Amino Acid Dysregulation Metabotypes: Potential Biomarkers for Diagnosis and Individualized Treatment for Subtypes of Autism Spectrum Disorder

Supplement 1

SUPPLEMENTAL METHODS

Mass Spectrometry

Mass spectroscopy (MS) was performed using electrospray ionization in positive ion mode with an Agilent QqQ 6490 triple quadrupole mass spectrometer. Analyte selectivity used a combination of product/precursor mass transitions and retention time (**Supplemental Tables S4 and S5**). Agilent MassHunter Quantitative Analysis software (version B.06.00) was used for quantitation of liquid chromatography (LC) MS/MS data. Dynamic Multiple Reaction Monitoring (MRM) was utilized to assign optimal dwell times for each analyte.

Stable isotope labeled (SIL) internal standards (**Supplemental Table S5**) were used to normalize the signal for each analyte to account for variations in the matrix and sample preparation. For analytes in which no SIL internal standard was available, a surrogate SIL internal standard was chosen based on the work of Gray et al. (1) using a structurally similar analyte (**Supplemental Table S4**).

Chromatographic separation was performed using reverse-phase chromatography on a HSS T3 2.1 x 150mm, 1.8µm column (Waters). Column temperature was maintained at 45ºC. The mobile phase was composed of 0.1% formic acid in water and 0.1% formic acid in acetonitrile. A gradient elution was performed which separates the analytes over the course of 7.5 minutes per injection using a flow rate of 0.6 ml/min.

Samples were evaluated relative to calibration standards measured in each analysis batch. Samples that measured below the lowest concentration level of the calibration standard were reported as having a concentration of 0.00 µM. Samples with an analyte(s) that quantified

above the highest concentration level calibration standard were diluted and reanalyzed to obtain a measurement within the range of valid quantification for that analyte.

SUPPLEMENTAL RESULTS

Abundance of Metabolite Ratios and Metabolites Used in the Ratios Are not Changed in AADM Positive and AADM Negative Subjects with Respect to Age and Sex

The mean levels of metabolite ratios or metabolites in the AADM positive and ADDM negative populations were not different (FDR > 0.05) in males and females indicating that the sex bias in detection of the $AADM_{glutamine}$, $AADM_{ornithine}$ and $AADM_{total}$ positive populations is not evident in the levels of metabolites within AADM positive and negative populations (**Supplemental Table S10**). Since there are slight differences between the age of TYP and ASD subjects (3.3 months) and between the age of ASD subjects in the training and test set (1.4 months), we tested if the age of the subject within AADM positive population was associated with the ratio of metabolite abundance levels. No differences in mean (FDR > 0.05) or correlations (FDR > 0.05) of abundance levels of metabolite ratios or individual metabolites of AADM positive populations were found in association with the age of the subjects (**Supplemental Table S11**).

Metabolite Ratios and Metabolites Used in the Ratios are Differentially Abundant in the $AADM_{total}$ Positive Population

Differential analysis of the $AADM_{total}$ positive and negative populations was performed to test if differences in the metabolites are present. The mean levels of metabolite ratios used to identify the $AADM_{total}$ population were increased by 47-82% (FDR < 0.001) in the $AADM_{total}$ positive population when compared to the $AADM_{total}$ negative population. The mean levels of numerator metabolites glutamine, glycine, and ornithine were increased by 16-38% (FDR < 0.001) and BCAA metabolites were decreased by 23-26% (FDR <0.001) in the $AADM_{total}$ positive population compared to the $AADM_{total}$ negative population (**Figure S7**, **Supplemental Table S12**).

