(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 November 2024 (28.11.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/243201 A2

(51) International Patent Classification:
Not classified

(21) International Application Number:
PCT/US2024/030358

(22) International Filing Date:
21 May 2024 (21.05.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/468,057 22 May 2023 (22.05.2023) US

(71) Applicants: **THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK** [US/US]; 412 Low Memorial Library, 535 W. 116th Street, New York, NY 10027 (US). **THE REGENTS OF THE UNIVERSITY OF CALIFORNIA** [US/US]; 1111 Franklin St., Twelfth Floor, Oakland, CA 94607-5200 (US). **NORWEGIAN INSTITUTE OF PUBLIC HEALTH** [NO/NO]; Lovisenberggata 8, 0456 Oslo (NO).

(72) Inventors: **CHE, Xiaoyu**; 739 Galda Rd., Paramus, NJ 07652 (US). **LIPKIN, Walter, Ian**; 45 W. 105th St., New York, NY 10025 (US). **FIEHN, Oliver**; UC Davis West Coast Metabolomics Center, UC, Davis, Genome Center, Room 1313, 451 Health Sci. Drive, Davis, CA 95616 (US). **STOLTENBERG, Camilla**; Krokusveien 6, 0875 Oslo (NO).

(74) Agent: **AMOS, Brian, J.**; Amster, Rothstein & Ebenstein LLP, 405 Lexington Avenue, Floor 48, New York, NY 10174 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

WO 2024/243201 A2

(54) Title: A METHOD OF PROVIDING EARLY INTERVENTION TO A NEWBORN OR INFANT WITH AUTISM SPECTRUM DISORDER

(57) Abstract: Described is use of biomarkers found in maternal mid-gestation (MMG) and cord blood (CB) plasma for early identification of children with autism spectrum disorder (ASD) or at risk for ASD. Using these biomarkers, newborns and infants with ASD or at high risk of ASD can be identified and interventions for ASD, which have been shown to be effective early in life, can be provided to a child as young as a newborn.

A METHOD OF PROVIDING EARLY INTERVENTION TO A NEWBORN OR INFANT WITH AUTISM SPECTRUM DISORDER

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit of U.S. Provisional Application No. 63/468,057, filed May 22, 2023, the contents of which are hereby incorporated by reference.

STATEMENT OF GOVERNMENT SUPPORT

[0002] This invention was made with government support under NS047537 and NS086122 awarded by the National Institutes of Health. The government has certain rights in the invention.

BACKGROUND

[0003] Approximately 1 in 44 children in the United States have an autism spectrum disorder (ASD) (1). Interventions using applied behavior analysis (ABA) have been proven to be most effective when implemented early (2, 3). However, the mean age for diagnosis is age 4-5 years (4), and the median age varies from 36 to 63 months (1). A recent study has shown that earlier intensive, individualized intervention led to greater improvement in ASD outcomes and further suggested a window of 18 months versus 27 months for the intervention can impact the outcomes. This study shows the need for early identification of children with ASD (124). Accordingly, identification of biomarkers for early diagnosis is a high priority. The pathogenesis of ASD may include genetic, environmental, and epigenetic factors (5).

[0004] It is hypothesized that clues to pathogenesis and biomarkers for diagnosis can be revealed through metabolomic analysis of samples from maternal mid-gestation (MMG) and children's cord blood (CB). The few studies of prenatal metabolomic profiling in ASD are inconsistent. One study of MMG sera reported dysregulations in lipid metabolism, pyrimidine metabolism, N-glycan pathway, and C21-steroid hormone biosynthesis (6). Another using maternal sera obtained during the 3rd trimester reported alterations in levels of fatty acid metabolites in the prostaglandin pathway (7). In a third, maternal dyslipidemia and increased levels of serum branched-chain amino acids (BCAA) were associated with an increased risk in boys (8). Metabolomic studies of neonatal samples are typically conducted using blood spots rather than umbilical CB (9, 10). There are no investigations that report metabolomic profiling of both MMG and CB plasma from the same subjects.

[0005] Thus, there is a need for markers to make an early diagnosis of ASD in order to provide intervention as early as possible, as early as birth.

SUMMARY

[0006] Described herein are results of metabolomic profiling in MMG and CB plasma drawn from the population-based Norwegian Autism Birth Cohort (ABC) (11). The findings described herein indicate that dysregulated MMG and CB metabolomic profiles are associated with ASD. These dysregulated profiles provide additional evidence for earlier work linking risk to inflammation during gestation as well as new findings consistent with disrupted membrane integrity, impaired neurotransmission, and neurotoxicity. Moreover, the imbalance of ether/non-ether phospholipids in the MMG of ASD girls may provide insight into the higher frequency of cognitive impairment in girls than in boys with ASD.

[0007] Disclosed herein are biomarkers that can be found in MMG and CB plasma, for early identification of children at risk for ASD. Using these biomarkers, intervention such as those utilizing ABA, which have been shown to be effective early in life, can be provided to a child as young as a newborn.

[0008] One embodiment of the current disclosure is a method for providing early intervention to newborn with autism spectrum disorder (ASD) or at risk for developing ASD comprising the steps of:

identifying that the newborn has ASD or is at risk of developing ASD by:

performing or having performed an ASD association assay on a sample, to identify if the newborn has ASD or is at risk of developing ASD;

identifying the newborn as having ASD or being at risk for developing ASD when there is a dysregulation of the newborn's metabolomic profile consistent with inflammation; and

providing early ASD intervention to the newborn.

[0009] In some embodiments, the newborn is a female. In some embodiments, the newborn is a male.

[0010] In some embodiments, the sample is blood from the mother at mid-gestation. In some embodiments, the sample is cord blood of the newborn.

[0011] In some embodiments, the metabolomic profile is compared to a control metabolomic profile. In some embodiments, the control is a newborn, infant or child not suffering from ASD.

[0012] In some embodiments where the newborn is a male and the sample is blood from the mother at mid-gestation, a decreased or lower level of the compound 17-hydroxy-4,7,10,13,15,19-docosahexaenoic acid or chemical cluster hydroxy eicosapentaenoic acid (HEPE) identifies the newborn as having a dysregulated metabolomic profile consistent with inflammation and as having ASD or at risk for developing ASD.

[0013] In some embodiments where the newborn is a female and the sample is blood from the mother at mid-gestation, an increased or higher ratio of arachidonic acid (AA)/eicosapentaenoic acid (EPA) identifies the newborn as having a dysregulated metabolomic profile consistent with inflammation and as having ASD or at risk for developing ASD.

[0014] In some embodiments where the sample is cord blood, an increased or higher level of the compound AA or ratio of AA/ Docosahexaenoic acid (DHA) identifies the newborn as having a dysregulated metabolomic profile consistent with inflammation and as having ASD or at risk for developing ASD, whether the newborn is male or female.

[0015] In some embodiments where the newborn is a male and the sample is cord blood, a decreased or lower level of chemical cluster epoxyeicosatrienoic acid (EpETrE) identifies the newborn as having a dysregulated metabolomic profile consistent with inflammation and as having ASD or at risk for developing ASD.

[0016] A further embodiment of the current disclosure is a method for providing early intervention to newborn with autism spectrum disorder (ASD) comprising the steps of:

identifying that the newborn has ASD or is at risk of developing ASD by:

performing or having performed an ASD association assay on a sample to identify if the newborn has ASD or is at risk of developing ASD; and

identifying the newborn as having ASD or being at risk for developing ASD when there is a dysregulation of the newborn's metabolomic profile consistent with disrupted membrane integrity; and

providing early ASD intervention to the newborn.

[0017] In some embodiments, the newborn is a female. In some embodiments, the newborn is a male.

[0018] In some embodiments, the sample is blood from the mother at mid-gestation. In some embodiments, the sample is cord blood of the newborn.

[0019] In some embodiments, the metabolomic profile is compared to a control metabolomic profile. In some embodiments, the control is a newborn, infant, or child not suffering from ASD.

[0020] In some embodiments where the sample is blood from the mother at mid-gestation, a decreased or lower level of chemical cluster phosphatidylcholines (PC), identifies the newborn as having a dysregulated metabolomic profile consistent with disrupted membrane integrity and as having ASD or at risk for developing ASD, whether the newborn is male or female.

[0021] In some embodiments where the newborn is a female and the sample is blood from the mother at mid-gestation, a decreased or lower level of compounds PC or phosphatidylethanolamines (PE) or chemical cluster PE, or an increased or higher level of chemical clusters PC-ether or PC ether-vlc identifies the newborn as having a dysregulated metabolomic profile consistent with disrupted membrane integrity and as having ASD or at risk for developing ASD.

