at 600 nm were determined. Cultures were centrifuged to pellet cells and the clarified media were transferred to microcentrifuge tubes as described (Szeto et al. 2010). Twelve biological replicates were collected for each strain.

Metabolite extraction

Worms were lysed on ice in trichloroacetic acid (TCA; 5% final concentration) by sonication. Clarified post growth yeast media were also treated with TCA (5% final concentration) and incubated on ice for 30 min to precipitate protein. Both worm and yeast samples were centrifuged at 14,000×*g* for 15 min at 4°C. Supernatants were recovered and pH neutralized with 5 M NaOH. Protein-free lysates were re-clarified by centrifugation. All samples were flash frozen in liquid nitrogen before lyophilization.

Sample preparation for NMR spectroscopy

Dry protein-free lysates were dissolved in 570 µl D_2O (99.9%; Isotec Inc., Miamisburg, OH). 30 µl of 5 mM 2,2-dimethyl-2-sila 3,3,4,4,5,5-hexadeutero-pentane sulphonic acid (DSS-d_6, Chenomx Inc., Edmonton, AB) were added as a chemical shift indicator and concentration standard for NMR analysis. The pH was recorded for calibration purposes and samples were centrifuged at 14,000×*g* for 3 min to remove particulate matter. 510 µl of supernatant were transferred to 5 mm diameter NMR tubes for data collection.

^{1}H NMR spectroscopy and NMR data processing

One-dimensional ^{1}H NMR spectra were acquired on a 600 MHz Varian Inova spectrometer (Varian Inc., Palo Alto, CA) at 30°C using a tnnoesy pulse sequence (circa Vnmr 6.1B software, Varian Inc.). Parameters were consistent with those previously described (Szeto et al. 2007, 2010; Reinke et al. 2010). Chenomx NMR Suite Professional software v5.1 (Chenomx Inc., Edmonton, AB) was used for metabolite identification and quantification. This software uses pattern recognition and lineshape deconvolution to fit spectra based on highly specific peak patterns (Weljie et al. 2006). The resonance linewidths are input from the reference standard, DSS. The spectral patterns of many metabolites often contain more than one peak throughout the spectrum, forming complex and distinctive sets of peaks. In addition, the spectral baseline often varies across regions but fitting can be achieved by using only the top portion of any peak as long as overall fit is consistent across all regions. As such, baseline inconsistencies do not affect quantification accuracy. For additional consistency, the same sets of peaks were always utilized for quantification of any one metabolite.

Metabolite data analysis

Worm metabolic data were normalized to protein contents. Yeast metabolic data were normalized to the mass of total dried protein-free lysate and to OD of the culture. Relative standard deviation (standard deviation/mean × 100%) was determined for each metabolite in each sample set. Box plots were generated as previously described (Cann 2003; Massart et al. 2005). The Mann–Whitney test was used to analyze whether the two worm metabolite RSD distributions differed significantly (Corder and Foreman 2009). Multivariate statistical analyses were performed using Simca P + v12.0.1 software (Umetrics, Umeå, Sweden) (Eriksson et al. 2001). Data were not transformed but were scaled to unit variance, dividing each variable by its standard deviation, and mean centered to provide equal importance to all variables. Data were visualized using both the unsupervised principal component analysis (PCA) and the supervised partial least square discriminate analysis (PLS-DA). R^2 and Q^2 values were used to assess the fit and predictability of the respective models. To examine the possibility of spurious model generation, validations of each group within each model were performed. The software randomly generated 999 permutations of y-variables (metabolites), while keeping x-variables (strain) intact, and facilitated a comparison of the measures of fit (R^2 and Q^2) from the generated permutations to those of the original. Statistical significance of the differences between groups in each model was determined using CV-ANOVA (analysis of variance testing of cross-validated predictive residuals).

Results

Biological variation observed in the worm and yeast model systems

We determined the RSD values for each metabolite to quantify spectrum-wide variability in the metabolomes