SUPPLEMENTAL FIGURES

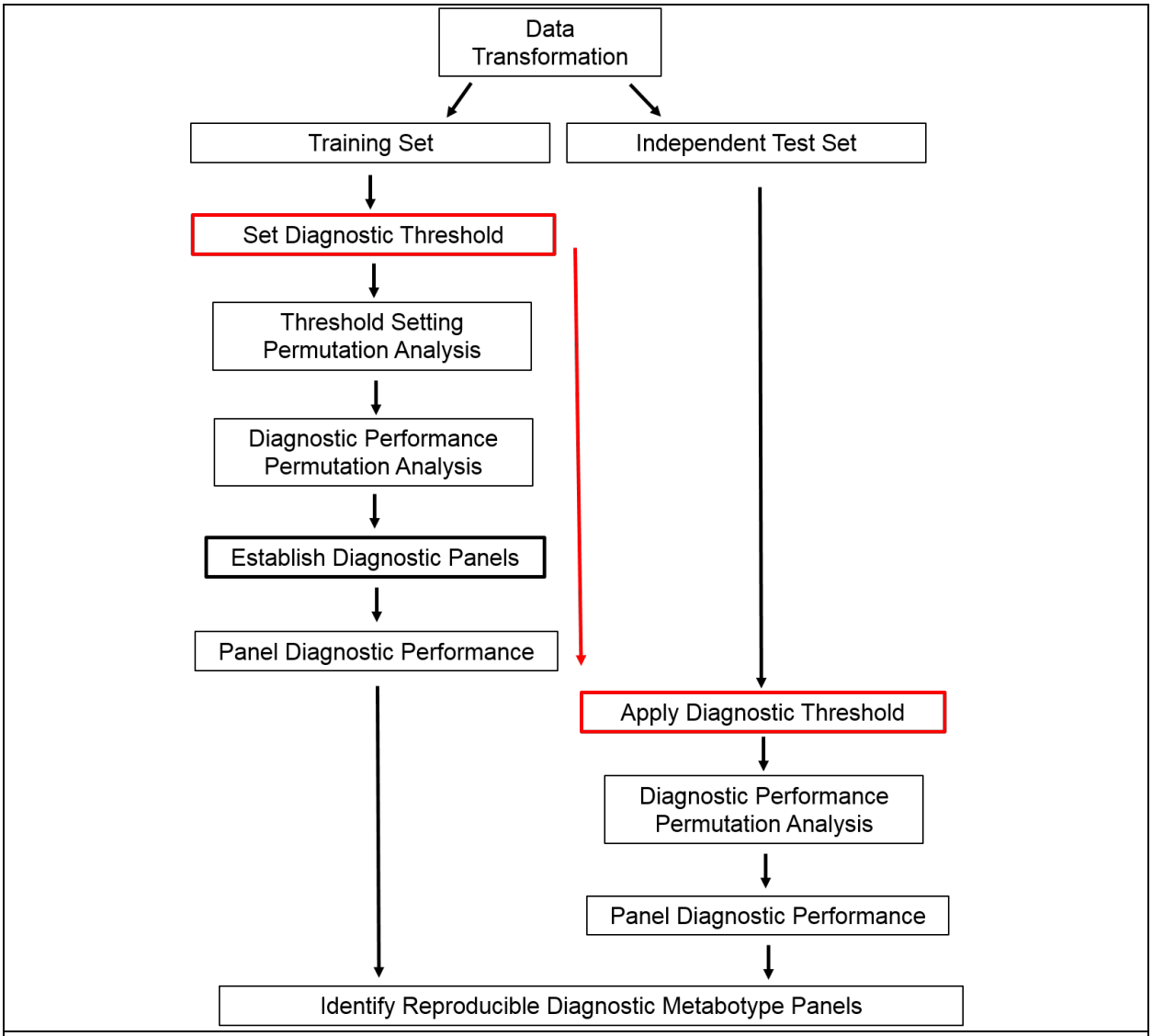

Figure S1. Outline of computational procedures utilized to set diagnostic thresholds and to evaluate diagnostic performance. Diagnostic thresholds were set for each metabolite or metabolite ratio in the training set and the threshold was applied to the training or test set to identify the affected subpopulation and determine the observed diagnostic performance. Permutation analysis was performed to evaluate the frequency at which the observed diagnostic performance occurred by chance. A diagnostic test was considered to identify a relevant metabolic subpopulation if the observed performance metrics occurred in less than 5% of 1000 iterations of random permutations of the subjects' diagnoses.

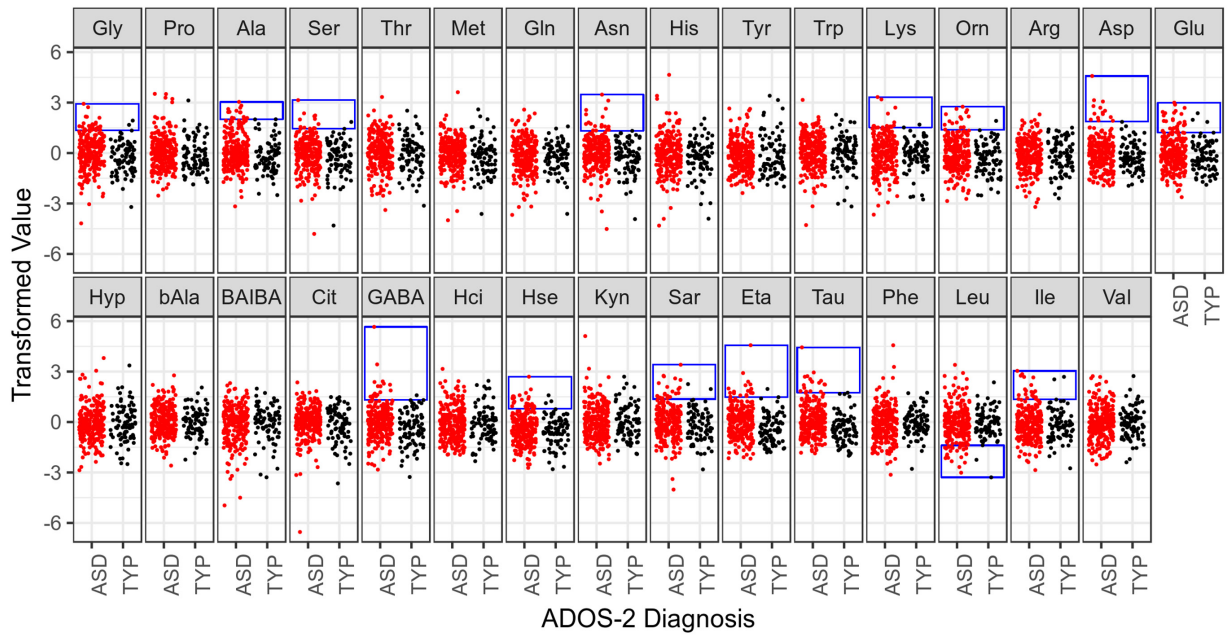

Figure S2. Scatter plot of the training set's transformed amine concentration values. Blue boxes indicate groups that are comprised of greater than 90% ASD subjects (90% PPV). These groups include at least 5% of the training set of ASD subjects (5% sensitivity). Glycine, alanine, asparagine, aspartic acid, GABA, glutamic acid, homoserine, ethanolamine, sarcosine, serine and taurine exhibited elevated metabolite levels in ASD subjects, while leucine exhibited decreased metabolite levels in ASD subjects. Red = ASD, Black = TYP; TYP, Typically Developing; ASD, Autism Spectrum Disorder. BAIBA, β-Aminoisobutyric Acid; GABA, γ-Aminobutyric acid, bAla, β-alanine; Hci, Homocitrulline; Hse, Homoserine; ETA, Ethanolamine; Sar, Sarcosine; Tau, Taurine; Hyp, 4-Hydroxyproline; Cit, Citrulline.