[0022] In some embodiments where the sample is cord blood, an increased or higher level of chemical cluster ceramide identifies the newborn as having a dysregulated metabolomic profile consistent with disrupted membrane integrity and as having ASD or at risk for developing ASD, whether the newborn is male or female.

[0023] In some embodiments where the newborn is a male and the sample is cord blood, a decreased or lower level of chemical cluster PC, or an increased or higher level of chemical clusters carnitine, long chain ceramides or unsaturated ceramides identifies the newborn as having a dysregulated metabolomic profile consistent with disrupted membrane integrity and as having ASD or at risk for developing ASD.

[0024] In some embodiments where the newborn is a female and the sample is cord blood, an increased or higher level of chemical cluster sphingomyelin, identifies the newborn as having a dysregulated metabolomic profile consistent with disrupted membrane integrity and as having ASD or at risk for developing ASD.

[0025] Yet a further embodiment of the current disclosure is a method for providing early intervention to newborn with autism spectrum disorder (ASD) comprising the steps of:

identifying that the newborn has ASD or is at risk of developing ASD by:

performing or having performed an ASD association assay on a sample to identify if the newborn has ASD or is at risk of developing ASD;

identifying the newborn as having ASD or being at risk for developing ASD when there is a dysregulation of the newborn's metabolomic profile consistent with impaired neurotransmission and neurotoxicity; and

providing early ASD intervention to the newborn.

[0026] In some embodiments, the newborn is a female. In some embodiments, the newborn is a male.

[0027] In some embodiments, the sample is blood from the mother at mid-gestation. In some embodiments, the sample is cord blood of the newborn.

[0028] In some embodiments, the metabolomic profile is compared to a control metabolomic profile. In some embodiments, the control is a newborn, infant or child not suffering from ASD.

[0029] In some embodiments where the sample is blood from the mother at mid-gestation, an increased or higher ratio of Glu/Gln, identifies the newborn as having a dysregulated metabolomic profile consistent with impaired neurotransmission and neurotoxicity and as having ASD or at risk for developing ASD, whether the newborn is male or female.

[0030] In some embodiments where the newborn is a female and the sample is blood from the mother at mid-gestation, an increased or higher level of compound Glu, identifies the newborn as having a dysregulated metabolomic profile consistent with impaired neurotransmission and neurotoxicity and as having ASD or at risk for developing ASD.

[0031] In some embodiments where the newborn is a male and the sample is blood from the mother at mid-gestation, a decreased or lower level of compound Gln, identifies the newborn as having a dysregulated metabolomic profile consistent with impaired neurotransmission and neurotoxicity and as having ASD or at risk for developing ASD.

[0032] In some embodiments where the newborn is a male and the sample is cord blood, a decreased or lower level of Gln, or an increased or higher ratio of Glu/Gln, or chemical clusters long chain ceramides or unsaturated ceramides, identifies the newborn as having a dysregulated metabolomic profile consistent with impaired neurotransmission and neurotoxicity and as having ASD or at risk for developing ASD.

[0033] Yet a further embodiment of the current disclosure is a method for providing early intervention to newborn with autism spectrum disorder (ASD) comprising the steps of:
identifying that the newborn has ASD or is at risk of developing ASD by:
performing or having performed an ASD association assay on a sample to identify if the newborn has ASD or is at risk of developing ASD;
identifying the newborn as having ASD or being at risk for developing ASD when there is a dysregulation of the newborn's metabolomic profile; and
providing early ASD intervention to the newborn.

[0034] In some embodiments, the newborn is a female. In some embodiments, the newborn is a male.

[0035] In some embodiments, the sample is blood from the mother at mid-gestation. In some embodiments, the sample is cord blood of the newborn.

[0036] In some embodiments, the metabolomic profile is compared to a control metabolomic profile. In some embodiments, the control is a newborn, infant or child not suffering from ASD.

[0037] In some embodiments where the newborn is a female and the sample is blood from the mother at mid-gestation, an increased or higher level of compounds pyroglutamic acid, propionic acid, N-methylalanine, pseudouridine, PC (p-40:3)/PC (o-40:4), or HexCer-NS, or chemical clusters galactosylceramides or alanine and derivatives, identifies the newborn as having a dysregulated metabolomic profile and as having ASD or at risk for developing ASD.
In some embodiments where the newborn is a female and the sample is blood from the mother at mid-gestation, a decreased or lower level of compounds 2,6-dihydroxybenzoic acid, or chemical clusters polyunsaturated fatty acid-containing phosphatidylcholine (PC-PUFA) or hydroxybenzoates, identifies the newborn as having a dysregulated metabolomic profile and as having ASD or at risk for developing ASD.

[0038] In some embodiments where the newborn is a male and the sample is blood from the mother at mid-gestation, an increased or higher level of compounds homo-gamma-linolenic acid or oxidized phosphatidylcholine (OxPC), or chemical clusters adipates or unsaturated fatty acids, identifies the newborn as having a dysregulated metabolomic profile consistent and as having ASD or at risk for developing ASD.

[0039] In some embodiments where the newborn is a male and the sample is blood from the mother at mid-gestation, a decreased or lower level of chemical clusters hydroxy fatty acid_22_6_1 (OH-FA_22_6_1) or basic amino acids, identifies the newborn as having a dysregulated metabolomic profile consistent and as having ASD or at risk for developing ASD.

[0040] In some embodiments where the newborn is a female and the sample is cord blood, an increased or higher level of compounds orotic acid, 2'-O-methylguanosine or eicosatrienoic acid, or chemical cluster purine nucleosides, identifies the newborn as having a dysregulated metabolomic profile and as having ASD or at risk for developing ASD.

[0041] In some embodiments where the newborn is a female and the sample is cord blood, a decreased or lower level of chemical clusters triacylglycerols or the majority of unsaturated triglycerides, identifies the newborn as having a dysregulated metabolomic profile and as having ASD or at risk for developing ASD.

[0042] In some embodiments where the newborn is a male and the sample is cord blood, an increased or higher level of compounds erythritol, alanine, glucose-6-phosphate, pseudouridine, methionine, succinic acid, 2'-deoxyadenosine-5'-monophosphate, or glycerophosphocholine, or chemical clusters aspartic acids, pyrimidinones, laurates, alanine and derivatives, malates, or sugar alcohols, identifies the newborn as having a dysregulated metabolomic profile consistent and as having ASD or at risk for developing ASD.

[0043] In some embodiments where the newborn is a male and the sample is cord blood, a decreased or lower level of compounds 1,4-cyclohexanedione or leukotriene B4, or chemical clusters of nitro compounds, dihydroxyeicosatrienoic acid (DiHETrE), epoxyoctadecadienoic acid (EpODE), PCs, or the majority of hexoses, identifies the newborn as having a dysregulated metabolomic profile consistent and as having ASD or at risk for developing ASD.

[0044] In a further embodiment of the current disclosure is a method for providing early intervention to newborn with autism spectrum disorder (ASD) comprising the steps of:

identifying that the newborn has ASD or is at risk of developing ASD by:

performing or having performed an ASD association assay on a sample to identify if the newborn has ASD or is at risk of developing ASD;

identifying the newborn as having ASD or being at risk for developing ASD when there is a dysregulation of the newborn's metabolomic profile, wherein the metabolomic profile is determined using one or more of the metabolites listed in Table 7; and
providing early ASD intervention to the newborn.

[0045] In some embodiments, the newborn is a female. In some embodiments, the newborn is a male.

[0046] In some embodiments, the sample is blood from the mother at mid-gestation. In some embodiments, the sample is cord blood of the newborn.

[0047] In some embodiments, the metabolomic profile is compared to a control metabolomic profile. In some embodiments, the control is a newborn, infant or child not suffering from ASD.

[0048] In a further embodiment of the current disclosure is a method for providing early intervention to a male newborn with autism spectrum disorder (ASD) comprising the steps of:
identifying that the newborn has ASD or is at risk of developing ASD by:
performing or having performed an ASD association assay on a sample to identify if the newborn has ASD or is at risk of developing ASD, wherein the sample is cord blood;
identifying the newborn as having ASD or being at risk for developing ASD when there is high number of dysregulated chemical clusters;
providing early ASD intervention to the newborn.

[0049] Lastly a further embodiment of the current disclosure is a method for providing early intervention to a female newborn with autism spectrum disorder (ASD) with increased impaired cognitive development comprising the steps of:
identifying that the newborn has ASD or is at risk of developing ASD with increased impaired cognitive development by:
performing or having performed an ASD association assay on a sample to identify if the newborn has ASD or is at risk of developing ASD with increased impaired cognitive development, wherein the sample is blood from the mother at mid-gestation;
identifying the newborn as having ASD or being at risk for developing ASD with increased impaired cognitive development when there is an imbalance of ether/non-ether phospholipids; and
providing early ASD intervention to the newborn.

