Archival Report

Biological Psychiatry: Celebrating 50 Years

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

Alan M. Smith, Joseph J. King, Paul R. West, Michael A. Ludwig, Elizabeth L.R. Donley, Robert E. Burrier, and David G. Amaral

ABSTRACT

BACKGROUND: Autism spectrum disorder (ASD) is behaviorally and biologically heterogeneous and likely represents a series of conditions arising from different underlying genetic, metabolic, and environmental factors. There are currently no reliable diagnostic biomarkers for ASD. Based on evidence that dysregulation of branched-chain amino acids (BCAAs) may contribute to the behavioral characteristics of ASD, we tested whether dysregulation of amino acids (AAs) was a pervasive phenomenon in individuals with ASD. This is the first article to report results from the Children's Autism Metabolome Project (CAMP), a large-scale effort to define autism biomarkers based on metabolomic analyses of blood samples from young children.

METHODS: Dysregulation of AA metabolism was identified by comparing plasma metabolites from 516 children with ASD with those from 164 age-matched typically developing children recruited into the CAMP. ASD subjects were stratified into subpopulations based on shared metabolic phenotypes associated with BCAA dysregulation.

RESULTS: We identified groups of AAs with positive correlations that were, as a group, negatively correlated with BCAA levels in ASD. Imbalances between these two groups of AAs identified three ASD-associated amino acid dysregulation metabotypes. The combination of glutamine, glycine, and ornithine amino acid dysregulation metabotypes identified a dysregulation in AA/BCAA metabolism that is present in 16.7% of the CAMP subjects with ASD and is detectable with a specificity of 96.3% and a positive predictive value of 93.5% within the ASD subject cohort.

CONCLUSIONS: Identification and utilization of metabotypes of ASD can lead to actionable metabolic tests that support early diagnosis and stratification for targeted therapeutic interventions.

Keywords: Amino acids, Autism, Biomarker, Diagnosis, Metabolomics, Metabotype

https://doi.org/10.1016/j.biopsych.2018.08.016

Autism spectrum disorder (ASD) is characterized by core symptoms of altered social communication and repetitive behaviors or circumscribed interests (1) and has a prevalence of 1:59 children in the United States. Affected individuals vary enormously in the severity of their autistic characteristics as well as in the occurrence of many comorbid conditions. Comorbid conditions include intellectual disability, which affects at least 40% of individuals with autism (2–4), as well as anxiety in approximately 50% (5), epilepsy in approximately 25% (3,4), and gastrointestinal disorders in approximately 25% of individuals with autism (6). Twin studies (7,8) have indicated that genetic factors play a prominent role in the etiology of ASD, although the genetics of autism appears to be extremely complex. There has been enormous progress in establishing the genetic architecture of ASD, and there are at least 100 genes known to confer risk of ASD (9,10). There is also increasingly strong evidence that environmental factors, alone or in conjunction with genotype, can contribute to the risk for ASD (11). These findings have led to a widespread consensus that there are different biological forms of ASD that may necessitate different diagnostic, preventive, and treatment strategies.

ASD is currently diagnosed based on behavioral characteristics exhibited by affected children (12). While specialist clinicians can confidently diagnose children as young as 24 months (13), the average age of diagnosis in the United States is over 4 years (2,14). Families often experience long waits to receive a definitive diagnosis owing to the paucity of trained clinicians able to perform diagnostic assessment. Early diagnosis is important not only because intensive behavioral therapies are effective in reducing disability in many children with autism (15–17) but also because the benefit of early intervention is greater the earlier the intervention is started.

ISSN: 0006-3223

Unfortunately, there is currently no reliable biomarker that can be used to identify children at risk for ASD (18). Because of the genetic complexity of ASD, there is currently no clinically meaningful genotyping carried out to detect ASD. There have been recent intriguing neuroimaging studies indicating that alterations of brain function or structure as early as 6 months of age may be valuable indicators of a higher risk for autism (19,20). However, it is unlikely that comprehensive structural and functional magnetic resonance imaging is a practical approach to detecting ASD in young children. Other more cost-effective and widely applicable biomarker strategies must be discovered.

We previously demonstrated that a metabolomics approach for the detection of autism risk holds substantial promise (21). In our preliminary study, we identified a subset of 179 features that classified children with ASD and typically developing (TYP) children with 81% accuracy. Metabolism-based analysis has the merit of being sensitive to interactions among the genome, gut microbiome, diet, and environmental factors that contribute to the unique metabolic signature of an individual (22). Metabolic testing can provide important biomarkers toward identifying the perturbations of biological processes underlying an individual's ASD. Past studies have been underpowered to identify metabolic perturbations that lead to actionable metabolic subtypes (23).

To test for metabolic imbalances that can reveal subtypes of subjects with ASD, we conducted the Children's Autism Metabolome Project (CAMP). CAMP recruited 1100 young children (18–48 months) with ASD, intellectual disability, or typical development. Research-reliable clinicians confirmed that children's diagnosis and blood samples were collected under protocols designed specifically for metabolomics analyses. The CAMP study is the largest metabolomics study of ASD to date.

The current study was motivated by observations of amino acid (AA) dysregulation by West *et al.* (21) and in preliminary analysis of the CAMP samples. The relevance of AA dysregulation to ASD was reinforced by Novarino *et al.* (24), who demonstrated loss of function mutations in the gene *BCKDK* (branched chain ketoacid dehydrogenase kinase), resulting in reductions of BCKDK messenger RNA and protein, E1a phosphorylation, and plasma branched-chain AAs in consanguineous families with autism, epilepsy, and intellectual disability. Follow-on studies by Tarlungeanu *et al.* (25) demonstrated that altered transport of branched-chain amino acids (BCAAs) across the blood-brain barrier led to dysregulation of AA levels and neurological impairments. We sought to determine whether dysregulation of AAs was a more pervasive phenomenon in individuals with ASD. Thus, we set out to identify metabotypes indicating the dysregulation of AAs in individuals with ASD and to determine whether these metabotypes might be diagnostic of subsets of individuals. A metabotype is a subpopulation defined by a common metabolic signature that can be differentiated from other members of the study population (26). Metabotypes of ASD can be useful in stratifying the broad autism spectrum into more biochemically homogeneous and clinically significant subtypes. Stratification of ASD based on distinct metabolism can inform pharmacological and dietary interventions that prevent or ameliorate clinical symptoms within a metabotype.