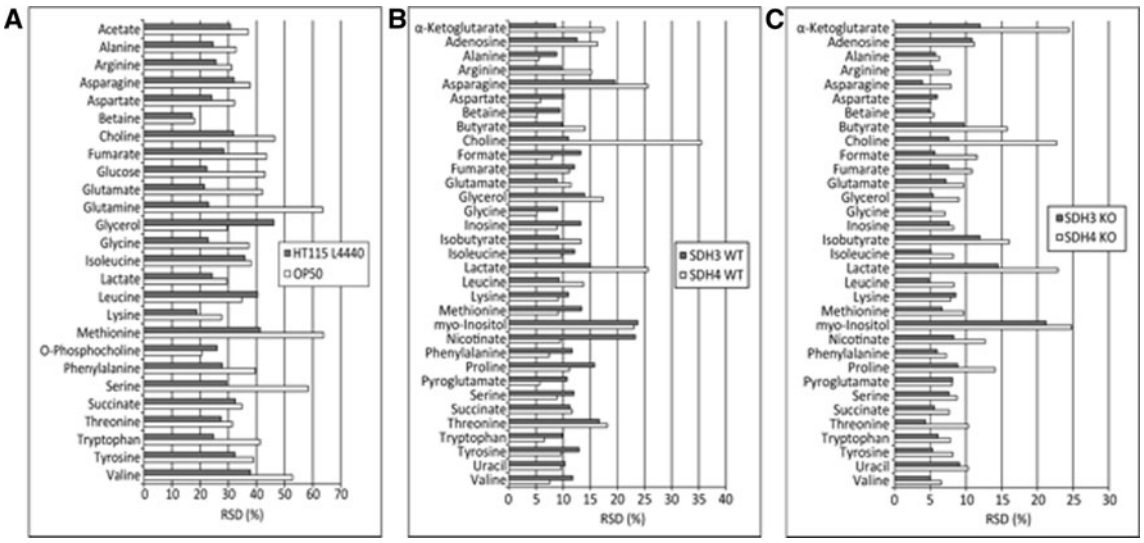

Fig. 1 RSD values determined for the metabolites identified in **a**, wild-type, whole worm, protein-free lysates from worms fed *E. coli* HT115 L4440, *grey* ($N = 6$) or fed *E. coli* OP50, *white* ($N = 7$); **b** exometabolome of the complemented *SDH3*, *grey* ($N = 12$) and *SDH4*, *white* ($N = 12$) yeast strains; **c** exometabolome of knockout *SDH3*, *grey* ($N = 12$) and *SDH4*, *white* ($N = 12$) yeast strains

Table 1 Summary of the metabolomic datasets used in this study

Metabolite Dataset	Metabolome type	No. of metabolites quantified	Average % RSD	Range of % RSD
C. elegans on OP50	Intracellular	26	39	18–64
C. elegans on HT115 L4440	Intracellular	26	29	17–46
S. cerevisiae SDH3 WT	Extracellular	33	12	9–24
S. cerevisiae SDH4 WT	Extracellular	33	12	5–36
S. cerevisiae SDH3 KO	Extracellular	33	8	4–21
S. cerevisiae SDH4 KO	Extracellular	33	11	5–25

(Fig. 1; Table 1) of whole worm, protein-free lysates and protein-free, exometabolome lysates of yeast. RSD values demonstrate reproducibility in large data sets. They can also offer a frame of reference when manipulating or developing new preparation methods for biological samples (Parsons et al. 2009). *C. elegans* are routinely cultured on NGM plates seeded with *E. coli* OP50. This strain is an uracil auxotroph and produces a thin lawn on NGM plates, allowing for easier visualization of the worms (Brenner 1974). Alternatively, another laboratory *E. coli* strain HT115 (DE3) is used for feeding-mediated gene suppression by RNA-interference (RNAi) (Kamath et al. 2001). HT115 is transformed with the empty vector L4440 as an RNAi control. Using six samples of wild-type worms fed *E. coli* HT115 L4440, variance in whole worm, protein-free lysates ranged from 17 to 46%, with an average RSD of 29%. Three metabolites glycerol, leucine and methionine had a coefficient of variance greater than 40%. Using seven samples, variance in worm lysates from wild-type worms fed *E. coli* OP50 ranged from 18 to 64%, with an average RSD of 39%. Nine metabolites had a coefficient of variance greater than 40%: choline, fumarate, glucose, glutamate, glutamine, methionine, serine, tryptophan and valine. The RSD distributions of these two metabolomes were compared and are significantly different ($p < 0.001$). The average RSD of the yeast exometabolomes were 12% for the complemented strains and 11% for the knockout strains. Although the average RSD values were quite similar, there was greater variance in the replicates of the complemented *SDH4* strain, exhibiting a range of 5–35%. The most variable metabolites in the complemented *SDH3* exometabolome were asparagine, myo-inositol and nicotinate. Choline, asparagine, lactate and myo-inositol were the most variable exometabolites of the complemented *SDH4* metabolome. Choline was the most variable metabolite, with an RSD greater than 35%. Myo-inositol was the only metabolite in the *SDH3* knockout strain that varied by more than 20% of the mean. However, α-ketoglutarate, choline, lactate and myo-inositol all exceeded 20% variation in the *SDH4* knockout strain. In addition, we did not observe any correlation between the RSD and the concentration of metabolites

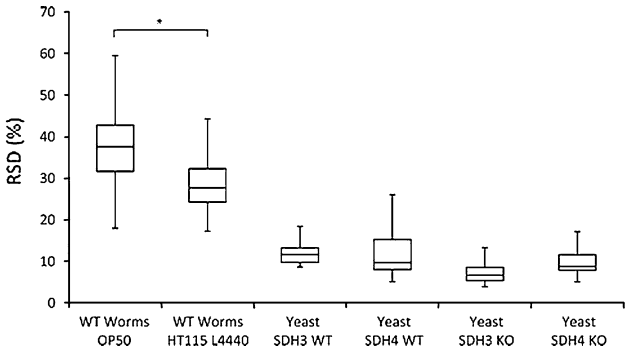

Fig. 2 Box plot of RSD values for metabolomic datasets described in this study. Values for lower quartile, median and upper quartile are shown. *Error bars* show the range of data. $*p < 0.001$

quantified. The RSD values were not significantly higher for metabolites that had either very small (<10 μM) or very large (>10 mM) concentrations (data not shown). These data are also represented as box plots shown in Fig. 2. The box plot yields a simple visual representation of the RSD data sets. The lower, middle and upper lines of the boxes represent the lower quartile, median and upper quartile of each RSD data set, respectively. The error bars show the range of the data. The yeast exometabolomes were much less variable than the whole-worm metabolome. Variance in the wild-type whole-worm metabolome was altered by food source.