[0050] As used herein, a control, e.g. a control level or amount or metabolomic profile, is a value or values decided or obtained, usually beforehand, from a sample(s) from a non-afflicted subject(s) as a reference. The concept of a control is well-established in the field, and can be determined, in a non-limiting example, empirically from non-afflicted subjects (versus afflicted subjects, including afflicted subjects having different grades of the relevant affliction), and may be normalized as desired (in non-limiting examples, for volume, mass, age, location, gender) to negate the effect of one or more variables. In embodiments, the control level or amount or metabolomic profile are determined from samples from subjects without ASD.

[0051] In all of the foregoing embodiments, an ASD association assay can be performed by any methods known in the art to quantify compounds and/or metabolites from a sample, including a biological sample, including blood.

[0052] In all of the foregoing embodiments, the intervention can continue as the newborn ages to an infant and then to a child.

BRIEF DESCRIPTION OF THE DRAWINGS

[0053] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0054] For the purpose of illustrating the invention, there are depicted in drawings certain embodiments of the invention. However, the invention is not limited to the precise arrangements and instrumentalities of the embodiments depicted in the drawings.

[0055] FIGS. 1A-1D. Chemical enrichment analyses reveal sex-specific altered chemical clusters in autism spectrum disorders (ASD). A bar restricted to the left of the centered vertical line indicates a metabolic cluster that is lower in ASD. A bar restricted to the right of the centered vertical line indicates a metabolic cluster that is higher in ASD. A bar that crosses the vertical line indicates a metabolic cluster that is dysregulated in mixed directions. Fig. 1A shows MMG female. Fig. 1B shows MMG male. Fig. 1C shows CB female. Fig. 1D shows CB male.

[0056] FIGS. 2A-2D. Autism spectrum disorders (ASD) predictive modeling. Fig. 2A shows ROC curves for MMG girls. Fig. 2B shows ROC curves for MMG boys. Fig. 2C shows ROC curves for CB girls. Fig. 2D shows ROC curves for CB boys. In each of the sex- and sample type-stratified

datasets, four machine learning algorithms were used: Lasso; AdaLasso; RF, and XGBoost. Feature selection of metabolites as predictors was conducted using Bayesian analysis and MX knockoffs. For each of the machine learning algorithms, four sets of predictors were considered: *Set 1*, all metabolites; *Set 2*, metabolites with BF > 3; *Set 3*, metabolites that were selected by MX knockoffs in more than one iteration; *Set 4*, metabolites that were selected by MX knockoffs in at least one iteration and had BF > 1. The predictive performance was evaluated in 500 iterations of random resampling CV with 80/20 training/testing split.

[0057] FIGS. 3A-3C. Compounds, chemical clusters, and metabolite ratios implicated in inflammation (Fig. 3A), membrane integrity (Fig. 3B), and neurotransmission and neurotoxicity (Fig. 3C). Blue arrows indicate elevated, and pink arrows indicate decreased levels of compounds, chemical clusters, and metabolite ratios.

[0058] FIG. 4. Pipeline for sample selection. MoBa, the Norwegian Mother, Father, and Child Cohort study, ABC the Norwegian Autism Birth Cohort study.

[0059] FIGS. 5A-5B Batch effect correction. Fig. 5A shows PCA scatter plots before correction for batch effect. Fig. 5B shows PCA scatter plots after correction for batch effect. Using scatter plots of the principal components, we detected a batch effect in the complex lipids data from MMG plasma samples that was linked to equipment failure resulting in a one-month delay in completing the lipidomic analyses. Comparisons of variance between batches suggested that mere shifting the means or medians was not sufficient; thus, we implemented a Bayesian framework to correct for the batch effect.

DETAILED DESCRIPTION

[0060] Autism spectrum disorder (ASD) is a neurological and developmental disorder that affects how people interact with others, communicate, learn, and behave. Although autism can be diagnosed at any age, it is described as a "developmental disorder" because symptoms generally appear in the first 2 years of life.

[0061] According to the *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*, a guide created by the American Psychiatric Association that health care providers use to diagnose mental disorders, people with ASD often have difficulty with communication and interaction with other

people, restricted interests and repetitive behaviors, and symptoms that affect their ability to function in school, work, and other areas of life.

[0062] Autism is known as a "spectrum" disorder because there is wide variation in the type and severity of symptoms people experience.

[0063] While much is not known about ASD, what is known is that the earlier intervention is provided to a child with ASD, the better the outcomes. Disclosed herein are methods of identifying newborns with ASD or at risk for developing ASD such that interventions can be started as early as at birth.

[0064] Early ASD interventions as contemplated herein are known in the art and include clinician-implemented, and caregiver/patent-implemented. Early ASD interventions can lead to improvements in developmental, adaptive behavior measures, and/or social skills, and include, for example, individual or group Early Social Interaction (ESI) model. In embodiments, by early identification, the methods permit application of early ASD interventions prior to 30 months of age, prior to 20 months of age, or prior to 10 months of age. Early intervention can focus or encompass one or more of: physical skills, thinking skills, communication skills, social skills, and emotional skills. Early intervention can also employ speech therapy, and/or hearing impairment therapy /treatment, and/or physical therapy.

[0065] In embodiment, the newborn is male. In embodiment, the newborn is female.

Definitions

[0066] The terms used in this specification generally have their ordinary meanings in the art, within the context of this invention and the specific context where each term is used. Certain terms are discussed below, or elsewhere in the specification, to provide additional guidance to the practitioner in describing the methods of the invention and how to use them. Moreover, it will be appreciated that the same thing can be said in more than one way. Consequently, alternative language and synonyms may be used for any one or more of the terms discussed herein, nor is any special significance to be placed upon whether or not a term is elaborated or discussed herein. Synonyms for certain terms are provided. A recital of one or more synonyms does not exclude the use of the other synonyms. The use of examples anywhere in the specification, including examples of any terms discussed herein, is illustrative only, and in no way limits the scope and meaning of

the invention or any exemplified term. Likewise, the invention is not limited to its preferred embodiments.

[0067] As used herein, the term "sample" means any substance containing or presumed to contain compounds and/or chemicals and/or metabolites. The sample can be a sample of tissue or fluid isolated from a subject including but not limited to, plasma, serum, whole blood, spinal fluid, semen, amniotic fluid, lymph fluid, synovial fluid, urine, tears, blood cells, organs, and tissue. One sample is blood.

[0068] The terms "healthy control" as used herein is a human subject who is not suffering from autism spectrum disorder or at risk for autism spectrum disorder. In addition, a healthy control can be aged matched to the subject being tested, and not suffering from other diseases or conditions.

[0069] The term "newborn" as used herein refers to a baby from birth to about two months of age.

[0070] The term "infant" as used herein means a baby from about one month to about one year in age.

Abbreviations

ASD	Autism Spectrum Disorder
ID	Intellectual Deficiency
ABC	Norwegian Autism Birth Cohort
MoBa	Norwegian Mother, Father, and Child Cohort
NPR	Norwegian Patient Register
MMG	maternal mid-gestation
CB	children's cord blood
ABA	applied behavior analysis
AdaLasso	adaptive Lasso
AUC	area under the curve value
BF	Bayesian factor
CV	cross-validation
Lasso	least absolute shrinkage and selection operator
MX	Model X
PCA	principal component analysis

RF	Random Forests
SD	standard deviation
DiHETrE	dihydroxyeicosatrienoic acid
EpETrE	epoxyeicosatrienoic acid
EpODE	epoxyoctadecadienoic acid
HEPE	hydroxy eicosapentaenoic acid
OH-FA_22_6_1	hydroxy fatty acid_22_6_1
PC-ether	ether-linked phosphatidylcholine
PC-ether-PUFA	polyunsaturated fatty acid-containing ether-linked phosphatidylcholine
PC-ether-vlc	very long-chain ether-linked phosphatidylcholine
PC-PUFA	polyunsaturated fatty acid-containing phosphatidylcholine
DHA	Docosahexaenoic acid
Glu	glutamate
Gln	glutamine
EPA	Eicosapentaenoic acid
AA	arachidonic acid
EPA	eicosapentaenoic acid
PC	phosphatidylcholine
PE	phosphatidylethanolamine
CL	complex lipid
PM	primary metabolites
BA	biogenic amines
OL	oxylipins
ETrE	eicosatrienoic acid
SM	sphingomyelins
LPC	saturated lysophosphatidylcholines

Tables

[0071] Table 1. Subject characteristics

[a]For continuous characteristics, t-test 2-sided p-value; For categorical characteristics, chi-squared 2-sided p-value. [b]Parental education missing: MMG, n=35; CB, n=27. [c]Folic acid supplements: maternal report of folate exposure between 0-8 weeks of gestation, no missing value. [d]Gestational age missing: MMG, n=4; CB, n=3. [e]Birth weight missing: MMG, n=2; CB, n=1. [f]Intellectual deficiency: We do not have information on ID within controls. *ASD* autism spectrum disorders, *CB* cord blood, *F* female, *ID* intellectual deficiency, *M* male, *MMG* maternal-mid-gestation, *N/A* not available, *SD* standard deviation.