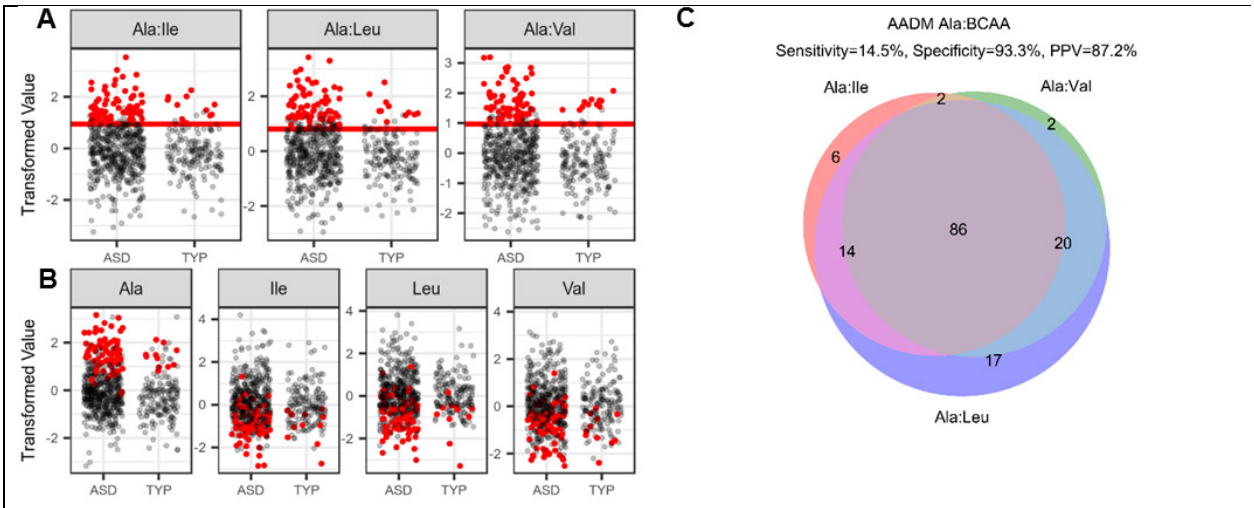

Figure S3. Alanine:BCAA diagnostic for $AADM_{alanine}$. A) Scatter plots of the ratios used to create $AADM_{alanine}$ diagnostic test. Red points represent $AADM_{alanine}$ positive subjects and black points represent $AADM_{alanine}$ negative subjects. The red horizontal line is the diagnostic threshold set in the training set. B) Scatter plots of individual metabolites used in the creation of the ratios. Red dots indicate $AADM_{alanine}$ positive subjects and black points represent those that are $AADM_{alanine}$ negative. C) Venn diagram of subjects identified by the three ratios. Each circle represents the subjects identified by the diagnostic threshold for a given ratio. The intersection of the Venn diagram indicates the subjects called $AADM_{alanine}$ positive (red dots). BCAA, branched chain amino acid; ASD, autism spectrum disorder; TYP, typically developing; AADM, amino acid dysregulation metabotype.

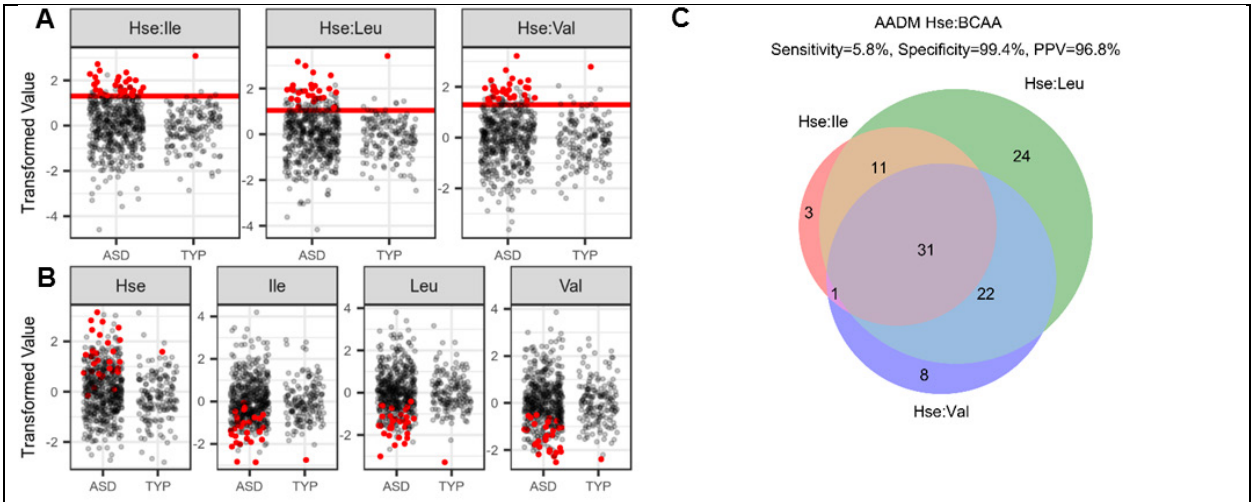

Figure S4. Homoserine:BCAA diagnostic for $AADM_{homoserine}$. A) Scatter plots of the ratios used to create $AADM_{homoserine}$ diagnostic test. Red points represent $AADM_{homoserine}$ positive subjects and black points represent $AADM_{homoserine}$ negative subjects. The red horizontal line is the diagnostic threshold set in the training set. B) Scatter plots of individual metabolites used in the creation of the ratios. Red dots indicate $AADM_{homoserine}$ positive subjects and black points represent those that are $AADM_{homoserine}$ negative. C) Venn diagram of subjects identified by the three ratios. Each circle represents the subjects identified by the diagnostic threshold for a given ratio. The intersection of the Venn diagram indicates the subjects called $AADM_{homoserine}$ positive (red dots). Hse, homoserine; BCAA, branched chain amino acid; ASD, autism spectrum disorder; TYP, typically developing; AADM, amino acid dysregulation metabotype.

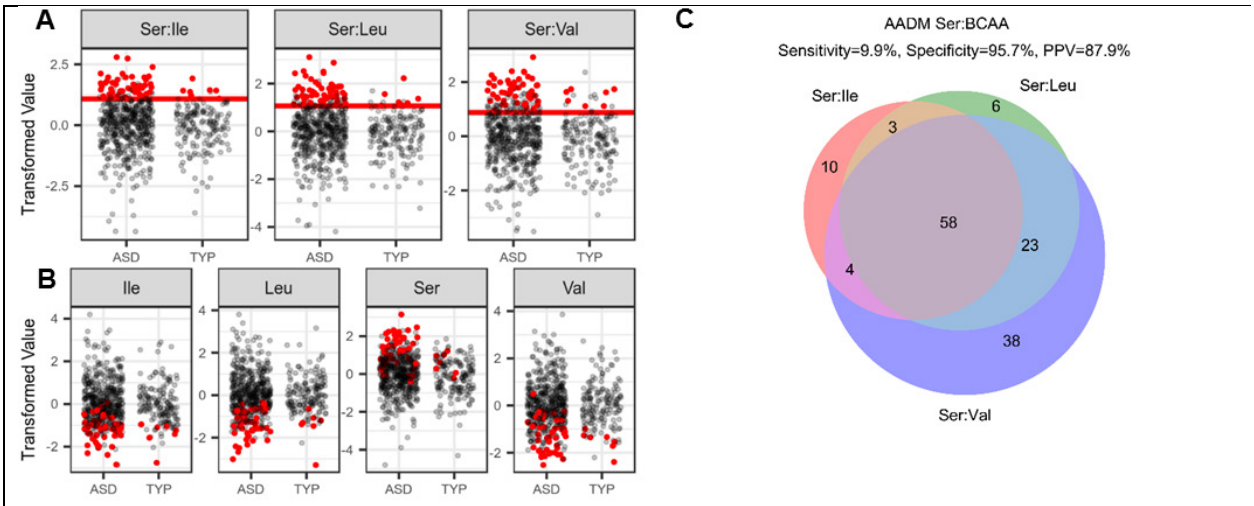

Figure S5. Serine:BCAA diagnostic for $AADM_{serine}$. A) Scatter plots of the ratios used to create $AADM_{serine}$ diagnostic test. Red points represent $AADM_{serine}$ positive subjects and black points represent $AADM_{serine}$ negative subjects. The red horizontal line is the diagnostic threshold set in the training set. B) Scatter plots of individual metabolites used in the creation of the ratios. Red dots indicate $AADM_{serine}$ positive subjects and black points represent those that are $AADM_{serine}$ negative. C) Venn diagram of subjects identified by the three ratios. Each circle represents the subjects identified by the diagnostic threshold for a given ratio. The intersection of the Venn diagram indicates the subjects called $AADM_{serine}$ positive (red dots). BCAA, branched chain amino acid; ASD, autism spectrum disorder; TYP, typically developing; AADM, amino acid dysregulation metabotype.

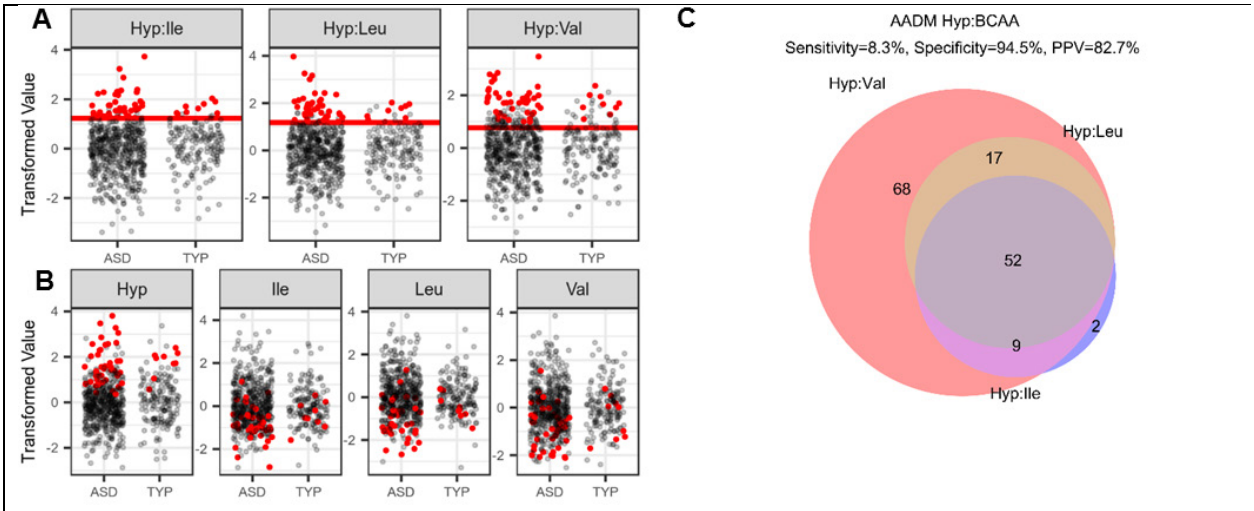

Figure S6. 4-Hydroxyproline:BCAA diagnostic for $AADM_{hydroxyproline}$. A) Scatter plots of the ratios used to create $AADM_{hydroxyproline}$ diagnostic test. Red points represent $AADM_{hydroxyproline}$ positive subjects and black points represent $AADM_{hydroxyproline}$ negative subjects. The red horizontal line is the diagnostic threshold set in the training set. B) Scatter plots of individual metabolites used in the creation of the ratios. Red dots indicate $AADM_{hydroxyproline}$ positive subjects and black points represent those that are $AADM_{hydroxyproline}$ negative. C) Venn diagram of subjects identified by the three ratios. Each circle represents the subjects identified by the diagnostic threshold for a given ratio. The intersection of the Venn diagram indicates the subjects called $AADM_{hydroxyproline}$ positive (red dots). BCAA, branched chain amino acid; ASD, autism spectrum disorder; TYP, typically developing; AADM, amino acid dysregulation metabotype.

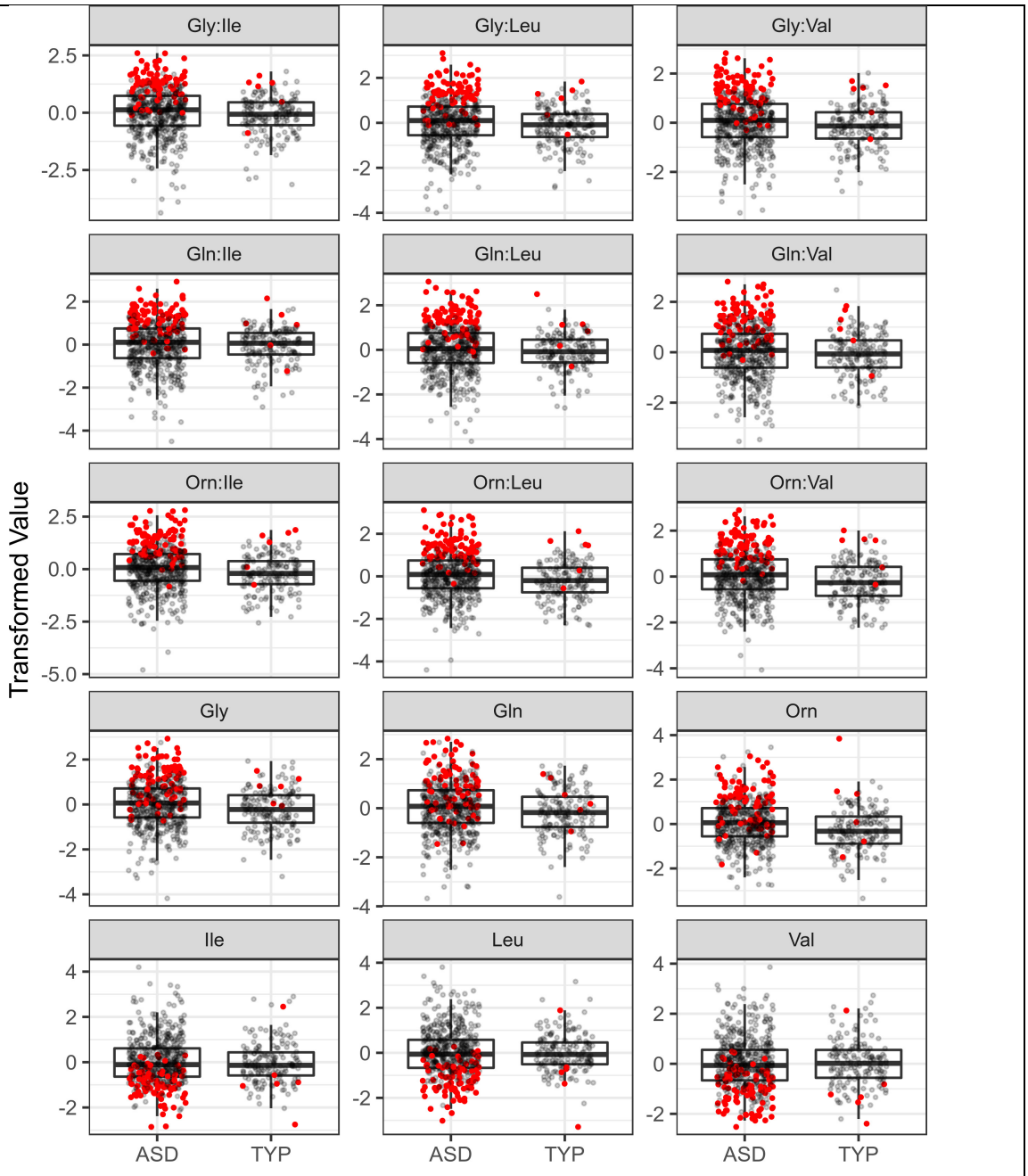

Figure S7. Scatter plots of the ratios and individual metabolites utilized in identification of AADMs. Red points are $AADM_{total}$ positive subjects. Black points $AADM_{total}$ negative subjects. ASD, autism spectrum disorder; TYP, typically developing; AADM, amino acid dysregulation metabotype.

SUPPLEMENTAL TABLES S1 – S13

See Supplement 2 (Excel file).

SUPPLEMENTAL REFERENCES

1. Gray N, Zia R, King A, Patel VC, Wendon J, McPhail MJ, et al. (2017): High-Speed Quantitative UPLC-MS Analysis of Multiple Amines in Human Plasma and Serum via Precolumn Derivatization with 6-Aminoquinolyl-N-hydroxysuccinimidyl Carbamate: Application to Acetaminophen-Induced Liver Failure. *Anal Chem*. 89:2478-2487.

2. Flahault A CM, Thomas G (2005): Sample size calculation should be performed for design accuracy in diagnostic test studies. *J Clin Epidemiol*. 58:859-862.