Multivariate analysis of yeast exometabolomes

We employed multivariate analysis to establish metabotype differences between previously characterized yeast strains. The metabotypes of two knockout strains (*SDH3* KO and *SDH4* KO) were compared to each other. This was also done with the respective complemented strains (knockout strains with plasmid-borne copies of the wild-type *SDH3* and *SDH4* genes, respectively). Partial least square discriminant analysis (PLS-DA) provided a supervised evaluation of the variation between the two data sets examined (Fig. 3). Each colour-coded dot represents a replicate metabotype in the representative PLS-DA score plots. The PLS-DA models indicated that the complemented strains could be discriminated from each other (Fig. 3a). A similar result was also obtained for the two knockout strains examined (Fig. 3b). PLS-DA yields an opportunity for statistical analysis of the models. The PLS-DA model comparing the two complemented strains resolved two components with R^2X, R^2Y, and Q^2 values of 0.688, 0.872 and 0.8, respectively. The PLS-DA model comparing the two knockout stains resolved three components with R^2X, R^2Y, and Q^2 values of 0.568, 0.98 and 0.913, respectively. CV-ANOVA tests were also performed on each model.

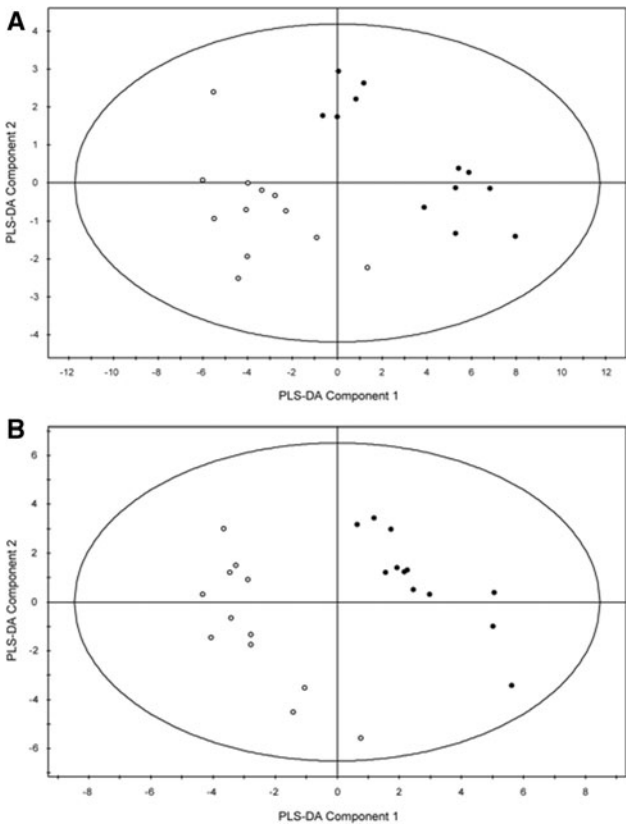

Fig. 3 Two component PLS-DA models of ^{1}H NMR derived exometabolome profiles of *S. cerevisiae* strains. PLS-DA score scatter plots and validations were performed using SIMCA P + v12.0.1 software. A, PLS-DA score plot for the complemented yeast strains. *Black dots*, *SDH3*; *white dots*, *SDH4*. The model is characterized by the following parameters using 2 components: R^2X_{cum}, 0.688, R^2Y_{cum}, 0.872, Q^2_{cum}, 0.8. $N = 12$ for both strains, CV-ANOVA $p = 1.69 \times 10^{-6}$. B, PLS-DA score plot for the yeast deletion strains. *Black dots SDH3* KO; *white dots SDH4* KO. The model is characterized by the following parameters using 3 components: R^2X_{cum}, 0.568, R^2Y_{cum}, 0.98, Q^2_{cum}, 0.913. $N = 12$ for both strains, CV-ANOVA $p = 5.67 \times 10^{-8}$. Each dot represents one yeast culture

These resulted in the values $p = 1.69 \times 10^{-6}$ for the complemented strain model and $p = 5.67 \times 10^{-8}$ for the knockout strain model. Validation plots were also generated for each strain within each model from 999 random permutations of the data using all components (Fig. 4). Two criteria indicate model validity. First, none of the permutations out-performed the original data set. The points representing the original data (the blue and green dots at the far right of the graph) are higher than the permutated data (all dots to the left of the original). Second, the Y-intercepts for the R^2 regression line (green) and the Q^2 regression line (blue) should be less than 0.4 and 0.05, respectively. Our data fit these criteria with the exception that the R^2Y-intercepts for the model comparing the knockout strains are higher than 0.4. However, when the