			MMG Plasma (n=408)			CB Plasma (n= 418)		
Subject Characteristics		**Sex**	**ASD Cases** n=203 45 F/158 M	**Non-Cases** n=205 47 F/158M	***p-value***[a]	**ASD Cases** n=204 49 F/155 M	**Non-Cases** n=214 50 F/164 M	***p-value***[a]
Maternal Characteristics								
Maternal age; **mean (SD)**		F	29.7 (5.4)	29.4 (4.1)	0.792	29.7 (5.2)	29.5 (4.0)	0.802
		M	29.7 (4.8)	30.2 (5.2)	0.392	29.9 (4.8)	30.4 (4.7)	0.353
Parental education[b]; **n (%)**	<12 years	F	3 (6.7%)	2 (4.3%)	0.928	4 (8.2%)	3 (6.0%)	0.956
		M	9 (5.7%)	6 (3.8%)	0.611	8 (5.2%)	6 (3.7%)	0.199
	12 years	F	7 (15.6%)	9 (19.1%)		7 (14.3%)	8 (16.0%)	
		M	47 (29.7%)	46 (29.1%)		46 (29.7%)	35 (21.3%)	
	13-16 years	F	14 (31.1%)	15 (31.9%)		16 (32.7%)	17 (34.0%)	
		M	56 (35.4%)	51 (32.3%)		55 (35.5%)	55 (33.5%)	
	≥17 years	F	15 (33.3%)	18 (38.3%)		16 (32.7%)	19 (38.0%)	
		M	37 (23.4%)	46 (29.1%)		36 (23.2%)	52 (31.7%)	
Folic acid supplements[c]; **n(%)**		F	23 (51.1%)	30 (63.8%)	0.217	24 (49.0%)	32 (64.0%)	0.132
		M	85 (53.8%)	100 (63.3%)	0.087	82 (52.9%)	104 (63.4%)	0.057
Child Characteristics								
Birth year; **n(%)**	2000	F	3 (6.7%)	0 (0.0%)	0.142	3 (6.1%)	0 (0.0%)	0.084
		M	4 (2.5%)	4 (2.5%)	0.150	4 (2.6%)	3 (1.8%)	0.010
	2001	F	1 (2.2%)	0 (0.0%)		2 (4.1%)	0 (0.0%)	

		M	12 (7.6%)	5 (3.2%)		10 (6.5%)	5 (3.0%)	
	2002	F	7 (15.6%)	5 (10.6%)		7 (14.3%)	7 (14.0%)	
		M	22 (13.9%)	10 (6.3%)		21 (13.5%)	8 (4.9%)	
	2003	F	12 (26.7%)	6 (12.8%)		13 (26.5%)	6 (12.0%)	
		M	32 (20.3%)	27 (17.1%)		31 (20.0%)	25 (15.2%)	
	2004	F	6 (13.3%)	10 (21.3%)		7 (14.3%)	9 (18.0%)	
		M	26 (16.5%)	32 (20.3%)		23 (14.8%)	21 (12.8%)	
	2005	F	8 (17.8%)	7 (14.9%)		8 (16.3%)	8 (16.0%)	
		M	23 (14.6%)	25 (15.8%)		23 (14.8%)	21 (12.8%)	
	2006	F	2 (4.4%)	5 (10.6%)		2 (4.1%)	6 (12.0%)	
		M	22 (13.9%)	28 (17.7%)		25 (16.1%)	40 (24.4%)	
	2007	F	5 (11.1%)	7 (14.9%)		6 (12.2%)	6 (12.0%)	
		M	12 (7.6%)	22 (13.9%)		13 (8.4%)	31 (18.9%)	
	2008	F	1 (2.2%)	7 (14.9%)		1 (2.0%)	8 (16.0%)	
		M	5 (3.2%)	4 (2.5%)		5 (3.2%)	8 (4.9%)	
	2009	F	0 (0.0%)	0 (0.0%)		0 (0.0%)	0 (0.0%)	
		M	0 (0.0%)	1 (0.6%)		0 (0.0%)	2 (1.2%)	
Birth season; **n(%)**	Winter (Dec-Feb)	F	11 (24.4%)	13 (27.7%)	0.889	13 (26.5%)	13 (26.0%)	0.825
		M	38 (24.1%)	47 (29.7%)	0.049	38 (24.5%)	46 (28.0%)	0.020
	Spring (Mar-May)	F	18 (40.0%)	18 (38.3%)		20 (40.8%)	22 (44.0%)	
		M	33 (20.9%)	36 (22.8%)		34 (21.9%)	46 (28.0%)	
	Summer (Jun-Aug)	F	10 (22.2%)	8 (17.0%)		10 (20.4%)	7 (14.0%)	
		M	31 (19.6%)	41 (25.9%)		31 (20.0%)	42 (25.6%)	
	Fall (Sept-Nov)	F	6 (13.3%)	8 (17.0%)		6 (12.2%)	8 (16.0%)	
		M	56 (35.4%)	34 (21.5%)		52 (33.5%)	30 (18.3%)	
Gestational age[d]; **n(%)**	< 37 weeks	F	3 (6.7%)	0 (0.0%)	0.081	3 (6.1%)	0 (0.0%)	0.185
		M	12 (7.6%)	4 (2.5%)	0.110	12 (7.7%)	2 (1.2%)	0.015
	37 - 42 weeks	F	42 (93.3%)	44 (93.6%)		45 (91.8%)	47 (94.0%)	
		M	127 (80.4%)	137 (86.7%)		124 (80.0%)	144 (87.8%)	
	≥42 weeks	F	0 (0.0%)	2 (4.3%)		1 (2.0%)	2 (4.0%)	
		M	17 (10.8%)	16 (10.1%)		18 (11.6%)	17 (10.4%)	
Birth weight[e]; **n(%)**	<1,500 g	F	0 (0.0%)	0 (0.0%)	0.505	0 (0.0%)	0 (0.0%)	0.459
		M	3 (1.9%)	0 (0.0%)	0.075	3 (1.9%)	0 (0.0%)	0.077
		F	2 (4.4%)	0 (0.0%)		2 (4.1%)	0 (0.0%)	

	1,500 - 2,500 g	M	6 (3.8%)	1 (0.6%)		5 (3.2%)	1 (0.6%)	
	2,500 - 4,000 g	F	38 (84.4%)	40 (85.1%)		41 (83.7%)	41 (82.0%)	
		M	106 (67.1%)	109 (69.0%)		106 (68.4%)	110 (67.1%)	
	≥4,000 g	F	5 (11.1%)	7 (14.9%)		6 (12.2%)	9 (18.0%)	
		M	42 (26.6%)	47 (29.7%)		41 (26.5%)	52 (31.7%)	
Intellectual Deficiency[f]; **n(%)**		F	12 (26.7%)	N/A	-	12 (24.6%)	N/A	-
		M	32 (20.3%)	N/A	-	29 (18.7%)	N/A	-

[0072] Table 2. Regression and Bayesian analyses

Adjusted logistic regression models with Bayesian inference were used to test for an association between each metabolite and ASD risk, separately for boys and girls and within each sample type. In MMG analyses, models were adjusted for maternal age, illnesses (fever, infection, inflammatory, autoimmune, allergic disorders), emotional distress scores (SCL-5), and non-NSAID antipyretic medications in pregnancy up until sample acquisition, as well as gestational age at MMG blood draw. In CB analyses, models were adjusted for maternal age, illnesses (fever, infection, inflammatory, autoimmune, allergic disorders), emotional distress scores (SCL-5), and non-NSAID antipyretic medications in pregnancy. Multiple comparisons were corrected using the Benjamini-Hochberg procedure, controlling the overall FDR at the level of 0.05. In the Bayesian analysis, we considered a metabolite to be associated with ASD risk if BayesFactor (BF) > 10 and 95% HDIs of the estimated aOR did not overlap with 1. Because we used weakly informative priors in Bayesian analysis, the 95% HDIs were similar to the 95% CIs. aOR adjusted odds ratio, *ASD* autism spectrum disorders, *CB* cord blood, *CI* confidence interval, *FDR* false discovery rate adjusted p-value, *HDI* highest density credible intervals, *MMG* maternal mid-gestation.

.

