Evidence for Metabolomic Phenotypes Based on Analysis of Plasma from the Autism Phenome Project Cohort

A. Smith[1], R. Burrier[1], J. King[1], P. West[1], S. Rogers[2], D. Li[2], D. Amaral[2], and E. Donley[1]
[1]Stemina Biomarker Discovery Inc., 504 S. Rosa Rd., Suite 150, Madison WI 53719
[2]University of California, Davis M.I.N.D. Institute

Overview

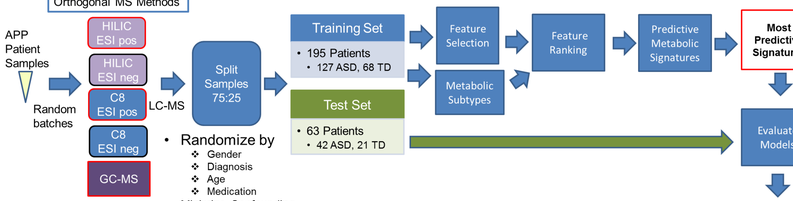
PURPOSE: Discover subtypes of ASD as defined by metabolic features.

METHODS: MS-based metabolomic analysis with heuristic and multivariate data analysis.

RESULTS: Multiple metabolic subtypes can discriminate ~ 45% ASD subjects from TD.

Introduction

Metabolomics can identify predictive and actionable biomarker profiles from a child's inherited biochemistry as well as the interactions of the gut microbiome with dietary and environmental factors that may contribute to ASD. Identification of metabolic profiles in children with ASD creates an opportunity to develop metabolic based diagnostics that enable early diagnosis and elucidate biochemical changes associated with ASD.



Methods

Subject Samples

- Diagnosis of autism using ADOS-G and ADI-R and criteria from the Collaborative Programs of Excellence in Autism
- TD children included if developmental scores were within 2 standard deviations of the mean on all subscales of the MSEL. TD exclusion criteria included mental retardation, pervasive developmental disorder, language impairment or other developmental, neurological, or behavioral problems. TD children were screened and excluded for autism with the Social Communication Questionnaire.
- Non-fasted blood was obtained in ACD tubes and the plasma was stored at -80°C.

Patient demographics

Demographic		TD	ASD	Overall
Group Size		93	180	273
Sex (male %)		69	83	78
Age (y)	Ave ± S.D.	3.0±0.4	3.1±0.5	3.0±0.5
DQ	Ave ± S.D.	106.4±11.9	62.5±20.8	77.3±27.8

Sample Preparation and Mass Spectrometry

- Small molecules extracted from plasma using 8:1 methanol:water solution at -20°C.
- C8 and HILIC LC-HRMS using positive and negative polarity was performed using an Agilent G6520 QTOF.
- GC-MS was performed at the West Coast Metabolomics Center (UC-Davis)

Experimental Workflow

Results

Feature Contribution by Analytical Platform

Platform	Features	After QC	P-value <= 0.05
HILIC +	2629	653	113
HILIC -	2364	565	71
C8 +	758	301	42
C8 -	736	235	66
GCMS	378	378	22
Total	6865	2132	314

Methods were optimized so that each of the analytical methods contributed to the overall orthogonal analytical approach. After QC, there were 314 features that were carried forward for classification modeling.

Classification Results for Both Training and Test Sample Sets

Three computational modeling methods (PLS-DA, RF, SVM) were used to identify the best performing feature subset and model method. Random forest yielded the best performing models having good accuracy, excellent sensitivity, and marginal specificity. The best performing model was chosen based on prediction of the validation set.

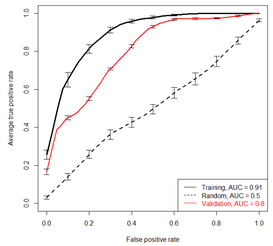
Training Set Results

Model	Feature No.	Accuracy	Sensitivity	Specificity	AUC
RF	120	0.82	0.98	0.52	0.91

Independent Test Set Results

Model	Feature No.	Accuracy	Sensitivity	Specificity	AUC
RF	120	0.79	0.84	0.48	0.80

ROC Analysis for Training and Validation Data Sets



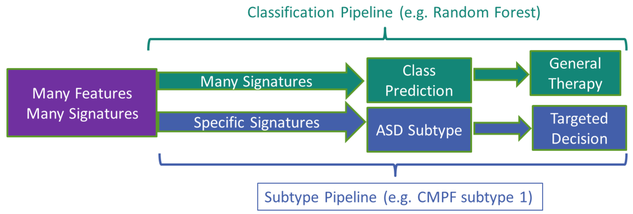
ROC analysis of the top performing model. The training set results (**black, solid**) are based on the average of the prediction from the hold out samples of 5-fold cross validation repeated 5 times. The validation data (**red**) were generated from the prediction of the validation set samples using 5 random seeds. A null model (**black, dotted**) was created by randomizing diagnosis to demonstrate that the ASD vs TD classification results were not obtained by chance.