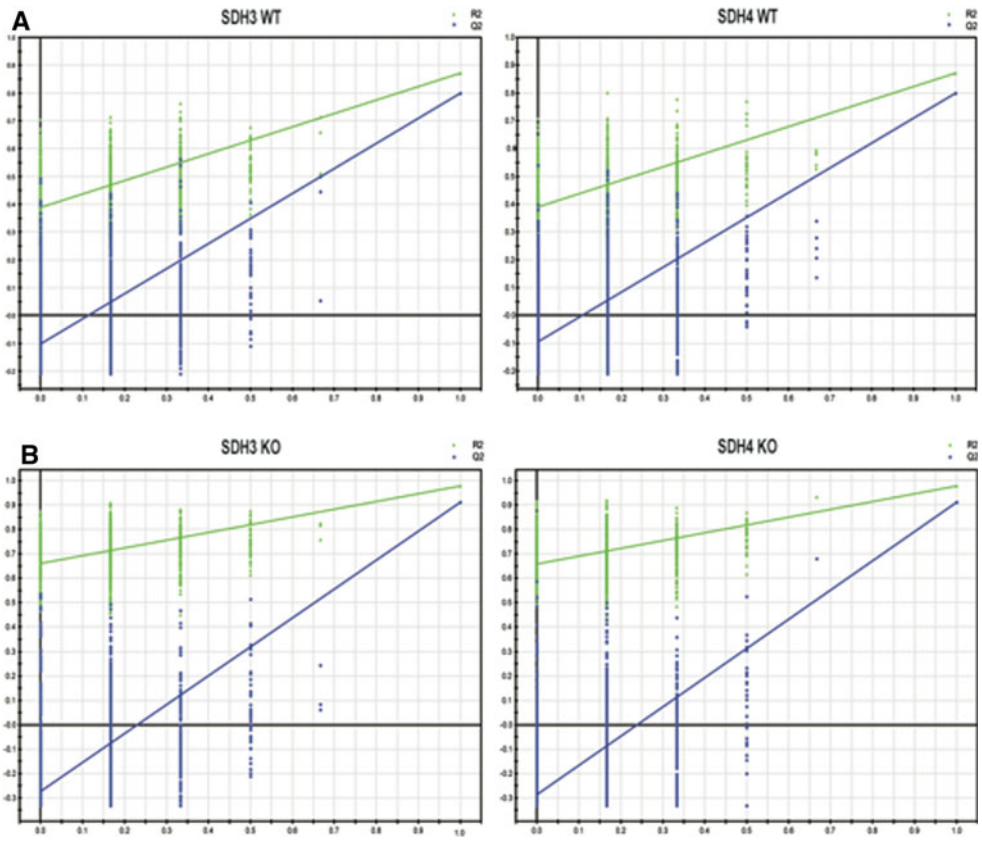

Fig. 4 Validation plots for the complemented and knockout strain models. **a** Validation model using 999 permutations across 2 components. **b** Validation model using 999 permutations across 3 components. The regression line represents the correlation coefficient between the original and permuted *Y*-variables against cumulative R^2 and Q^2. *Green*, R2; *blue*, Q2

high R^2 and Q^2 values for both models, the visual separation seen between strains in each model, the low CV-ANOVA *p* values and the results of the validation permutations are considered, we suggest that the models are valid. To ensure that supervised evaluation did not falsely differentiate between the various yeast strains, principal component analysis (PCA) was also performed (Fig. 5). PCA is unsupervised and treats each observation equally. Again, each colour-coded dot represents a replicate metabotype. In both models, visual separation between the strains in the score plots is evident. The PCA model comparing the two complemented strains resolved two components with R^2X and Q^2 values of 0.694 and 0.516, respectively. The model comparing the two knockout strains resolved three components with R^2X and Q^2 values of 0.608 and 0.314, respectively.

Discussion

The metabolome is an extremely sensitive monitor of physiological state. As such, metabolomic studies of human biofluids can show immense variability in the concentration of an individual metabolite. Concentrations of metabolites in human urine can vary by more than 200% of the mean (Saude et al. 2007). Model systems are highly conducive to metabolomic studies because genetic background and growth conditions can be controlled. The metabotypes of model systems under such conditions should be less variable than those of their human counterparts. We show that in genetically identical animals grown under carefully controlled conditions, individual metabolite concentrations in *C. elegans* can vary by as much as 64% (Figs. 1, 2; Table 1). The less complex unicellular model system *S. cerevisiae* exhibited decreased biological variation in its exometabolomes, with a maximum RSD of 36%. However, the range of variation between the complemented strains differed. Whereas the complemented *SDH3* strain had a RSD range of 15%, the complemented *SDH4* strain had a RSD range of 31%. The yeast knockout strains were more similar in their RSD ranges: 17% for *SDH3* KO and 20% for *SDH4* KO, respectively. The observation that substantial biological variations exists in organisms such as yeast may not be a surprising result considering clonal populations of *E. coli* have also been shown to exhibit substantial phenotypic variation (Elowitz et al. 2002).