MMG Girls						
Panel	**Metabolite**	**aOR**	**95% CI**	***p-value***	**FDR**	**BayesFactor**
Biogenic Amines	2,6-Dihydroxybenzoic acid	0.388	(0.213, 0.707)	0.002	0.533	39.673
	Glutamic acid - BA	2.295	(1.343, 3.921)	0.002	0.533	34.655
	Pyroglutamic acid	2.121	(1.265, 3.556)	0.004	0.533	23.775
	Propionic acid	2.165	(1.154, 4.060)	0.016	0.533	21.482

	N-Methylalanine	2.210	(1.266, 3.855)	0.005	0.533	17.279
	Pseudouridine - BA	2.105	(1.236, 3.586)	0.006	0.533	10.659
Complex Lipids	PC (p-40:3)/PC (o-40:4)	2.709	(1.360, 5.399)	0.005	0.533	118.302
	TAG (56:9)	0.452	(0.235, 0.869)	0.017	0.533	19.200
	PE (36:5); PE (16:0-20:5)	0.424	(0.246, 0.732)	0.002	0.533	18.499
	PC (40:8) - ESI (+)	0.435	(0.238, 0.797)	0.007	0.533	17.088
	PC (37:5)	0.454	(0.254, 0.811)	0.008	0.533	15.595
	HexCer-NS (d34:1)	2.245	(1.286, 3.921)	0.004	0.533	12.344
MMG Boys						
Panel	**Metabolite**	**aOR**	**95% CI**	***p-value***	**FDR**	**BayesFactor**
Complex Lipids	FA (20:3) (homo-gamma-linolenic acid)	1.548	(1.229, 1.950)	0.000	0.251	65.460
	OxPC (34:2+1O); OxPC (16:0-18:2+1O)	1.523	(1.157, 2.004)	0.003	0.483	18.754
Oxylipins	17-hydroxy-4,7,10,13,15,19-docosahexaenoic acid	0.667	(0.523, 0.850)	0.001	0.483	10.980
CB Girls						
Panel	**Metabolite**	**aOR**	**95% CI**	***p-value***	**FDR**	**BayesFactor**
Primary Metabolites	Orotic acid - PM	2.214	(1.161, 4.224)	0.016	0.743	11.582
Biogenic Amines	2'-O-Methylguanosine	2.739	(1.299, 5.775)	0.008	0.743	23.019
Complex Lipids	DAG (36:1)	2.425	(1.419, 4.145)	0.001	0.743	69.630
	TAG (56:9)	0.445	(0.266, 0.745)	0.002	0.743	19.417
	TAG (56:3)	2.298	(1.298, 4.068)	0.004	0.743	18.639
	TAG (56:8) A	0.489	(0.292, 0.818)	0.006	0.743	13.941
	TAG (58:10)	0.472	(0.281, 0.792)	0.004	0.743	13.223
	FA (20:3) (eicosatrienoic acid)	2.139	(1.269, 3.605)	0.004	0.743	11.598
CB Boys						
Panel	**Metabolite**	**aOR**	**95% CI**	***p-value***	**FDR**	**BayesFactor**
Primary Metabolites	Erythritol	1.641	(1.289, 2.089)	0.000	0.035	242.458
	Alanine - PM	1.601	(1.209, 2.121)	0.001	0.116	26.894
	Glucose-6-phosphate - PM	1.504	(1.184, 1.911)	0.001	0.116	18.599
	Pseudouridine - PM	1.804	(1.226, 2.655)	0.003	0.138	13.556

	Methionine - PM	1.489	(1.172, 1.893)	0.001	0.116	13.064
	Succinic acid	1.452	(1.147, 1.837)	0.002	0.116	11.444
Biogenic Amines	1,4-Cyclohexanedione	0.651	(0.509, 0.834)	0.001	0.100	19.679
	2'-Deoxyadenosine-5'-monophosphate	1.692	(1.262, 2.269)	0.000	0.100	18.289
	Glutamine - BA	0.668	(0.532, 0.839)	0.001	0.100	17.540
	Glycerophosphocholine	1.460	(1.153, 1.849)	0.002	0.102	13.239
Complex Lipids	FA (20:4) (arachidonic acid)	1.675	(1.322, 2.123)	0.000	0.023	786.450
Oxylipins	Leukotriene B4	0.640	(0.494, 0.830)	0.001	0.100	24.866

[0073] Table 3. Chemical enrichment analyses

Chemical enrichment analyses of the results from the logistic regression models were performed using ChemRICH to determine chemical clusters that were significantly altered between ASD and control groups. *ASD* autism spectrum disorders, *CB* cord blood, *DiHETrE* dihydroxyeicosatrienoic acid, *EpETrE* epoxyeicosatrienoic acid, *EpODE* epoxyoctadecadienoic acid, *HEPE* hydroxy eicosapentaenoic acid, *MMG* maternal mid-gestation, *OH-FA_22_6_1* hydroxy fatty acid_22_6_1, *PC-ether* ether-linked phosphatidylcholine, *PC-ether-PUFA* polyunsaturated fatty acid-containing ether-linked phosphatidylcholine, *PC-ether-vlc* very long-chain ether-linked phosphatidylcholine, *PC-PUFA* polyunsaturated fatty acid-containing phosphatidylcholine.

MMG Girls							
Cluster name	**Cluster size**	***p-value***	**Key compound**	**Altered**	**Increased**	**Decreased**	**Altered Ratio**
PC-PUFA	36	4.10E-11	PC (40:8) - ESI (+)	12	0	12	0.3
Galactosylceramides	6	3.10E-08	GlcCer (d42:2) - ESI (-)	5	5	0	0.8
PC-ether	23	9.80E-07	PC (o-34:0)	3	3	0	0.1

PC-ether-vlc	14	2.70E-06	PC (p-40:3)/PC (o-40:4)	6	6	0	0.4
Phosphatidylethanolamines	44	0.0033	PE (36:5); PE (16:0-20:5)	5	0	5	0.1
Phosphatidylcholines	53	0.0072	PC (37:2) - ESI (+)	10	0	10	0.2
PC-ether-PUFA	22	0.013	PC (p-42:4)/PC (o-42:5) - ESI (-)	5	5	0	0.2
Alanine and derivatives	6	0.014	N-Methylalanine	2	2	0	0.3
Hydroxybenzoates	6	0.024	2,6-Dihydroxybenzoic acid	2	0	2	0.3
MMG Boys							
Cluster name	**Cluster size**	***p-value***	**Key compound**	**Altered**	**Increased**	**Decreased**	**Altered Ratio**
1-Acyl-sn-glycero-3-phosphocholines	15	2.60E-04	LPC (18:2) - ESI (+)	4	2	2	0.3
Phosphatidylcholines	53	5.00E-04	OxPC (34:2+1O); OxPC (16:0-18:2+1O)	4	1	3	0.08
HEPE	3	0.0016	12-Hydroxy-5,8,10,14,17-eicosapentaenoic acid	2	0	2	0.7
Unsaturated fatty acids	26	0.005	FA (20:3) (homo-gamma-linolenic acid)	4	4	0	0.2
Saturated lysophosphatidylcholines	6	0.010	LPC (14:0) - ESI (-)	2	1	1	0.3
Basic amino acids	11	0.022	Asparagine - BA	4	0	4	0.4
Adipates	3	0.028	Adipic acid - BA	2	2	0	0.7
OH-FA_22_6_1	3	0.040	17-Hydroxy-4,7,10,13,15,19-docosahexaenoic acid	2	0	2	0.7
CB Girls							
Cluster name	**Cluster size**	***p-value***	**Key compound**	**Altered**	**Increased**	**Decreased**	**Altered Ratio**
Sphingomyelins	37	2.30E-04	SM (d42:2) B - ESI (-)	3	3	0	0.08
Unsaturated triglycerides	60	0.0012	TAG (56:3)	11	4	7	0.2
Purine nucleosides	8	0.014	2'-O-Methylguanosine	2	2	0	0.2
Ceramides	6	0.023	Ceramide (d42:2) A - ESI (+)	2	2	0	0.3
Triacylglycerols	45	0.045	TAG (56:9)	6	0	6	0.1