Confirmed Feature Annotations from Most Predictive Metabolic Signature

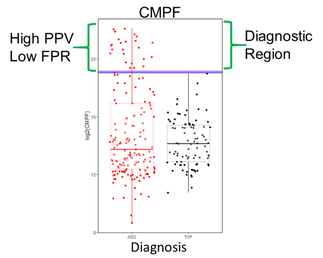
Metabolite	LC-MS Method	Fold (ASD/TD)	P-Value	FDR
2-Hydroxy-2-methylbutyric acid	C8 ESIneg	0.84	0.006	0.282
3-Methyl-2-oxovaleric acid	C8 ESIneg	0.87	0.013	0.412
Salicylic acid	C8 ESIneg	0.77	0.002	0.174
Gentisic acid	C8 ESIneg	0.71	0.000	0.036
CMPF related metabolite	C8 ESIneg	8.43	0.046	0.637
DHEA sulfate	C8 ESIneg	2.63	0.001	0.127
Pregnenolone sulfate	C8 ESIneg	1.71	0.000	0.029
LysoPE(22:6)	C8 ESIpos	1.38	0.000	0.028
Glycine	HILIC ESIpos	1.34	0.000	0.069
L-Alanine or Sarcosine	HILIC ESIpos	1.17	0.005	0.255
Proline betaine	HILIC ESIpos	0.72	0.000	0.069

Outcomes from Computational Modeling Applied to Discovery Metabolomic Workflow

- Biochemical differences are present in ASD that may be useful in understanding the disorder.
- Identified metabolites have previously been associated with ASD whereas others novel.
- Predictive signatures contained many features and many of these exhibit small changes.
- Procedure reveals metabolites with potential to diagnose ASD but few with immediately actionable results.
- Two class models may have limited effectiveness in modeling a heterogeneous disorder like ASD.
- Subtype analysis was conducted to overcome the limitations of the mean-based feature selection .



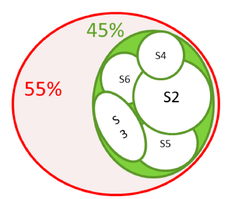
Multiple Potential Metabolic Subtypes of ASD Identified



- **Identify Features** where the distribution of abundance of ASD and TD do not overlap.
- **Identify a threshold** that produces a high PPV in a small population of ASD.
- **Evaluate Candidates** in an independent sample set to confirm.
- Develop **Quantitative Assay** for clinical applications.

Feature-Defined Subtypes in APP

Subtype Biomarker	Prevalence in APP	PPV
CMPF	*14%*	*100%*
S2	23%	98%
S3	16%	90%
S4	13%	82%
S5	10%	81%
S6	5%	100%
All 6 combined, diagnose **51.4%** of ASD subjects in APP		
The 5 without CMPF, diagnose **45.3%** of ASD subjects in APP		



5 metabolic features can identify 45% of ASD subjects.

- Subtypes yield specific information on a biochemical pathway or metabolic abnormality
- Subtypes can be combined into a panel that increases population that can be diagnosed.
- Recently CMPF was found to be associated with fish oil supplementation.

Conclusions

- Computational modeling identifies a signature that can discriminate ASD from TD individuals with 79% accuracy.
- Metabolites include lysophospholipids, organic acids, hormone sulfates, furans, and amino acids some which have previously been associated with ASD.
- Metabolic subtypes be may present that can be utilized to create metabolic panels for diagnosis of ASD.
- The subtype approach yields more actionable results that computational modeling.
- Larger study populations will allow detection of less prevalent metabolic subtypes.

Future plans include:

1) Working to assign chemical structures to subtype related features.
2) Validate metabolic subtypes of ASD in the CAMP Study.
3) Develop subtypes into a panel of diagnostic tests which may enable personalized treatment.
4) Establish CLIA service to provide access to metabolic panels

Children's Autism Metabolome Project (CAMP) Clinical Study

- A clinical study powered for successful identification of subtypes
- Recruiting 1500 children (ages 18-48 months)
  - 500 Autism Spectrum Disorder
  - 500 Developmental Delay
  - 500 Typically Developing
- ADOS-2, DSM-V, and MSEL as behavioral measures
- Enrollment underway at 8 centers located in CA, OH, TN, AR, AZ, MA,

Acknowledgments

- Agilent Technologies for LC-MS Instrumentation and Software Support

Funded By:
NIMH and Nancy Lurie Marks Family Foundation