The greater variance in *C. elegans* than in *S. cerevisiae* may be attributed to four factors. First, *C. elegans* are social feeders. Despite adequate food and space for the whole population, local subpopulations may experience food limitation (de Bono and Bargmann 1998). Second,

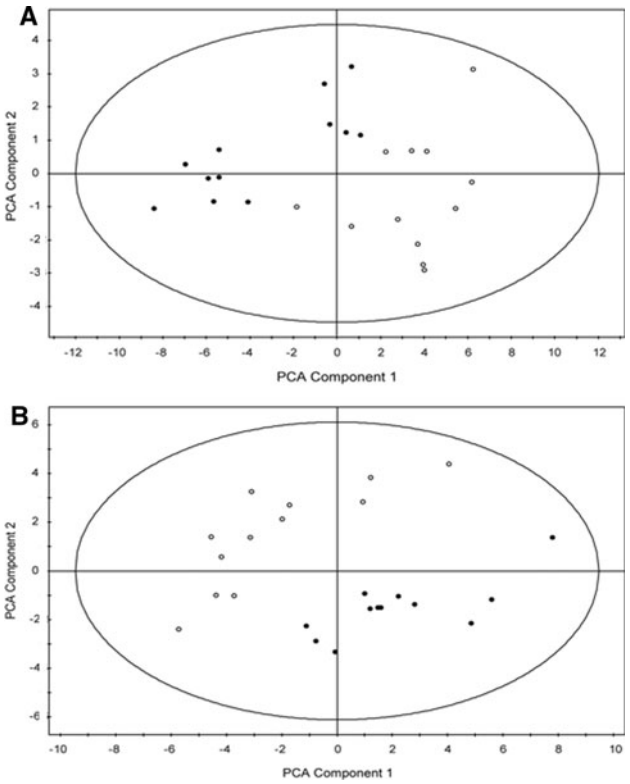

Fig. 5 Two component PCA models of ^{1}H NMR derived exometabolome profiles of *S. cerevisiae* strains. PCA score scatter plots were performed using SIMCA P + v12.0.1 software. A, PCA score plot for the complemented yeast strains. *Black dots*, *SDH3*; *white dots*, *SDH4*. The model is characterized by the following parameters using 2 components: R^2X_{cum}, 0.694, Q^2_{cum}, 0.516. $N = 12$ for both strains. B, PCA score plot for the yeast deletion strains. *Black dots*, *SDH3* KO; *white dots*, *SDH4* KO. The model is characterized by the following parameters using 3 components: R^2X_{cum}, 0.608, Q^2_{cum}, 0.314. $N = 12$ for both strains. Each dot represents one yeast culture

C. elegans develop through four larval stages before entering adulthood. Metabolic differences among the different stages have been documented (O'Riordan and Burnell 1989, 1990; Bart et al. 2009). The distribution of larval stages will vary somewhat between cultures. Third, the worms were cultured on NGM plates seeded and grown overnight with *E. coli*. Chemical composition will vary between plates and *E. coli* lawns. Additionally, several plates of worms are needed to obtain a sufficient number of worms on which to perform metabolic studies. For logistical reasons, we did not evaluate the metabolic composition of either NGM media or of the *E. coli*. Fourth, intracellular metabolites were measured in worms and suboptimal worm lysis could also contribute to variation between cultures. However, we have optimized the sonication protocol to minimize incomplete lysis. The metabolic complexity of *C. elegans* and the inherently more variable culture conditions it demands likely account for most of the sample variation we detected. Yeast also exhibits an appreciable amount of biological variation despite being grown under well-defined and controlled conditions. As all experiments were carried out with great care and consistency, we proffer that the variation in yeast samples may be primarily due to stochastic gene expression (Raser and O'Shea 2005).

Technical or analytical variation is an important consideration in metabolomics studies and can be a significant contributing factor to the observed metabolic variation. Sources of this variation include instrumentation, sample preparation and NMR data analysis. In our studies, we examined the contributions of technical variation to the observed variation using two approaches. First, we manually fitted and quantified the DSS concentration reference peak in all of the yeast exometabolite samples. DSS is added at a constant concentration to all samples and variations in the quantified values are a reflection of the precision of the NMR spectrometer. Since these spectra were collected at various times over a number of days, variations in DSS peak values may reflect subtle changes in the instrument that occur during operation. The average RSD value for the DSS peaks from the spectral dataset was 2.1% (data not shown). The second approach involved quantifying the metabolite profile from a yeast exometabolite sample prepared in triplicate, with the ^{1}H NMR spectra from these samples being collected over the course of three consecutive days. The average RSD value for the metabolites quantified from these profiles was 3.8% with a range of 0.8–9.1% (data not shown) and these values are significantly lower than that observed for any of the yeast exometabolite datasets. Our findings are consistent with those in several past studies that indicate that technical variability is not a significant contributing factor to the variation observed in the model systems examined (Parsons et al. 2009; Fiehn et al. 2000; Crews et al. 2009). Our results reaffirmed the conclusion by Parsons et al. that quantitation using 1D NMR offers excellent analytical precision. In their comparative study of a number of analytical approaches, 1D NMR generated data with the lowest RSD values (median RSD of 3.1%) out of all the analytical approaches examined (Parsons et al. 2009).

Our previous work showed that the two *E. coli* strains used as *C. elegans* food sources produce distinct and significant metabolic profile differences (Reinke et al. 2010). Our present findings show that biological variation can also be significantly affected by diet. This observation highlights the metabolome's sensitivity to external factors. Despite both diets being *E. coli*, a number of factors may contribute to the metabolic variation. First, *E. coli* strains may vary in nutritional composition, and specific amino acids can act as signaling molecules in *C. elegans* (Kang and Avery 2009). Second, worm nutrition may be influenced by the ease of ingestion and digestion of the *E. coli*

strains. Third, the two *E. coli* strains are genetically different and likely respond differentially to growth media. Bacterial responses to stress may impact the host metabolome. Fourth, *E. coli* strains exhibit profound metabolic variability (Maharjan and Ferenci 2005). Fifth, *E. coli* can not only be a source of nutrition for *C. elegans*, but also a pathogen (Darby 2005). Bacteria have been shown to colonize the gut of *C. elegans*, especially in aging worms. Colonization could cause long term damage to the gut lining, thus impeding nutritional absorption. Finally, NGM plates containing *E. coli* HT115 have IPTG and carbenicillin added to induce RNAi and maintain plasmid selection, respectively. As the metabolome is extremely sensitive to environmental factors, these additives may have played a role in metabolic variability.

We also examined whether this biological variation occurs between yeast strains with very similar genetic backgrounds. In previous studies, the complemented SDH strains were considered phenotypically wild-type and identical. However, multivariate analysis using PLS-DA discriminated between these two strains (Fig. 3a). A similar situation is also observed for the comparison between the two knockout strains (Fig. 3b). To rule out the possibility of spurious model generation, PCA and model validations were performed (Figs. 4, 5). The results suggest the models are valid. It has been demonstrated previously that metabolic profiling can discriminate between yeast mutants that are otherwise phenotypically indistinguishable (Allen et al. 2003; Szeto et al. 2010). The results presented in this study further substantiate this notion. There may be several reasons why these pairs of yeast strains are metabolically different. In the case of the complemented SDH strains, there may be subtle differences in the expression of the respective SDH subunits. Upstream or downstream regulatory elements may not have been cloned into the respective plasmids, affecting the plasmid-borne expression of the subunits. Another factor could be plasmid copy number. The yeast strains were each transformed with centromeric plasmids that are maintained at 1–2 copies per cell; however, precise copy number is subject to stochastic variation as cells divide. Increased heterogeneity of gene expression was observed in *E. coli* cells carrying a plasmid-borne copy of the *lacI* gene compared to expression from the chromosomal copy (Elowitz et al. 2002). Subtle changes in the activity of an enzyme can be amplified into larger changes in the metabotype, as we have shown previously (Szeto et al. 2010). The observation that the metabotypes of the knockout strains could also be discriminated is a surprising result considering that both knockouts result in the complete absence of SDH (Szeto et al. 2007; Oyedotun and Lemire 1997, 1999). The different metabotypes of the two knockouts suggest that the Sdh3p and Sdh4p subunits may have additional roles beyond their presence in SDH. It was recently demonstrated that the deletion of *SDH3* results in transcriptional changes associated with fatty acid and sterol metabolism despite being in glucose-repressed conditions. This evidence suggests that nuclear gene signaling is responsive to TCA cycle dysfunction even in situations when this pathway is underutilized (Cimini et al. 2009).

In this study, we examined whether the biological variation observed in the metabolomic studies of other organisms is also observed in simpler model systems, such as yeast and nematodes. Both model systems reveal appreciable amounts of biological variation, with *C. elegans* showing greater metabolic variance. The variation in *C. elegans* was also modulated by the food source. We also examined the biological variation between phenotypically identical yeast strains and were able to discriminate between them by their metabotypes using multivariate analysis. Our results highlight the sensitivity of the metabolome, even in simpler model systems grown under controlled conditions, and the usefulness of metabolic profiling to determine these subtle differences. They also emphasize the importance of appropriate statistical analysis for interpreting data.

Acknowledgments We thank Dr. Brian D. Sykes for the usage of the 600 MHz spectrometer. S.N.R. was supported by an AHFMR Studentship. This work is supported by Canadian Institutes of Health Research Grants MT-15336 and MT-15290 to BDL.

References

Allen J, Davey HM, Broadhurst D, Heald JK, Rowland JJ, Oliver SG, Kell DB (2003) High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. Nat Biotechnol 21(6):692–696. doi:10.1038/nbt823

Bart P, Braeckman BP, Houthoofd K, Vanfleteren JR (2009) Intermediary metabolism. WormBook. doi:10.1895/wormbook.1.146.1

Bogdanov M, Matson WR, Wang L, Matson T, Saunders-Pullman R, Bressman SS, Beal MF (2008) Metabolomic profiling to develop blood biomarkers for Parkinson's disease. Brain 131(2):389–396. doi:10.1093/brain/awm304

Bollard ME, Stanley EG, Lindon JC, Nicholson JK, Holmes E (2005) NMR-based metabonomic approaches for evaluating physiological influences on biofluid composition. NMR Biomed 18(3):143–162. doi:10.1002/nbm.935

Brenner S (1974) The genetics of *Caenorhabditis elegans*. Genetics 77(1):71–94

Cann AJ (2003) Maths from scratch for biologists. Wiley Sons, Chichester

Carrola J, Rocha CuM, Barros AnS, Gil AM, Goodfellow BJ, Carreira IM, Bernardo Jo, Gomes A, Sousa V, Carvalho L, Duarte IF (2010) Metabolic signatures of lung cancer in biofluids: NMR-based metabonomics of urine. J Proteome Res 10(1):221–230. doi:10.1021/pr100899x

Cimini D, Patil K, Schiraldi C, Nielsen J (2009) Global transcriptional response of *Saccharomyces cerevisiae* to the deletion of *SDH3*. BMC Syst Biol 3(1):17. doi:10.1186/1752-0509-3-17

Corder G, Foreman D (2009) Nonparametric statistics for non-statisticians: a step-by-step approach. Wiley, Hoboken

Crews B, Wikoff WR, Patti GJ, Woo HK, Kalisiak E, Heideker J, Siuzdak G (2009) Variability analysis of human plasma and cerebral spinal fluid reveals statistical significance of changes in mass spectrometry-based metabolomics data. Anal Chem 81(20):8538–8544. doi:10.1021/ac9014947

Darby C (2005) Interactions with microbial pathogens. WormBook. doi:10.1895/wormbook.1.21.1

de Bono M, Bargmann CI (1998) Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. Cell 94(5):679–689. doi:10.1016/S0092-8674(00)81609-8

de Kok J, Muller JLM, Slater EC (1975) EPR studies on the respiratory chain of wild-type *Saccharomyces cerevisiae* and mutants with a deficiency in succinate dehydrogenase. Biochim Biophys Acta 387:441–450. doi:10.1016/0005-2728(75)90084-5

Dowlatabadi R, Weljie AM, Thorpe TA, Yeung EC, Vogel HJ (2009) Metabolic footprinting study of white spruce somatic embryogenesis using NMR spectroscopy. Plant Physiol Biochem 47(5):343–350. doi:10.1016/j.plaphy.2008.12.023

Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. Science 297(5584):1183–1186. doi:10.1126/science.1070919

Eriksson LJE, Kettaneh-Wold N, Trygg J, Wikström C, Wold S (2001) Multi- and megavariate data analysis. Part 1: principles and applications, 2nd edn. Umetrics academy, Umeå

Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L (2000) Metabolite profiling for plant functional genomics. Nat Biotechnol 18(11):1157–1161. doi:10.1038/81137

Fraser AG, Kamath RS, Zipperlen P, Martinez-Campos M, Sohrmann M, Ahringer J (2000) Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. Nature 408(6810):325–330. doi:10.1038/35042517

Gietz D, St Jean A, Woods RA, Schiestl RH (1992) Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res 20(6):1425

Grad LI, Sayles LC, Lemire BD (2007) Isolation and functional analysis of mitochondria from the nematode *Caenorhabditis elegans*. In: Leister D, Herrmann J (eds) Mitochondria: practical protocols. Humana Press, Totowa, pp 51–66

Holmes E, Wilson ID, Nicholson JK (2008) Metabolic phenotyping in health and disease. Cell 134(5):714–717. doi:10.1016/j.cell.2008.08.026

Jansson J, Willing B, Lucio M, Fekete A, Dicksved J, Halfvarson J, Tysk C, Schmitt-Kopplin P (2009) Metabolomics reveals metabolic biomarkers of Crohn's disease. PLoS ONE 4(7):e6386. doi:10.1371/journal.pone.0006386

Kamath RS, Martinez-Campos M, Zipperlen P, Fraser AG, Ahringer J (2001) Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. Genome Biol 2(1):research0002.0001–0010. doi:10.1186/gb-2000-2-1-research0002

Kang C, Avery L (2009) Systemic regulation of starvation response in *Caenorhabditis elegans*. Genes Dev 23(1):12–17. doi:10.1101/gad.1723409

Kell DB, Brown M, Davey HM, Dunn WB, Spasic I, Oliver SG (2005) Metabolic footprinting and systems biology: the medium is the message. Nat Rev Microbiol 3(7):557–565. doi:10.1038/nrmicro1177

Lanza IR, Zhang S, Ward LE, Karakelides H, Raftery D, Nair KS (2010) Quantitative metabolomics by ^{1}H-NMR and LC-MS/MS confirms altered metabolic pathways in diabetes. PLoS ONE 5(5):e10538

Lenz EM, Bright J, Wilson ID, Morgan SR, Nash AF (2003) A ^{1}H NMR-based metabonomic study of urine and plasma samples obtained from healthy human subjects. J Pharm Biomed Anal 33(5):1103–1115. doi:10.1016/S0731-7085(03)00410-2

Lewis JA, Fleming JT (1995) Basic culture methods. Methods Cell Biol 48:3–29

MacIntyre DA, Jiménez B, Lewintre EJ, Martin CR, Schäfer H, Ballesteros CG, Mayans JR, Spraul M, Garcia-Conde J, Pineda-Lucena A (2010) Serum metabolome analysis by ^{1}H-NMR reveals differences between chronic lymphocytic leukaemia molecular subgroups. Leukemia 24(4):788–797. doi:10.1038/leu.2009.295