CB Boys							
Cluster name	**Cluster size**	***p-value***	**Key compound**	**Altered**	**Increased**	**Decreased**	**Altered Ratio**
Hexoses	7	5.20E-06	Mannose - BA	5	1	4	0.7
Unsaturated fatty acids	26	6.30E-06	FA (20:4) (arachidonic acid)	7	5	2	0.3
Long-chain ceramides	3	6.70E-06	Ceramide (d36:1) - ESI (+)	3	3	0	1
DiHETrE	5	1.70E-05	19,20-Dihydroxydocosa-4,7,10,13,16-pentaenoic acid	3	0	3	0.6
Carnitines	9	3.70E-05	Propionylcarnitine	2	2	0	0.2
Phosphatidylcholines	53	4.70E-05	PC (32:2) - ESI (-)	10	0	10	0.2
Ceramides	6	6.90E-05	Ceramide (d42:2) A - ESI (+)	5	5	0	0.8
Laurates	7	7.80E-04	Acylcarnitine (14:1)	4	4	0	0.6
Hexosephosphates	7	0.0014	Glucose-6-phosphate - PM	4	3	1	0.6
EpODE	7	0.0024	11-Hydroxy-14,15-epoxyeicosatrienoic acid	3	0	3	0.4
EpETrE	7	0.0065	12(13)-Epoxy-9,15-octadecadienoic acid	4	0	4	0.6
Unsaturated ceramides	16	0.0069	Ceramide (d36:1) - ESI (-)	3	3	0	0.2
Malates	4	0.010	Citramalic acid - PM	2	2	0	0.5
Sugar alcohols	14	0.018	Erythritol	5	5	0	0.4
Pyrimidinones	5	0.028	Uracil - BA	3	3	0	0.6
Dipeptides	26	0.030	Glycyl-leucine	5	4	1	0.2
Alanine and derivatives	6	0.040	Alanine - PM	3	3	0	0.5
Aspartic acid	3	0.047	Aspartic acid - PM	2	2	0	0.7
Nitro compounds	3	0.049	10-Nitrolinoleic acid	2	0	2	0.7

[0074] Table 4. Number of Outliers

Table 4 Number of outliers.

Outliers		ASD Cases	Non-Cases
MMG Male	Primary Metabolites	2	5
	Biogenic Amines	3	7
	Complex Lipids	2	4
	Oxylipins	8	5
MMG Female	Primary Metabolites	0	0
	Biogenic Amines	1	0
	Complex Lipids	0	0
	Oxylipins	3	2
CB Male	Primary Metabolites	4	3
	Biogenic Amines	6	9
	Complex Lipids	3	4
	Oxylipins	1	4
CB Female	Primary Metabolites	3	0
	Biogenic Amines	3	1
	Complex Lipids	1	1
	Oxylipins	3	1

ASD autism spectrum disorders, *CB* cord blood, *MMG* maternal mid-gestation.

[0075] Table 5. ASD associations with metabolite ratios

Adjusted logistic regression models were used to test for ASD association with ratios of metabolites. We analyzed ratios of AA/EPA and AA/DHA implicated in inflammation, and ratios of GLU/GLN implicated in neurotransmission. We consider a ratio to be associated with ASD risk (bold) if p-value < 0.05. AA arachidonic acid, aOR adjusted odds ratio, ASD autism spectrum disorders, CI confidence interval, DHA docosahexaenoic acid, EPA eicosapentaenoic acid, GLN glutamine, GLU glutamate.

Table 5 ASD is associated with differences in metabolite ratios.

MMG Girls				
Ratio	aOR	95% CI		p-value
AA/EPA	**1.893**	**1.133**	**3.162**	**0.015**
AA/DHA	1.552	0.936	2.573	0.088
GLU/GLN	**1.767**	**1.066**	**2.928**	**0.027**
MMG Boys				
Ratio	aOR	95% CI		p-value
AA/EPA	0.916	0.725	1.157	0.461
AA/DHA	1.028	0.817	1.292	0.815
GLU/GLN	**1.402**	**1.121**	**1.754**	**0.003**
CB Girls				
Ratio	aOR	95% CI		p-value
AA/EPA	1.511	0.933	2.446	0.093
AA/DHA	**1.720**	**1.072**	**2.759**	**0.025**
GLU/GLN	0.888	0.569	1.385	0.600
CB Boys				
Ratio	aOR	95% CI		p-value
AA/EPA	1.166	0.926	1.468	0.191
AA/DHA	**1.352**	**1.069**	**1.710**	**0.012**
GLU/GLN	**1.435**	**1.150**	**1.791**	**0.001**

[0076] Table 6. Chemical enrichment analyses for sex x metabolite interactions

[0077] In MMG and CB analyses, separately, we combined metabolomic data from boys and girls and repeated the adjusted logistic regression models including sex as a covariate, together with the interaction term between sex and each metabolite. ChemRICH analysis was conducted using the estimates of the sex-metabolite interaction terms. CB cord blood, MMG maternal mid-gestation, PC-ether ether-linked phosphatidylcholine, PC-ether-vl c very long-chain ether-linked phosphatidylcholine, PC-PUFA polyunsaturated fatty acid-containing phosphatidylcholine.

Table 6 Chemical enrichment analyses for sex x metabolite interactions.

MMG							
Cluster name	Cluster size	*p-value*	Key compound	Altered	Increased	Decreased	Altered Ratio
PC-PUFA	36	1.9E-14	PC (40:5) A - ESI (+)	16	16	0	0.4
Galactosylceramides	6	2.10E-07	GlcCer (d40:1) - ESI (-)	4	0	4	0.7
PC-ether-vic	14	1.7E-06	PC (p-38:4)/PC (o-38:5) B - ESI (-)	3	0	3	0.2
PC-ether	23	1.8E-06	PC (p-38:4)/PC (o-38:5) - ESI (-)	2	0	2	0.09
Phosphatidylinositols	15	0.0019	PI (32:0); PI (16:0-16:0)	2	2	0	0.1
Phosphatidylethanolamines	44	0.007	PE (36:5); PE (16:0-20:5)	3	3	0	0.07
Glycosphingolipids	5	0.01	HexCer-NS (d40:1)	2	0	2	0.4
CB							
Cluster name	Cluster size	*p-value*	Key compound	Altered	Increased	Decreased	Altered Ratio
Hexosephosphates	7	0.0047	Galactosamine	4	3	1	0.6
Cyclohexanes	6	0.017	1,4-Cyclohexanedione	2	1	1	0.3

[0078] **Table 7**. Feature selection using MX knockoffs

[0079] MX knockoffs teases apart the important variables from noise in high-dimensional datasets while controlling the FDR. We used MX knockoffs with RF, and the procedure was repeated for 500 iterations controlling the FDR at the level of 0.1. The variable importance was measured as the number of iterations in which a metabolomic analyte was selected. CB cord blood, FDR false discovery rate, MMG maternal mid-gestation, MX Model-X, RF- Random Forests.

[0080]

Table 7 Feature selection using MX knockoffs.

MMG Girls		MMG Boys	
Metabolite	Frequency	Metabolite	Frequency
2,6-Dihydroxybenzoic acid	88	OxPC (34:2+1O); OxPC (16:0-18:2+1O)	299
PE (36:2) - ESI (+)	56	FA (20:3) (homo-gamma-linolenic acid)	76
SM (d36:0) - ESI (+)	28	Isomaltose A - BA	64
TAG (50:6)	20	Honokiol	47
PE (36:5); PE (16:0-20:5)	16	beta-Hydroxymyristic acid	47
14(15)-epoxy-5,8,11,17-eicosatetraenoic acid	15	FA (20:4) (arachidonic acid)	39
SM (d44:2) - ESI (-)	12	5'-Methylthioadenosine	36
TAG (48:4) A	11	PI (36:2); PI (18:0-18:2)	25
SM (d38:1) - ESI (-)	10	Adipic acid - BA	24
4-Pyridoxic acid	8	Glu-Gln	17
5,6-dihydroxyeicosa-8,11,14-trienoic acid	8	1,3 Cyclohexanedione	7
FA (24:0) (lignoceric acid)	7	Glycerophosphocholine	6
Ceramide (d42:2) A - ESI (-)	7	Imatinib	5
PC (33:2) - ESI (+)	6	TAG (50:3) B	5
PC (34:3) A	5	Asparagine - BA	3
11,12-Epoxyeicosa-5,8,14-trienoic acid	5	FA (24:1) (nervonic acid)	3
N-Methylhistidine	4	N-Methylalanine	2
PE (40:5); PE (18:0-22:5)	4	S-Adenosyl-homocysteine	2
PC (35:4) - ESI (+)	4	Benidipine	2
1-Methyl-L-histidine	3	CE (18:1)	2
PC (40:8) - ESI (+)	3	glucose-6-phosphate - PM	1
PC (37:5)	3	N-Acetylalanine	1
PC (37:3) A	3	Betaine	1
PC (32:2) - ESI (+)	3	Liquiritigenin	1
Glutamic acid - BA	2	Orotic acid - BA	1
Erythronolactone	2	Pipecolic acid	1
N-Acetylphenylalanine	2	Riboflavin	1
PC (36:5) D - ESI (+)	2	1,2-Benzenedicarboxylic acid	1
PC (34:3) B	2	TAG (51:5)	1
PE (34:2) - ESI (+)	2	TAG (60:4)	1
SM (d42:2) B - ESI (+)	2	PC (36:2) - ESI (-)	1
TAG (50:4)	2	SM (d42:2) B - ESI (+)	1
PC (34:4) - ESI (+)	2	FA (22:1) (erucic acid)	1
PE-Cer (d42:2)	2	DAG (34:2)	1
1-methylgalactose	1	PC (40:4) - ESI (-)	1
glycerol-alpha-phosphate	1	CE (20:2)	1
Isomaltose A - BA	1	TAG (53:0)	1
Glycyl-leucine	1	TAG (58:4)	1
Methylacetate	1		
Picrotin	1		
Acetaminophen glucuronide	1		
Kynurenine - BA	1		
N-Acetylcystidine	1		
PC (37:3) B	1		
PE (p-36:5)/PE (o-38:6) - ESI (-)	1		
PC (36:4) B - ESI (+)	1		
PE (34:2) - ESI (-)	1		
PC (p-38:4)/PC (o-38:5) B - ESI (-)	1		
PC (38:7) - ESI (+)	1		
FA (22:2) (docosadienoic acid)	1		
SM (d42:2) B - ESI (-)	1		
Ceramide (d38:1) - ESI (-)	1		
PC (p-40:1)/PC (o-40:2) - ESI (+)	1		
14-hydroxydocosa-4,7,10,12,16,19-hexaenoic acid	1		
5-Hydroxy-6,8,11,14-eicosatetraenoic acid	1		