METHODS AND MATERIALS

CAMP Participants

The CAMP study recruited children, ages 18 to 48 months, from eight centers across the United States (Supplemental Table S1). Informed consent of a parent or legal guardian was obtained for each participant. The study protocol was approved and monitored by local institutional review boards at all the sites. Enrollment was limited to one child per household to minimize genetic or family environmental effects. Children participating in other clinical studies could not have used any investigational agent within 30 days of participation. Children were excluded from the study if they were previously diagnosed with a genetic condition such as fragile X syndrome, Rett syndrome, Down syndrome, tuberous sclerosis, or inborn errors of metabolism. Subjects who had fetal alcohol syndrome or other serious neurological, metabolic, psychiatric, cardiovascular, or endocrine system disorders were also excluded. In addition, children exhibiting signs of illness within 2 weeks of enrollment, including vomiting, diarrhea, fever, cough, and ear infection, were rescheduled. Each participant underwent physical, neurological, and behavioral examinations. Metadata were obtained about children's birth, developmental, medical, and immunization histories, dietary supplements, and medications. Parents' medical histories were also obtained.

The Autism Diagnostic Observation Schedule–Second Version (ADOS-2) was performed by research-reliable clinicians to confirm an ASD diagnosis. The Mullen Scales of Early Learning was administered to establish a developmental quotient for all children in the study. A prior ADOS-2 or Mullen Scales of Early Learning assessment was accepted if performed within 90 days of enrollment by qualified personnel. Children without ASD receiving a clinical diagnosis of developmental delay were not included in the current study. Children entering the study as TYP were not routinely administered the ADOS-2. The Social Communications Questionnaire was administered to a subset of 65 TYP children as a screen for ASD. Of these, 9 children were referred for subsequent ADOS-2 evaluation, with 4 receiving a diagnosis of ASD (and being included in this study) and 5 receiving a diagnosis of TYP.

Training and Test Sets

A training set was used to identify metabotypes associated with ASD, and a test set was used to evaluate the reproducibility of the metabotypes. The sample size of the training set was designed to detect metabotypes with a sensitivity (metabotype prevalence) $>$ 3% and specificity $>$ 85% with a power of 0.90 (Supplemental Table S2). The training set (n = 338; ASD = 253, TYP = 85) was established and analyzed, and then, as recruitment continued, the test set (n = 342; ASD = 263, TYP = 79) was established when sufficient subjects were available to match the training set demographics (Supplemental Table S3).

Phlebotomy Procedures

Blood was collected from subjects after at least a 12-hour fast by venipuncture into 6-mL sodium heparin tubes on wet ice. A minimum of a 2-mL blood draw was required for sample

inclusion in the computational analyses. The plasma was obtained by centrifugation (1200*g* for 10 minutes) and frozen to −70°C within 1 hour.

Triple Quadrupole Liquid Chromatography–Tandem Mass Spectrometry Method for Quantitative Analysis of Biological Amines

The Waters AccQTag Ultra kit (Waters Corporation, Milford, MA), which employs derivatization of amine moieties with 6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate, was employed for all samples prior to multiple reaction monitoring on a liquid chromatography–mass spectrometry system consisting of an Agilent 1290 ultra-high-performance liquid chromatograph coupled with an Agilent G6490 Triple Quadrupole Mass Spectrometer (Agilent Technologies, Santa Clara, CA) (Supplemental Methods and Supplemental Tables S4 and S5).

Bioinformatic Analysis

The concentration values of each metabolite were log base 2 transformed and *Z* score normalized prior to analyses. Analysis of covariance and pairwise Pearson correlation analysis were performed on each amine compound. False discovery rates were controlled for multiple testing using the Benjamini and Hochberg (27) method of *p*-value correction. A comparison was considered significant if the corrected *p* value was less than .05. A dissimilarity measurement of 1 − absolute value of Pearson correlation coefficient (ρ) was used to calculate distances for clustering. Wards' method was used for hierarchical clustering. Bootstrap analysis of the clustering result was performed using the pvclust package in R (28). A cluster was considered significant if the unbiased *p* value was ≥ .95. The nonlinear iterative partial least squares algorithm was used for principal component analysis (PCA), and confidence intervals were drawn at the 95th percentile of the PCA scores using Hotelling's T2 statistic with the package PCAmethods (29). Welch *t* tests were used to test for differences in study populations. The independence of subject metadata relative to the metabotypes was tested using the Fisher exact test statistic with an alpha of .05 to reject the null hypothesis. These analyses were conducted with R (version 3.4.1).

Establishing and Assessing Diagnostic Thresholds

A heuristic algorithm was developed to identify individual biomarkers able to discriminate ASD subpopulations, indicative of a metabotype, using a threshold (Supplemental Figure S1) for metabolite abundance or ratios. Diagnostic thresholds were established in the training set to generate a subpopulation with at least 10% of the ASD population while minimizing the number of TYP subjects in the subpopulation. Subjects exceeding the diagnostic threshold were scored as metabotype-positive subjects with ASD, and the remaining subjects were scored as metabotype negative. Diagnostic performance metrics of specificity (detection of TYP), sensitivity (detection of ASD), and positive predictive value (PPV), i.e., percentage of metabotype-positive subjects with ASD, were calculated based on metabotype status (positive or negative) and ADOS-2 diagnosis (ASD or TYP) of the CAMP subjects. PPV was not adjusted for prevalence of ASD in the general population or for populations at greater risk of ASD. The PPV reported is based on the CAMP study population.

Permutation analysis was performed to test the probability that the observed diagnostic performance values from threshold setting and subpopulation prediction could be due to chance. Chance was assessed using 1000 permutations of subject diagnoses in the training set for threshold setting and subpopulation prediction or test set following subpopulation prediction. In both permutation procedures, the probability that observed biomarker performance metrics were due to chance was calculated based on the frequency that the observed sensitivity, specificity, and PPV were met or exceeded in the random permutation set.

When the diagnostic ratios were combined into panels of ratios to test for ASD-associated metabotypes, the minimum performance required to consider a metabotype as reproducible was a sensitivity ≥ 5%, a specificity ≥ 95%, and a PPV ≥ 90% in both the training and test sets.

RESULTS

CAMP Study Demographics

The training and test sets of subjects were chosen with appropriate power and randomization (Supplemental Table S3). The ASD prevalence, developmental quotient, and sex composition between the training and test sets were equivalent ($p > .05$). However, the ASD population contained 16.5% more boys than the TYP population ($p < .01$). The ASD population was slightly older than the TYP population by 3.3 months ($p < .01$), and subjects with ASD in the training set were 1.4 months older than subjects with ASD in the test set ($p < .01$).

Analysis of Amine-Containing Metabolites Between ASD and TYP Study Populations

Analysis of covariance was performed on 31 amine-containing metabolites in the training set of subjects to test the effect of sex or diagnosis, controlling for subject age on metabolite means. No significant differences were identified in metabolite abundance values for diagnosis, age, sex, or their interactions (Supplemental Table S6). These results suggest that within the demographic ranges in this study, the differences in subject age or sex had little impact on metabolite abundance. Therefore, the differences in the composition of ASD and TYP study populations were unlikely to have significant impact on study results.

Metabolite Correlations Within ASD Reveal Distinct Clusters of Amine Metabolites

We then examined the relationship among the amine metabolites in the training set of subjects with ASD by pairwise Pearson correlation analysis and hierarchical clustering to identify metabolites with coregulated metabolism (Figure 1). Three clusters of metabolites with positive correlations were identified. Cluster 1 contains the metabolites serine, glycine, ornithine, 4-hydroxyproline, alanine, glutamine, homoserine, and proline (i.e., the glycine cluster—mean ρ of 0.45 ± 0.02). Cluster 2 contains the BCAAs (leucine, isoleucine, and valine) and phenylalanine, where the BCAAs are more highly correlated with each other (mean pairwise ρ of 0.86 ± 0.02) than the BCAAs are with phenylalanine (mean pairwise ρ of 0.56 ± 0.02)

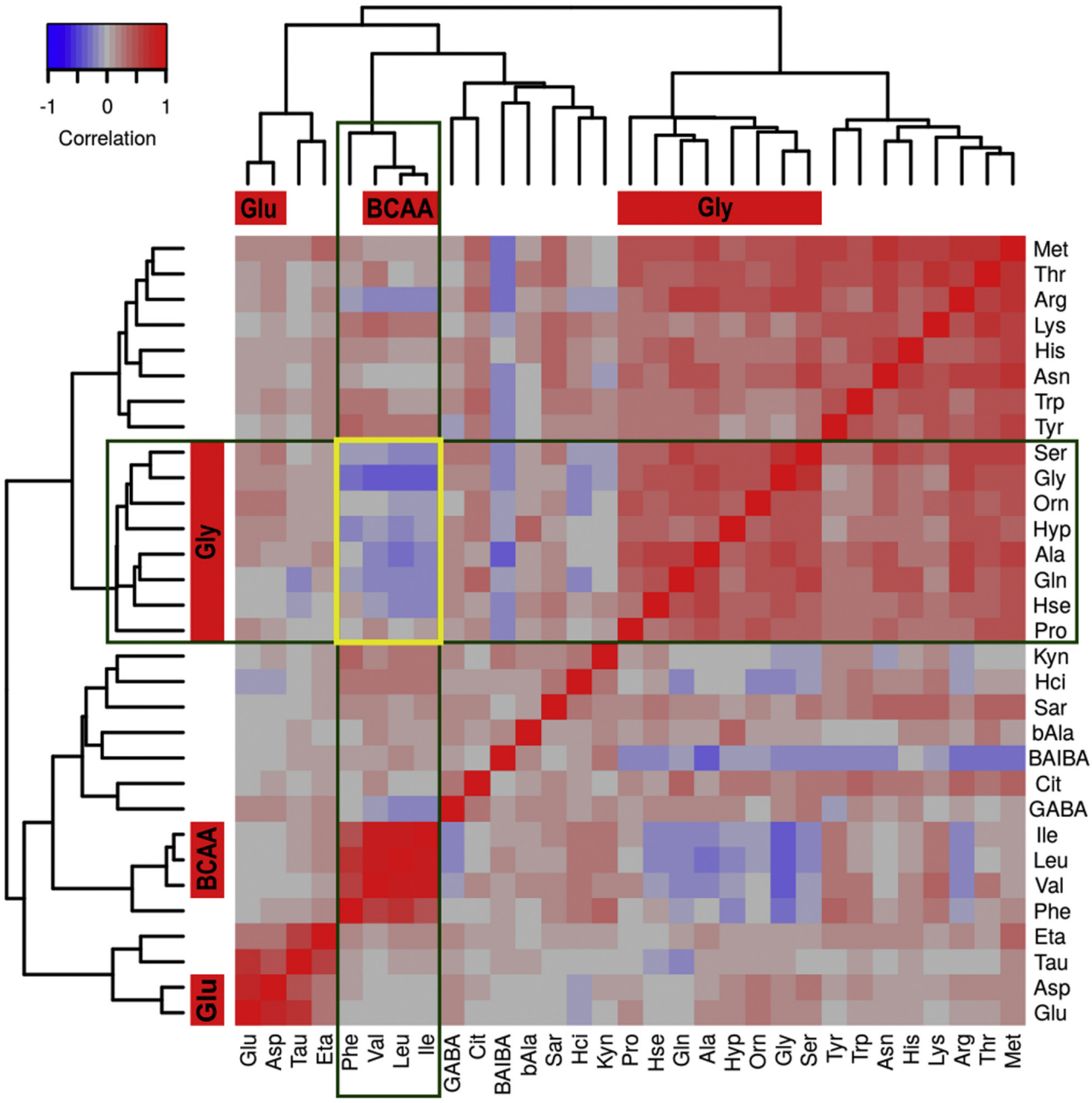


Figure 1. Heat map with hierarchical clustering dendrograms from pairwise Pearson correlations of metabolite abundances for the training set subjects with autism spectrum disorder. Red-filled boxes associated with the dendrograms identify statistically significant clusters following bootstrap resampling. The names of these clusters appear within the red boxes. Green open boxes highlight the branched-chain amino acid (BCAA) cluster in the columns and the glycine (Gly) cluster in the rows. The intersection of the two green boxes, marked by a yellow open rectangle, identifies the block of negative correlations shared by the Gly and BCAA clusters. Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartate; BAIBA, β-aminoisobutyric acid; bAla, β-alanine; Cit, citrulline; Eta, ethanolamine; GABA, gamma-aminobutyric acid; Gln, glutamine; Glu, glutamic acid; Hci, homocitrulline; His, histidine; Hse, homoserine; Hyp, 4-hydroxyproline; Ile, isoleucine; Kyn, kynurenine; Leu, leucine; Lys, lysine; Met, methionine; Orn, ornithine; Phe, phenylalaline; Pro, proline; Sar, sarcosine; Ser, serine; Tau, taurine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; Val, valine.

(i.e., the BCAA cluster red boxes in Figure 1). Cluster 3 contains glutamate and aspartate (i.e., the glutamate cluster—ρ of 0.78) (Figure 1). The intersection of the glycine and BCAA clusters yielded a block of negative correlations (intersection of boxes in Figure 1). We decided to focus our analysis on the glycine cluster metabolites that are negatively correlated with BCAA metabolites. Proline was removed from further analysis because it was not negatively correlated with the BCAAs.

Phenylalanine was removed because it is not a BCAA metabolite.

Identification of AA:BCAA Imbalance Metabotypes Associated With ASD

The negative correlation between the BCAAs and glycine cluster led us to evaluate ratios of these AAs as predictors of ASD diagnosis. Ratios can uncover biological properties not evident with individual metabolites and increase the signal when two metabolites with a negative correlation are evaluated. This strategy, for example, formed the basis of the standard phenylketonuria diagnostic using a ratio of phenylalanine and tyrosine (30). Based on analysis of box plots (Supplemental Figure S2), we created ratios with one of the BCAAs in the denominator and one of the glycine cluster metabolites in the numerator. Thresholds for the ratios were set in the training set and evaluated in the test set of subjects (Table 1). The BCAA ratios of glutamine, glycine, and ornithine identified reproducible subpopulations of subjects associated with an ASD diagnosis that met minimum performance criteria in both the training and test sets (Supplemental Tables S7 and S8).

The correlation of the BCAAs with each other (ρ = 0.86 ± 0.02) and the overlap of affected subjects (Venn diagrams in Figure 2) identified by the AA:BCAA ratios suggested that a combination of ratios containing a single numerator and each of the three BCAAs as a denominator could uncover BCAA metabolic dysregulation. Exploiting the positive correlation among the BCAAs in this way improves the specificity and PPV. For example, each of the glycine:BCAA ratios (i.e., glycine:leucine, glycine:valine, or glycine:isoleucine) results in a specificity of 94.1% and a PPV of 91.1% (Supplemental Table S9). However,

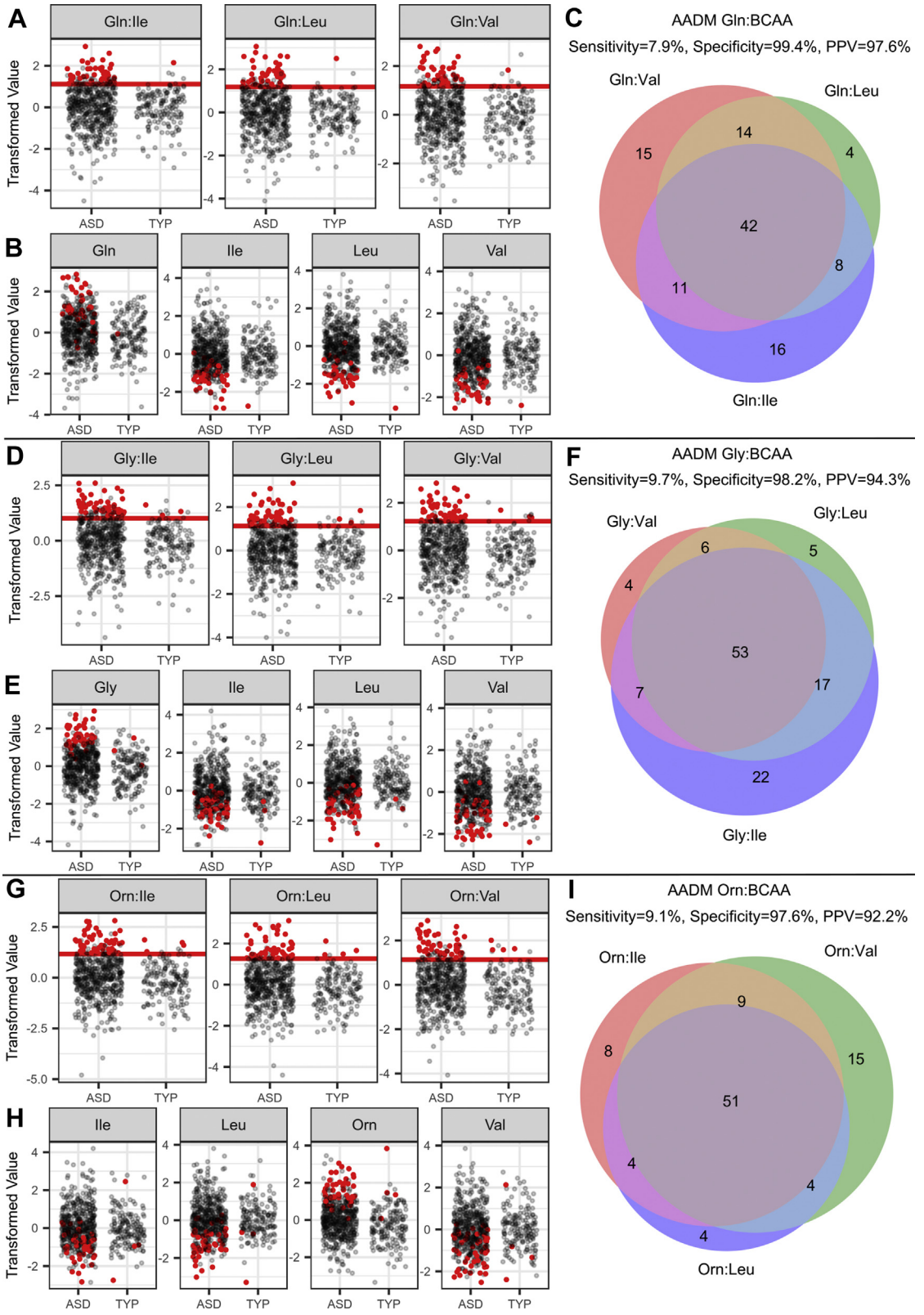
Table 1. Diagnostic Performance Metrics of Amine Ratios to Discriminate Subpopulations of Subjects With ASD in the Training and Test Sets of Subjects

	Sensitivity		Specificity		PPV		Permutation Test	
Ratio	Train	Test	Train	Test	Train	Test	Train	Test
Ratios Used to Create $AADM_{alanine}$								
Ala:Ile	.202	.163	.894	.937	.850	.896	3.40×10^{-2}	1.30×10^{-2}
Ala:Leu	.245	.217	.894	.886	.873	.864	2.00×10^{-3}	2.40×10^{-2}
Ala:Val	.178	.186	.918	.886	.866	.845	2.20×10^{-2}	9.90×10^{-2}
Ratios Used to Create $AADM_{glutamine}$								
Gln:Ile[a]	.126	.152	.976	.962	.941	.930	3.00×10^{-3}	2.00×10^{-3}
Gln:Leu[a]	.103	.148	.976	.987	.929	.975	9.00×10^{-3}	0
Gln:Val	.130	.160	.976	.937	.943	.894	1.00×10^{-3}	1.50×10^{-2}
Ratios Used to Create $AADM_{glycine}$								
Gly:Ile[a]	.174	.175	.953	.937	.917	.902	0	9.00×10^{-3}
Gly:Leu	.146	.133	.953	.937	.902	.875	8.00×10^{-3}	5.50×10^{-2}
Gly:Val	.126	.114	.976	.924	.941	.833	3.00×10^{-3}	2.22×10^{-1}
Ratios Used to Create $AADM_{homoserine}$								
Hse:Ile	.063	.103	.976	.949	.889	.964	1.16×10^{-1}	9.00×10^{-3}
Hse:Leu	.107	.202	.953	.975	.871	.930	8.20×10^{-2}	0
Hse:Val	.067	.141	.965	.962	.850	.881	2.09×10^{-1}	4.30×10^{-2}
Ratios Used to Create $AADM_{ornithine}$								
Orn:Ile[a]	.115	.137	.965	.949	.906	.900	2.10×10^{-2}	2.50×10^{-2}
Orn:Leu[a]	.103	.122	.965	.975	.897	.941	3.70×10^{-2}	3.00×10^{-3}
Orn:Val	.119	.160	.953	.962	.882	.933	3.70×10^{-2}	4.00×10^{-3}
Ratios Used to Create $AADM_{serine}$								
Ser:Ile	.130	.129	.953	.949	.892	.895	2.00×10^{-2}	2.20×10^{-2}
Ser:Leu	.138	.167	.941	.924	.875	.880	3.00×10^{-2}	3.80×10^{-2}
Ser:Val	.190	.228	.941	.873	.906	.857	1.00×10^{-3}	3.90×10^{-2}
Ratios Used to Create $AADM_{hydroxyproline}$								
Hyp:Ile	.111	.087	.941	.911	.848	.767	1.14×10^{-1}	6.03×10^{-1}
Hyp:Leu	.115	.095	.918	.899	.806	.758	2.69×10^{-1}	6.64×10^{-1}
Hyp:Val	.241	.213	.871	.772	.847	.757	1.70×10^{-2}	6.64×10^{-1}

The ratios all include branched-chain amino acid (BCAA) values in the denominator and negatively correlated glycine cluster metabolites in the numerator.

AADM, amino acid dysregulation metabotype; Ala, alanine; ASD, autism spectrum disorder; Gln, glutamine; Gly, glycine; Hse, homoserine; Hyp, 4-hydroxyproline; Ile, isoleucine; Leu, leucine; Orn, ornithine; PPV, positive predictive value; Ser, serine; Test, test set; Train, training set; Val, valine.

[a]Ratios are reproducible metabotypes that are identified across training and test populations with a sensitivity greater than 5% and a PPV greater than 90%.


A
Gln:Ile
Gln:Leu
Gln:Val
Transformed Value
ASD
TYP
B
Gln
Ile
Leu
Val
C
AADM Gln:BCAA
Sensitivity=7.9%, Specificity=99.4%, PPV=97.6%
Gln:Val
Gln:Leu
Gln:Ile
15
14
4
42
8
11
16
D
Gly:Ile
Gly:Leu
Gly:Val
E
Gly
F
AADM Gly:BCAA
Sensitivity=9.7%, Specificity=98.2%, PPV=94.3%
Gly:Val
Gly:Leu
Gly:Ile
6
5
4
53
7
17
22
G
Orn:Ile
Orn:Leu
Orn:Val
H
Orn
I
AADM Orn:BCAA
Sensitivity=9.1%, Specificity=97.6%, PPV=92.2%
Orn:Ile
Orn:Val
Orn:Leu
9
8
15
51
4
4
4

requiring that subjects be positive for all three glycine:BCAA ratios results in a specificity of 98.8% and a PPV of 96.0%. Through this process, we identified groups of subjects who exhibited an amino acid dysregulation metabotype (AADM). Subjects were identified by AADM when they exceeded an established threshold for all three AA:BCAA ratios. Because the nomenclature for these biomarkers can quickly become confusing, we have designated different AADMs using the numerator metabolite, for example, $AADM_{glutamine}$ (Figure 2). Not all AA:BCAA ratios resulted in diagnostic differences between the ASD and TYP groups (Table 2 and Supplemental Figures S3–S6). We focused, therefore, on those AA:BCAA ratios that had the greatest predictive power, including glutamine AADM ($AADM_{glutamine}$) (Figure 2A–C), glycine AADM ($AADM_{glycine}$), (Figure 2D–F), and ornithine AADM ($AADM_{ornithine}$) (Figure 2G–I).

AADMs Define a Diagnostic for BCAA Dysregulation Associated With ASD

The subjects with ASD identified by each AADM were evaluated to assess the extent of overlap. We found that there is substantial overlap of the subjects identified by each of the metabotypes (Figure 3). However, each of the metabotypes also identifies a unique group of subjects. The $AADM_{glutamine}$ identified 7.9% of the subjects with ASD in the total CAMP population, the $AADM_{glycine}$ identified 9.7%, and the $AADM_{ornithine}$ identified 9.1%, with PPVs of 97.6%, 94.3%, and 92.2%, respectively. Combining all three AADM subtypes together ($AADM_{total}$) identified 16.7% of subjects with ASD in the CAMP population with a specificity of 96.3% and a PPV of 93.5% (Figure 3A). PCA of the metabolite ratios used in $AADM_{glycine}$, $AADM_{glutamine}$, and $AADM_{ornithine}$ was performed to test whether an unsupervised method could identify subjects with AA dysregulation. A majority (80%, 74 of 92) of the AADM-positive subjects were separated from the unaffected subjects (Figure 3B).

$AADM_{ornithine}$ and $AADM_{glutamine}$ Are More Sensitive at Detecting Girls With ASD

Because the composition of subject sex and age differed between the ASD and TYP populations, the impact of these variables was evaluated in the AADM-positive and AADM-negative populations. Differential analysis of reproducible AADM-positive and AADM-negative subjects' metabolite levels with respect to age or sex did not identify statistically significant changes in abundance (Supplemental Results and Supplemental Tables S11 and S12). Girls with ASD were 2.1-fold (odds ratio = 2.8, $p = .002$) more likely to be identified by $AADM_{ornithine}$ and $AADM_{glutamine}$ than would be expected by chance (Supplemental Table S13). $AADM_{glycine}$ did not demonstrate a predictive sex bias.

Table 2. Diagnostic Performance Metrics of AADMs

	Sensitivity		Specificity		PPV	
AADM Diagnostic	Train	Test	Train	Test	Train	Test
Ala:BCAA	.150	.141	.929	.937	.864	.881
Gln:BCAA[a]	.079	.080	.988	1.000	.952	1.000
Gly:BCAA[a]	.095	.099	.988	.975	.960	.929
Hse:BCAA	.036	.080	.988	1.000	.900	1.000
Orn:BCAA[a]	.079	.103	.976	.975	.909	.931
Ser:BCAA	.091	.106	.965	.949	.885	.875
Hyp:BCAA	.087	.080	.965	.924	.880	.778

Each amino acid dysregulation metabotype (AADM) consists of three ratios with a different branched-chain amino acid (BCAA) in the denominator.

Ala, alanine; Gln, glutamine; Gly, glycine; Hse, homoserine; Hyp, 4-hydroxyproline; Orn, ornithine; PPV, positive predictive value; Train, training set; Test, test set.

[a]AADMs are reproducible metabotypes that are identified across training and test populations with a sensitivity greater than 5% and a PPV greater than 90%.

DISCUSSION

CAMP is the largest study of the metabolism of children with ASD and age-matched TYP children carried out to date. Metabolomics offers the opportunity to examine associations between small-molecule abundance levels and the presence of a disorder such as ASD as well as influences such as sex, severity of the disorder, comorbid conditions, diet, supplements, and other environmental factors. Given the known heterogeneity of ASD, the size of CAMP offers the prospect of identifying metabolically defined subtypes (or metabotypes) that can identify groups with a prevalence as low as 5%. Diagnostic tests for metabotypes of ASD create an opportunity for earlier diagnosis and the potential to inform more targeted treatment.

Our goal was to analyze data from the CAMP population to identify metabotypes associated with ASD that could enable stratification of the disorder based on shared metabolic characteristics. Based on our own observations and growing literature (23–25,31,32) reporting a dysregulation of AA metabolism associated with ASD, we began our analysis by studying free plasma amine levels. A simple analysis of the mean concentrations of free plasma amines did not reveal meaningful differences between the ASD and TYP populations of children. However, scatter plots of amine levels indicated that there were subsets of children with ASD with amine levels at the extreme upper or lower end of the abundance distribution. Moreover, correlation analyses revealed two negatively correlated clusters of related metabolites. We tested whether ratios of these metabolites could identify subpopulations that exhibit dysregulation of AA metabolism associated with ASD.

Figure 2. Levels of amino acid ratios and of individual amino acids. Venn diagrams of the metabotype-positive subjects identified by each of the amino acid:branched-chain amino acid (AA:BCAA) diagnostics for amino acid dysregulation metabotype $(AADM)_{glutamine}$ **(A–C)**, $AADM_{glycine}$ **(D–F)**, and $AADM_{ornithine}$ **(G–I)** in the training and test sets are shown. **(A, D, G)** Scatter plots of the AA:BCAA ratios used to create an AADM diagnostic test. Red points represent AADM-positive subjects, and black points represent AADM-negative subjects. The red horizontal line is the diagnostic threshold set in the training set. **(B, E, H)** Scatter plots of individual amino acids used in the creation of the ratios. Red dots indicate AADM-positive subjects, and black points represent AADM-negative subjects. **(C, F, I)** Venn diagram of metabotype-positive subjects identified by the three ratios used to create each AA:BCAA diagnostic. 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 positive (red dots in scatter plots). Performance metrics above the Venn diagram represent entire study population (training and test sets). ASD, autism spectrum disorder; Gln, glutamine; Gly, glycine; Ile, isoleucine; Leu, leucine; Orn, ornithine; PPV, positive predictive value; TYP, typically developing; Val, valine.

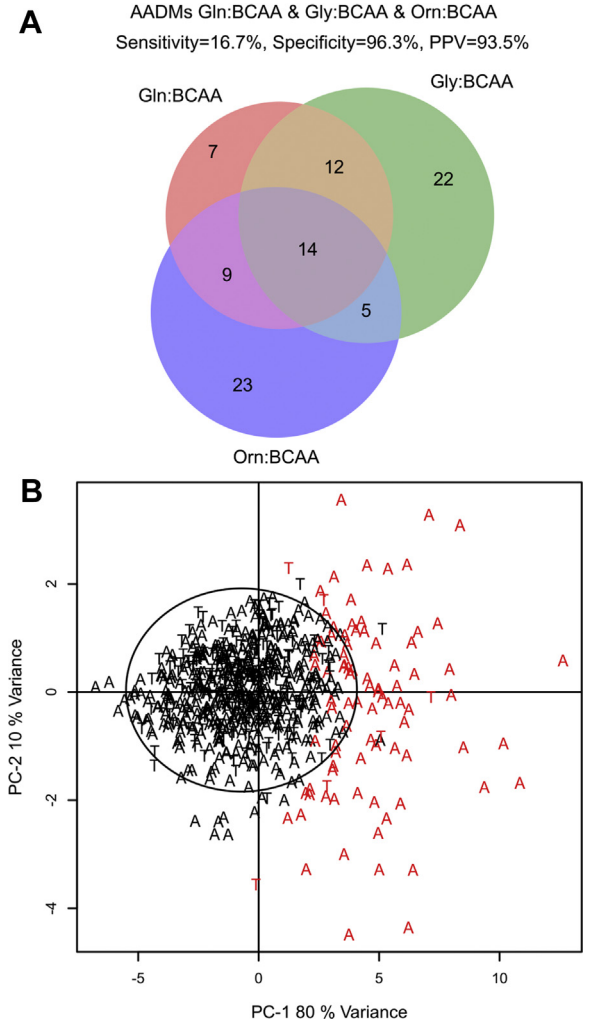


Figure 3. **(A)** Venn diagram of the 92 amino acid dysregulation metabotype $(AADM)_{total}$ subjects identified by each of the AADMs. At least 50% of the subjects identified by one AADM were identified by the other two AADMs. The $AADM_{total}$ population is composed of 86 subjects with autism spectrum disorder (ASD) and 6 typically developing (TYP) subjects. The overall prevalence of metabolic dysregulation in the Children's Autism Metabolome Project (CAMP) ASD population is 16.7% (86 $AADM_{total}$ ASD/516 CAMP ASD), with a specificity of 96.3% (158 AADM-negative TYP/164 CAMP TYP) and a PPV of 93.5% (86 $AADM_{total}$ ASD/92 $AADM_{total}$). **(B)** Principal component (PC) analysis of the metabolite ratios used in the metabolic signature of the reproducible AADMs creating the $AADM_{total}$ estimates in the CAMP study population. The black circle is the 95% confidence interval from the Hotelling's T2. Red letters are $AADM_{total}$ positive (N = 92), and black letters are $AADM_{total}$ negative (N = 588). A, ASD; BCAA, branched-chain amino acid; Gln, glutamine; Gly, glycine; Orn, ornithine; PPV, positive predictive value; T, TYP.

Diagnostic thresholds established in the training set of subjects using ratios of glutamine, glycine, and ornithine, with leucine, isoleucine, and valine (BCAAs) reproducibly detected subpopulations in an independent test set. Three AADMs based on an imbalance of glutamine, glycine, or ornithine with the BCAAs were reproduced across training and test sets of subjects. Separately, each AADM identified subjects with ASD with 7% to 10% sensitivity and 92% to 98% PPVs. Taken together, all AADMs identified an altered metabolic phenotype of imbalanced BCAA metabolism in 16.7% of CAMP subjects with ASD with a specificity of 96.3% and a PPV of 93.5%. We wish to note that our use of PPV was not adjusted for prevalence of ASD in the general population or for populations at greater risk of ASD. The PPV reported is based on the CAMP study population and was used as one facet of the criteria to define metabotypes. We do not imply that a general population screen using the AADMs would achieve a similar specificity or PPV.

Identification of children with ASD and altered AADMs represents an important step toward understanding the etiology of one form of ASD. Imbalances in BCAAs in plasma have been shown to alter not only brain levels of BCAAs but also other brain AAs important for key metabolic processes, including intermediary metabolism, protein synthesis, and neurotransmission. For example, when plasma BCAA levels are reduced owing to a rare genetic defect in *BCKDK* (24), leading to accelerated BCAA degradation, the transporters that are normally responsible for their import into the brain transport an excess of other AAs instead. In addition, this condition is associated with ASD (24). Similarly, Tarlungeanu *et al.* (25) demonstrated that rare disruption of AA transport associated with defects in the L-type amino acid transporter 1 reduced the uptake of BCAAs into the brain; again this was associated with ASD-like symptoms. Interestingly, neither study reported elevated plasma levels of glycine, ornithine, or glutamine. The imbalance of AA levels in CAMP strongly suggests that other perturbations in BCAA metabolism may be a risk factor for the development of ASD. Importantly, the metabolomic results reported here provide a mechanism for stratifying the larger group of children with ASD into an AADM-positive subgroup to enable a more targeted approach to understanding the etiology of this form of ASD. For example, the AADMs we identified may reveal a disruption of the mechanistic target of rapamycin complex 1 system, which could be an underlying reason for lower free plasma BCAA levels. Cellular levels of BCAAs as well as other AAs are maintained through signaling-associated mechanistic target of rapamycin complex 1 and activating transcription factor 4 (33). Dysregulation of the mechanistic target of rapamycin kinase pathway is an underlying cause of AA dysregulation that is associated with ASD and tuberous sclerosis (34).

The AADMs provide one pathway to much earlier diagnosis of a substantial subset of children with ASD. Earlier diagnosis may also provide the opportunity for earlier biological intervention. BCAA supplementation or high-protein diet has been used in mouse models (24) and human patients (31) with BCKDK deficiency to successfully reduce ASD symptoms and improve cognitive function. Defining a group of AADM-positive children may enable stratification of the autistic population as a precursor to targeted intervention through dietary supplementation or specialized diet. Currently, clinical trials of common therapies such as vitamin and mineral supplements, carnitine, and gluten-free casein-free diets apply these therapies to all participants. Metabotyping subjects prior to treatment and monitoring metabolite levels provides the opportunity to assess patient compliance and response and to adjust treatment based on objective measurement of the metabolic profiles of individual subjects. It is likely that this strategy would substantially improve positive treatment outcomes.

This study does have some limitations. The levels of blood plasma amine metabolites are not directly relatable to brain

levels (35), making direct association of changes in plasma levels with changes in brain levels difficult. The CAMP study focused on recruitment of a large sample of children with ASD and age-matched TYP control subjects. Logistical and financial constraints precluded our ability to recruit a large enough sample of children with developmental delays without ASD. Thus, the specificity of AADMs for ASD relative to other neurodevelopmental disorders is currently unclear. This is an important issue that will need to be resolved in future studies. In addition, longitudinal samples are not available to analyze whether AADMs are stable over time. Finally, this study lacks animal models or tissue samples that could be used to dissect enzymatic and expression analysis to identify the molecular mechanisms underlying AADMs. While we cannot explain the alterations in metabolism, we have demonstrated that our approach provides stratification of subjects for which future studies and perhaps targeted treatments could be carried out.

This study demonstrates one approach to analyzing the metabolism of ASD to successfully identify reproducible metabotypes. Analysis of the CAMP study samples is ongoing, and there will be additional metabotypes that will be diagnostic for subsets of children with ASD. Stratifying ASD based on metabotypes offers an opportunity to identify efficacious interventions within metabotypes that can lead to more precise and individualized treatment. The hope is that by combining the established metabotypes into a more comprehensive diagnostic system, a substantial percentage of children at risk for ASD will be identifiable at a very early age.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the National Institutes of Health (Grant No. NIH 5R44MH10712403 [to REB] and Grant No. 1R01MH103371 [to DGA]). Other financial support was provided by the Nancy Lurie Marks Family Foundation and the Robert E. and Donna Landreth Family Fund.

We are grateful to Dana Kelly and Gina Wentling for administrative and social media support of subject recruitment; DelRay Sugden for subject sample intake, sample handling, and integrity quality control; and Lindsay Feuling for development and maintenance of study subject and sample databases. We are particularly indebted to the families and children who participated in the CAMP study; without their collaboration, this study would not have been possible.

AMS, MAL, and REB are employees of, JJK and PRW were employees of, and ELRD is an equity owner in Stemina Biomarker Discovery Inc. AMS, JJK, PRW, MAL, ELRD, and REB are inventors on provisional patent application 62/623,153 titled "Amino Acid Analysis and Autism Subsets" filed January 29, 2018. DGA receives research funding from the National Institutes of Health, the Simons Foundation, and Stemina Biomarker Discovery Inc. He is on the scientific advisory boards of Stemina Biomarker Discovery Inc. and Axial Therapeutics.

ClinicalTrials.gov: Children's Autism Metabolome Project; https://clinicaltrials.gov/ct2/show/NCT02548442; NCT02548442.

ARTICLE INFORMATION

From Stemina Biomarker Discovery Inc. (AMS, JJK, PRW, MAL, ELRD, REB), Madison, Wisconsin, and The MIND Institute and Department of Psychiatry and Behavioral Sciences (DGA), University of California, Davis, Davis, California.

The Children's Autism Metabolome Project (CAMP) investigators group includes David G. Amaral from The MIND Institute (University of California, Davis), Amanda E. Bennett from Children's Hospital of Philadelphia, Daniel L. Coury from Nationwide Children's Hospital, Richard E. Frye from The University of Arkansas for Medical Sciences, Raun D. Melmed from The Melmed Center, Ann Neumeyer from Massachusetts General Hospital, Kevin B. Sanders from Vanderbilt University Medical Center, and Logan K. Wink and Craig Erickson from Cincinnati Children's Hospital. We appreciate the CAMP investigators and their staff's contributions to making CAMP a successful study.

REB and DGA contributed equally to this work.

Address correspondence to David G. Amaral, Ph.D., The MIND Institute, University of California, Davis, 2825 50th Street, Sacramento, CA 95817; E-mail: dgamaral@ucdavis.edu.

Received Mar 16, 2018; revised Aug 3, 2018; accepted Aug 22, 2018.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.biopsych.2018.08.016.

REFERENCES

1. American Psychiatric Association (2013): Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Arlington, VA: American Psychiatric Publishing.
2. Baio J, Wiggins L, Christensen DL, Maenner MJ, Daniels J, Warren Z, *et al.* (2018): Prevalence of autism spectrum disorder among children aged 8 years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. MMWR Surveill Summ 67:1–23.
3. El Achkar CM, Spence SJ (2015): Clinical characteristics of children and young adults with co-occurring autism spectrum disorder and epilepsy. Epilepsy Behav 47:183–190.
4. Spence SJ, Schneider MT (2009): The role of epilepsy and epileptiform EEGs in autism spectrum disorders. Pediatr Res 65:599–606.
5. Kerns CM, Kendall PC, Berry L, Souders MC, Franklin ME, Schultz RT, *et al.* (2014): Traditional and atypical presentations of anxiety in youth with autism spectrum disorder. J Autism Dev Disord 44:2851–2861.
6. Buie T, Campbell DB, Fuchs GJ 3rd, Furuta GT, Levy J, Vandewater J, *et al.* (2010): Evaluation, diagnosis, and treatment of gastrointestinal disorders in individuals with ASDs: A consensus report. Pediatrics 125(Suppl 1):S1–S18.
7. Bourgeron T (2016): Current knowledge on the genetics of autism and propositions for future research. C R Biol 339:300–307.
8. Sandin S, Lichtenstein P, Kuja-Halkola R, Hultman C, Larsson H, Reichenberg A (2017): The heritability of autism spectrum disorder. JAMA 318:1182–1184.
9. Geschwind DH, State MW (2015): Gene hunting in autism spectrum disorder: On the path to precision medicine. Lancet Neurol 14: 1109–1120.
10. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, *et al.* (2015): Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. Neuron 87:1215–1233.
11. Mandy W, Lai MC (2016): Annual research review: The role of the environment in the developmental psychopathology of autism spectrum condition. J Child Psychol Psychiatry 57:271–292.
12. Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, *et al.* (2000): The Autism Diagnostic Observation Schedule–Generic: A standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 30:205–223.
13. Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A (2006): Autism from 2 to 9 years of age. Arch Gen Psychiatry 63:694–701.
14. Centers for Disease Control and Prevention (2009): Prevalence of autism spectrum disorders—Autism and Developmental Disabilities Monitoring Network, United States, 2006. MMWR Surveill Summ 58(10):1–20.
15. Dawson G (2010): Recent advances in research on early detection, causes, biology, and treatment of autism spectrum disorders. Curr Opin Neurol 23:95–96.
16. Dawson G, Rogers S, Munson J, Smith M, Winter J, Greenson J, *et al.* (2010): Randomized, controlled trial of an intervention for toddlers with autism: The Early Start Denver model. Pediatrics 125:e17–e23.
17. Ganz ML (2007): The lifetime distribution of the incremental societal costs of autism. Arch Pediatr Adolesc Med 161:343–349.
18. McPartland JC (2017): Developing clinically practicable biomarkers for autism spectrum disorder. J Autism Dev Disord 47:2935–2937.

19. Hazlett HC, Gu H, Munsell BC, Kim SH, Styner M, Wolff JJ, *et al.* (2017): Early brain development in infants at high risk for autism spectrum disorder. Nature 542:348–351.
20. Shen MD, Kim SH, McKinstry RC, Gu H, Hazlett HC, Nordahl CW, *et al.* (2017): Increased extra-axial cerebrospinal fluid in high-risk infants who later develop autism. Biol Psychiatry 82:186–193.
21. West PR, Amaral DG, Bais P, Smith AM, Egnash LA, Ross ME, *et al.* (2014): Metabolomics as a tool for discovery of biomarkers of autism spectrum disorder in the blood plasma of children. PLoS One 9:e112445.
22. Beger RD, Dunn W, Schmidt MA, Gross SS, Kirwan JA, Cascante M, *et al.* (2016): Metabolomics enables precision medicine: "A White Paper, Community Perspective." Metabolomics 12:149.
23. Żurawicz E, Kałużna-Czaplińska J (2015): Analysis of amino acids in autism spectrum disorders. Trends Analyt Chem 73:91–118.
24. Novarino G, El-Fishawy P, Kayserili H, Meguid NA, Scott EM, Schroth J, *et al.* (2012): Mutations in BCKD-kinase lead to a potentially treatable form of autism with epilepsy. Science 338:394–397.
25. Tarlungeanu DC, Deliu E, Dotter CP, Kara M, Janiesch PC, Scalise M, *et al.* (2016): Impaired amino acid transport at the blood brain barrier is a cause of autism spectrum disorder. Cell 167:1481–1494.e1418.
26. Riedl A, Gieger C, Hauner H, Daniel H, Linseisen J (2017): Metabotyping and its application in targeted nutrition: An overview. Br J Nutr 117:1631–1644.
27. Benjamini Y, Hochberg Y (1995): Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Statist Soc B Statist Methodol 57:289–300.
28. Suzuki R, Shimodaira H (2006): Pvclust: An R package for assessing the uncertainty in hierarchical clustering. Bioinformatics 22:1540–1542.
29. Stacklies W, Redestig H, Scholz M, Walther D, Selbig J (2007): pcaMethods—A bioconductor package providing PCA methods for incomplete data. Bioinformatics 23:1164–1167.
30. Eastman JW, Sherwin JE, Wong R, Liao CL, Currier RJ, Lorey F, *et al.* (2000): Use of the phenylalanine:tyrosine ratio to test newborns for phenylketonuria in a large public health screening programme. J Med Screen 7:131–135.
31. Garcia-Cazorla A, Oyarzabal A, Fort J, Robles C, Castejon E, Ruiz-Sala P, *et al.* (2014): Two novel mutations in the BCKDK (branched-chain keto-acid dehydrogenase kinase) gene are responsible for a neurobehavioral deficit in two pediatric unrelated patients. Hum Mutat 35:470–477.
32. Zheng HF, Wang WQ, Li XM, Rauw G, Baker GB (2017): Body fluid levels of neuroactive amino acids in autism spectrum disorders: A review of the literature. Amino Acids 49:57–65.
33. Park Y, Reyna-Neyra A, Philippe L, Thoreen CC (2017): mTORC1 balances cellular amino acid supply with demand for protein synthesis through post-transcriptional control of ATF4. Cell Rep 19:1083–1090.
34. Maynard TM, Manzini MC (2017): Balancing act: Maintaining amino acid levels in the autistic brain. Neuron 93:476–479.
35. Scholl-Burgi S, Haberlandt E, Heinz-Erian P, Deisenhammer F, Albrecht U, Sigl SB, *et al.* (2008): Amino acid cerebrospinal fluid/plasma ratios in children: Influence of age, gender, and antiepileptic medication. Pediatrics 121:e920–e926.