Maharjan RP, Ferenci T (2005) Metabolomic diversity in the species *Escherichia coli* and its relationship to genetic population structure. Metabolomics 1(3):235–242. doi:10.1007/s11306-005-0002-2

Mapelli V, Olsson L, Nielsen J (2008) Metabolic footprinting in microbiology: methods and applications in functional genomics and biotechnology. Trends Biotechnol 26(9):490–497. doi:10.1016/j.tibtech.2008.05.008

Massart D, Smeyers-Verbeke J, Capron X, Schlesier K (2005) Visual presentation of data by means of box plots. LC-GC Eur 18(4):215–218

Nicholson JK (2006) Global systems biology, personalized medicine and molecular epidemiology. Mol Syst Biol 2:52. doi:10.1038/msb4100095

O'Riordan VB, Burnell AM (1989) Intermediary metabolism in the dauer larva of the nematode *Caenorhabditis elegans*—1. Glycolysis, gluconeogenesis, oxidative phosphorylation and the tricarboxylic acid cycle. Comp Biochem Physiol Part B: Biochem Mol Biol 92(2):233–238. doi:10.1016/0305-0491(89)90271-X

O'Riordan VB, Burnell AM (1990) Intermediary metabolism in the dauer larva of the nematode *Caenorhabditis elegans*—II. The glyoxylate cycle and fatty-acid oxidation. Comp Biochem Physiol Part B: Biochem Mol Biol 95(1):125–130. doi:10.1016/0305-0491(90)90258-U

Oyedotun KS, Lemire BD (1997) The carboxyl terminus of the *Saccharomyces cerevisiae* succinate dehydrogenase membrane subunit, Sdh4p, is necessary for ubiquinone reduction and enzyme stability. J Biol Chem 272(50):31382–31388. doi:10.1074/jbc.272.50.31382

Oyedotun KS, Lemire BD (1999) The *Saccharomyces cerevisiae* succinate-ubiquinone oxidoreductase. Identification of Sdh3p amino acid residues involved in ubiquinone binding. J Biol Chem 274(34):23956–23962. doi:10.1074/jbc.274.34.23956

Parsons HM, Ekman DR, Collette TW, Viant MR (2009) Spectral relative standard deviation: a practical benchmark in metabolomics. Analyst 134(3):478–485. doi:10.1039/b808986h

Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. Nat Biotechnol 19(1):45–50. doi:10.1038/83496

Raser JM, O'Shea EK (2005) Noise in gene expression: origins, consequences, and control. Science 309(5743):2010–2013. doi:10.1126/science.1105891

Reinke SN, Hu X, Sykes BD, Lemire BD (2010) *Caenorhabditis elegans* diet significantly affects metabolic profile, mitochondrial DNA levels, lifespan and brood size. Mol Genet Metab 100(3):274–282. doi:10.1016/j.ymgme.2010.03.013

Saude E, Adamko D, Rowe B, Marrie T, Sykes B (2007) Variation of metabolites in normal human urine. Metabolomics 3(4):439–451. doi:10.1007/s11306-007-0091-1

Shaham O, Slate NG, Goldberger O, Xu Q, Ramanathan A, Souza AL, Clish CB, Sims KB, Mootha VK (2010) A plasma signature of human mitochondrial disease revealed through metabolic profiling of spent media from cultured muscle cells. Proc Natl Acad Sci USA 107(4):1571–1575. doi:10.1073/pnas.0906039107

Solanky KS, Bailey NJ, Beckwith-Hall BM, Davis A, Bingham S, Holmes E, Nicholson JK, Cassidy A (2003) Application of biofluid

^{1}H nuclear magnetic resonance-based metabonomic techniques for the analysis of the biochemical effects of dietary isoflavones on human plasma profile. Anal Biochem 323(2):197–204. doi:10.1016/j.ab.2003.08.028

Szeto SS, Reinke SN, Sykes BD, Lemire BD (2007) Ubiquinone-binding site mutations in the *Saccharomyces cerevisiae* succinate dehydrogenase generate superoxide and lead to the accumulation of succinate. J Biol Chem 282(37):27518–27526. doi:10.1074/jbc.M700601200

Szeto SS, Reinke SN, Sykes BD, Lemire BD (2010) Mutations in the *Saccharomyces cerevisiae* succinate dehydrogenase result in distinct metabolic phenotypes revealed through ^{1}H NMR-based metabolic footprinting. J Proteome Res 9(12):6729–6739. doi:10.1021/pr100880y

Walsh MC, Nugent A, Brennan L, Gibney MJ (2008) Understanding the metabolome–challenges for metabolomics. Nutrition Bulletin 33(4):316–323. doi:10.1111/j.1467-3010.2008.00732.x

Weljie AM, Newton J, Mercier P, Carlson E, Slupsky CM (2006) Targeted profiling: quantitative analysis of ^{1}H-NMR metabolomics data. Anal Chem 78(13):4430–4442. doi:10.1021/ac060209g

Zira AN, Theocharis SE, Mitropoulos D, Migdalis V, Mikros E (2010) ^{1}H NMR metabonomic analysis in renal cell carcinoma: a possible diagnostic tool. J Proteome Res 9(8):4038–4044. doi:10.1021/pr100226m