[0081] Table 8. AUC values of predictive models

[0082] In each of the sex- and sample type-stratified datasets, we employed four machine learning algorithms: Lasso, AdaLasso, RF, and XGBoost. Feature selection of metabolites as predictors

was conducted using Bayesian analysis and MX knockoffs. For each of the machine learning algorithms, we considered four sets of predictors: ***Set 1***, all metabolites; ***Set 2***, metabolites with BF > 3; ***Set 3***, metabolites that were selected by MX knockoffs in more than one iteration; ***Set 4***, metabolites that were selected by MX knockoffs in at least one iteration and had BF > 1. The predictive performance was evaluated in 500 iterations of random resampling CV with 80/20 training/testing split. AdaLasso adaptive Lasso, AUC area under the curve value, CB cord blood, CI confidence interval, CV cross-validation, Lasso least absolute shrinkage and selection operator, MMG maternal mid-gestation, RF Random Forests.

CB Girls		CB Boys	
Metabolite	Frequency	Metabolite	Frequency
PC (34:0) - ESI (-)	55	erythritol	211
PE (34:3e); PE (16:1e/18:2)	38	Galactosamine	88
PE (p-34:2)/PE (o-34:3)	31	FA (20:4) (arachidonic acid)	87
2'-O-Methylguanosine	28	N-acetylornithine	67
DAG (36:1)	19	9-ketooctadeca-10,12-dienoic acid	57
Ala-Ala	16	FA (20:3) (eicosatrienoic acid)	56
PE (p-38:2)/PE (o-38:3) - ESI (+)	16	PC (p-34:2)/PC (o-34:3) - ESI (-)	56
8-Oxo-2-deoxyadenosine	9	TAG (60:3)	55
Phosphomycin	9	Glutamine - BA	53
Acylcarnitine (12:1)	9	PC (33:2) - ESI (-)	53
Ribose 5 - BA	8	hypoxanthine - PM	28
Nateglinide	8	Inosine - PM	28
LPE (20:4) - ESI (+)	8	Maltose - BA	22
Formoterol	7	Ceramide (d36:1) - ESI (+)	19
pseudouridine - PM	6	TAG (64:2)	17
DAG (36:2)	6	12,13-dihydroxyoctadeca-9,15-dienoic acid	17
Glucuronic acid A - BA	4	Ferulic acid	16
TAG (56:3)	4	Ceramide (d36:2)	16
2-deoxytetranic acid	3	3-(1-Pyrazolyl)-alanine	10
PC (p-42:5)/PC (o-42:6) - ESI (-)	3	Cyclo(Leu-Pro)	10
SM (d42:2) B - ESI (-)	3	2'-Deoxyadenosine-5'-monophosphate	9
PC (44:6e)	3	Leukotriene B4	9
glycerol	2	N-Methylalanine	8
Kaempferol-3-O-rutinoside	2	3-Hydroxyphenylacetic acid	7
N-Acetylcytidine	2	11,16-Dihydroxyicosa-5,8,14-trienoic acid	7
PI (32:1); PI (16:0-16:1)	2	oxoproline	6
PC (38:3) D - ESI (+)	2	Histidine - BA	6
TAG (58:2)	2	Acylcarnitine (10:0)	6
PI (34:1); PI (16:0-18:1)	2	Hypoxanthine - BA	6
PC (p-36:1)/PC (o-36:2) - ESI (-)	2	3'-O-Methylguanosine	5
PC (38:6) A - ESI (+)	2	Ceramide (d42:2) A - ESI (+)	5
PI (38:1); PI (18:0-18:1)	2	14,15-dihydroeicosa-5,8,11-trienoic acid	5
6-trans-Leukotriene B4	2	8-Oxo-2-deoxyadenosine	4
Carbobyquotan	1	N-alpha-(tert-Butoxycarbonyl)-L-histidine	4
2,6-Dihydroxybenzoic acid	1	18-[3-ethyloxiran-2-yl]octadeca-4,7,10,13,16-pentaenoic a	4
1,4-Cyclohexanedicarboxylic acid	1	14-hydroxydocosa-4,7,10,12,16,19-hexaenoic acid	4
Piroxicam	1	pseudouridine - PM	3
3-Indolepropionic acid	1	inulin	3
Methylgallate	1	glycerol-alpha-phosphate	3
Ergothioneine	1	1,4-Cyclohexanedione	3
Glucose-1-phosphate	1	Cinnamic acid - BA	3
Corticosterone	1	Fenfluramine B	3
Thr-Glu	1	PI (37:4)	3
Tomeloxacet	1	xanthine	2
TAG (56:5)	1	2-Hydroxyphenylacetic acid	2
TAG (54:8)	1	Benzphetamine	2
TAG (54:9)	1	Inosine - BA	2
PC (46:6) A - ESI (-)	1	PC 34:3	2
PC (p-34:1)/PC (o-34:2) - ESI (+)	1	TAG (54:8)	2
PE (p-38:5)/PE (o-38:6) - ESI (+)	1	12(13)-epoxy-9,16-octadecadienoic acid	2
PC (o-34:0)	1	12-Hydroxy-5,8,10,14-eicosatetraenoic acid	2
LPE (22:6)	1	aspartic acid - PM	1
SM (d44:2) - ESI (-)	1	sucrose	1
PC (34:3) B	1	glycerol	1
		adipic acid - PM	1
		cerotinic acid	1
		Glycerophosphocholine	1
		Tryptophan - BA	1
		Uracil - BA	1
		Glycyl-leucine	1
		Dienylaldehyde	1
		Inosine-5-monophosphate	1
		Docosahexaenoic acid	1
		Gly-val	1
		Orotic acid - BA	1
		N-Acetylglycine	1
		3-Methyladipic acid	1
		Ceramide (d34:1) - ESI (+)	1
		PC (32:2) - ESI (-)	1
		PC (36:2) - ESI (-)	1
		PC (36:5) A - ESI (+)	1
		FA (14:1) (physeteric acid)	1
		PE (36:2) - ESI (+)	1
		15(16)-epoxy-9,12-octadecadienoic acid	1
		Thromboxane B2	1
		Leukotriene B5	1

MX knockoffs tease apart the important variables from noise in high-dimensional datasets while controlling the FDR. We used MX knockoffs with RF, and the procedure was repeated for 500 iterations controlling the FDR at the level of 0.1. The variable importance was measured as the number of iterations in which a metabolomic analyte was selected. *CB* cord blood, *FDR* false discovery rate, *MMG* maternal mid-gestation, *MX* Model-X, *RF* Random Forests.

Markers of Autism Spectrum Disorder (ASD)

[0083] Described herein are results of metabolomic profiling in MMG and CB plasma drawn from the population-based Norwegian Autism Birth Cohort (ABC) (11). The findings described herein indicate that dysregulated MMG and CB metabolomic profiles are associated with ASD. These dysregulated profiles provide additional evidence for earlier work linking risk to inflammation during gestation as well as new findings consistent with disrupted membrane integrity, impaired neurotransmission, and neurotoxicity. Moreover, the imbalance of ether/non-ether phospholipids in the MMG of ASD girls may provide insight into the higher frequency of cognitive impairment in girls than in boys with ASD.

[0084] The discovery of prenatal and neonatal molecular biomarkers has the potential to yield insights into autism spectrum disorder (ASD) and facilitate early diagnosis. Metabolomic profiles in ASD were analyzed using plasma samples collected in the Norwegian Autism Birth Cohort from mothers at weeks 17-21 gestation (maternal mid-gestation, MMG, n=408) and from children on the day of birth (cord blood, CB, n=418). Associations were analyzed using sex-stratified adjusted logistic regression models with Bayesian analyses. Chemical enrichment analyses (ChemRICH) were performed to determine altered chemical clusters. Machine learning algorithms were also used to assess the utility of metabolomics as ASD biomarkers.

[0085] ASD was associated with a variety of chemical compounds including arachidonic acid, glutamate, and glutamine, and metabolite clusters including hydroxy eicospentaenoic acids, phosphatidylcholines, and ceramides in MMG and CB plasma that are consistent with inflammation, disruption of membrane integrity, and impaired neurotransmission and neurotoxicity. Girls with ASD have disruption of ether/non-ether phospholipid balance in the MMG plasma that is similar to that found in other neurodevelopmental disorders. ASD boys in the CB analyses had the highest number of dysregulated chemical clusters. These findings may provide new insights into the sex-specific differences in ASD and have implications for discovery of biomarkers that may enable early detection and intervention.

[0086] Specifically, the following dysregulation of compounds, ratios and chemical clusters were found in the MMG and CB of subjects with ASD in the study described herein:

Compounds and Ratios

Females-MMG

-increased or higher levels of compounds: glutamic acid; pyroglutamic acid; propionic acid; N-methylalanine; pseudouridine; PC (p-40:3)/PC (o-40:4),, HexCer-NS; AA/ EPA, and Glu/Gln

-decreased or lower levels of compounds: 2,6-dihydroxybenzoic acid; triglyceride (TAG); PC; and PE/PE

Females-CB

increased or higher levels of compounds: orotic acid; 2'-O-methylguanosine; EtrE; TAG; and diglyceride (DAG); and ratios of AA/ DHA

decreased or lower levels of TAGs

Males-MMG

-increased or higher levels of compounds homo-gamma-linolenic acid; and OxPC; and ratios of Glu/Gln

-decreased or lower levels of 17-hydroxy-4,7,10,13,15,19-docosahexaenoic acid

Males- CB

-increased or higher levels of compounds: AA; erythritol; alanine; glucose-6-phosphate; pseudouridine; methionine; succinic acid; 2'-deoxyadenosine-5'-monophosphate; and glycerophosphocholine; and ratios of AA/DHA; and Glu/Gln

-decreased or lower levels of compounds glutamine, 1,4-cyclohexanedione and leukotriene B4

Chemical Clusters

Females-MMG

-increased or higher levels of galactosylceramides, very long-chain ether-linked phosphatidylcholine (PC-ether-vlc), alanine and derivatives, polyunsaturated fatty acid-containing ether-linked phosphatidylcholine (PC-ether-PUFA), and ether-linked phosphatidylcholine (PC-ether)

-decreased or lower levels of polyunsaturated fatty acid-containing phosphatidylcholine (PC-PUFA), hydroxybenzoates, phosphatidylcholines (PC), and phosphatidylethanolamines (PE)

Females-CB

-increased or higher levels of ceramides, purine nucleosides, and sphingomyelins (SM)

-decreased or lower levels of triacylglycerols and the majority of unsaturated triglycerides

Males-MMG

-increased or higher levels of adipates and unsaturated fatty acids

-decreased or lower levels of hydroxy eicosapentaenoic acid (HEPE), hydroxy fatty acid_22_6_1 (OH-FA_22_6_1), basic amino acids, and the majority of phosphatidylcholines (PC)

mixed direction levels of 1-acyl-sn-glycero-3-phosphocholines, and saturated lysophosphatidylcholines (LPC)

Males- CB

-increased or higher levels of long-chain ceramides, ceramides, aspartic acids, pyrimidinones, laurates, alanine and derivatives, malates, sugar alcohols, unsaturated ceramides, carnitines, and the majority of dipeptides, unsaturated fatty acids, and hexosephosphates

-decreased or lower levels of nitro compounds, dihydroxyeicosatrienoic acid (DiHETrE), epoxyeicosatrienoic acid (EpETrE), epoxyoctadecadienoic acid (EpODE), PCs, and the majority of hexoses

[0087] Additionally, certain chemical clusters were found that were significantly different with regard to sex. In MMG, girls with ASD had lower levels of PC-PUFA, galactosylceramides, PC-ether-vlc, PC-ether, phosphatidylinositols (PI), PEs, and glycosphingolipids with no differences in

boys. In CB, two chemical clusters with sex differences in ASD association: hexosephosphates and cyclohexanes. The ASD associations with these metabolites were in opposite directions between boys and girls (increased in ASD boys but decreased in ASD girls; or decreased in ASD boys but increased in ASD girls).

[0088] MMG was a focus based on studies that indicate the early gestational environment is an important factor in ASD outcome: Swedish registry studies wherein first trimester use of the antiseizure medication valproic acid was associated with ASD risk (44), increased risk with thalidomide exposure at 20-24 days gestation (45) and earlier work in the ABC showing a robust protective effect of dietary folate in women who initiated supplementation before conception or during early gestation (43). Further support came from rodent models for ASD based on dam early-midgestational exposure to infection, stimulants of innate immunity or toxins (46, 47, 48).

[0089] The majority of papers describing metabolomic profiles in maternal plasma represent results from high-risk cohorts where the pathogenesis of ASD may be skewed towards specific genetic mechanisms (7, 49, 50, 51, 52) or environmental exposures (53). An exception is a population-based study in California where findings included dysregulations in lipid metabolism, pyrimidine metabolism, N-glycan pathway, and C21-steroid hormone biosynthesis (6).

[0090] The key findings shown herein in boys and girls with ASD include dysregulated MMG and CB metabolomic profiles consistent with inflammation (Fig. 3A), disruption of membrane integrity (Fig. 3B), and impaired neurotransmission and neurotoxicity (Fig. 3C). Girls with ASD also have disruption of ether/non-ether phospholipid balance in the MMG that is reminiscent of other neurodevelopmental disorders (54, 55, 56). The sample size was smaller for girls than boys; however, subsampling and *post hoc* power analyses suggested that the differences in findings between boys and girls were not due to differences in sample size.

Inflammation

[0091] A recent literature review of publications reporting analyses of more than 411 ASD subjects and 596 healthy controls through post-mortem histology and *in vivo* positron emission tomography revealed microglial activation in ASD (60). RNASeq analyses of human cortex from 47 ASD cases and 57 controls found evidence of microglial activation and type 1 interferon responses (61).

[0092] In MMG studies reported herein, analytes in the HEPE chemical cluster were decreased in ASD boys, but not in girls. ASD boys also had decreased levels of 17-hydroxy-4,7,10,13,15,19-docosahexaenoic acid compared to controls. HEPE and DHA are omega-3 (n-3) polyunsaturated fatty acids (PUFA) that are precursors to anti-inflammatory leukotrienes, resolvins and protectins and contribute to resolution of inflammation (62). Supplementation of HEPEs and DHAs reduce TNF-α expression (63). Studies in animal models have revealed that HEPEs regulate inflammation by induction of regulatory T cells (Tregs) (64), and that DHAs inhibit LPS-induced TNF-α production in macrophages (65).

[0093] In CB, AA was elevated in ASD boys, but not in girls. AA was also elevated (BF=4.21) in ASD boys in MMG but did not meet the criterion for BF > 10. AAs are omega-6 (n-6) PUFAs that are metabolized by cyclooxygenases (COX) and lipoxygenases (LOX) to produce pro-inflammatory eicosanoids (66). In human cell lines, AAs stimulate the c-jun amino-terminal kinase (JNK) cascade and NF-kB signaling (67), and promote TNF-α production (68). Levels of analytes in the EpETrE chemical cluster were reduced in ASD boys, but not in girls. EpETrEs are anti-inflammatory mediators derived from AAs through the action of cytochrome P450 (CYP450) epoxygenases (69). EpETrEs inhibit NF-κB signaling, prevent the accumulation of NF-κB subunit Rel A, and regulate TNF-α-induced inflammation in human endothelial cells (69).

[0094] A previous study reported imbalance in the n-6 PUFA/n-3 PUFA ratio in plasma of ASD subjects that is consistent with inflammation (39). Accordingly, ratios of AA/DHA and AA/EPA between ASD and controls were compared in each of the sex- and sample type-stratified datasets (Table 5). In MMG, AA/EPA ratios were elevated in ASD girls, but not in boys. