Supporting information for

## <sup>13</sup>C NMR Metabolomics: Applications at Natural Abundance

Chaevien S. Clendinen<sup>1‡</sup>, Brittany Lee-McMullen<sup>1‡</sup>, Caroline Williams<sup>2†</sup>, Gregory S. Stupp<sup>1</sup>, Krista Vandenborne<sup>3</sup>, Daniel A. Hahn<sup>2</sup>, Glenn A. Walter<sup>4,5</sup>, Arthur S. Edison\*<sup>1,5</sup>

1) Department of Biochemistry & Molecular Biology, 2) Department of Entomology and Nematology, 3) Department of Physical Therapy, 4) Department of Physiology and Functional Genomics, 5) Southeast Center for Integrated Metabolomics, University of Florida, Gainesville FL 32610-0245.

<sup>†</sup> Present address: Department of Integrative Biology, University of California, Berkeley CA 94720.

\*Corresponding Author

E-mail: aedison@ufl.edu



Figure S1. H-<sup>1</sup>H STOCSY of Synthetic Mixtures. Normalized and scaled 1D <sup>1</sup>H spectra were used to make <sup>1</sup>H-<sup>1</sup>H correlation maps (STOCSY). Peak lists generated from these correlations were used for database searches.



Figure S2. PCA separation of the mixtures A (blue) and B (red). <sup>13</sup>C spectra were peak picked and <sup>1</sup>H spectra were binned prior to analysis. Metabolite peaks are indicated by the letter abbreviations given in Table S1. Both mixtures were separated using PCA. The <sup>13</sup>C loadings plot (A) of PC1 provided unambiguous peak assignment. The <sup>1</sup>H loadings plot (B) of PC1, though providing adequate separation, some metabolites, such as Lys, would be difficult to identify because its peaks load differently due to overlap.

	Metabolite	<sup>13</sup> C chemical shifts (ppm)
1	Alanine	18.84, 53.22
2	AMP	66.34, 73.36,77.26, 87.6,89.54
3	Arginine	26.56, 30.25, 43.18, 57.01,
4	Glucose	63.45, 72.35, 74.18, 75.46, 76.84, 78.46,
		78.65, 94.79, 98.61
5	Glutamine	28.91, 33.53, 56.71
6	Glutamate	29.65, 36.23, 57.37
7	Glycine	44.14
8	Histidine	30.55, 57.35, 119.61, 138.7
9	Maltose	72.02, 72.64, 73.99, 74.36, 75.39, 75.55,
		76.70, 77.17, 78.91, 79.58, 94.58,
		98.46,102.32
10	DL Methionine	26.53, 50.98, 56.07
	Sulfoxide	
11	*O-Phosphocholine	69.22, 60.76
12	Proline	26.45, 31.69, 48.77, 63.93
13	Taurine	38.08, 50.17

**Table S1)** <sup>13</sup>C fly metabolite and chemical shifts in ppm.



Figure S3. PCA of <sup>13</sup>C (A) and <sup>1</sup>H (B) 1D spectra from cold hardy and cold susceptible flies. PCA scores separated hardy from susceptible flies much better in <sup>13</sup>C data when compared to <sup>1</sup>H data. <sup>13</sup>C loadings plot (A) showed more changes between cold susceptible (blue) and cold hardy (red) flies. <sup>1</sup>H loadings plot (B) showed little differences between the groups, though sugars seemed to load better with the cold susceptible flies. Red and positive peaks indicate resonances that were correlated with cold hardy flies and blue and negative resonances indicate resonances that correlated well with the cold susceptible. Annotations are given for 1D <sup>13</sup>C loadings plot (A).



Figure S4. 2D statistical correlations from mouse serum.  ${}^{13}C{}^{-13}C$  STOCSY (A) and  ${}^{13}C{}^{-1}H$  SHY (B) show highly correlated peaks representing metabolites identified within the mixture. Using the driver peak at 29.8 ppm and the correlating peaks in the  ${}^{13}C$  (A) and the  ${}^{1}H$  (B), we identified linoleic acid as the specific fatty acid contributing to that signal. The BMRB spectra of linoleic acid are shown above as a reference (black). Note that the proton SHY correlations are less helpful than the  ${}^{13}C$  STOCSY correlations in database matching.



Figure S5. <sup>1</sup>H and <sup>13</sup>C NMR and statistical analysis of mouse samples. Scores plot (insets) separate the control (blue) and *mdx* (red) serum. Loadings from the PLS-DA component 1 for the <sup>13</sup>C (A) and <sup>1</sup>H (B) data indicate metabolites that underlie the scores. Unlike the synthetic mixtures or fly data in Figures 1 and 3, PCA did not produce an adequate separation of groups in the mouse samples for either nucleus (data not shown). Thus, <sup>1</sup>H and <sup>13</sup>C datasets were analyzed using PLS-DA. The <sup>13</sup>C data led to a more robust PLS-DA result than <sup>1</sup>H data, with a separation along PC1 with a Q<sup>2</sup> value of 0.67 and R<sup>2</sup> of 0.79 compared to the <sup>1</sup>H results of a Q<sup>2</sup> value of 0.46 and R<sup>2</sup> value of 0.66.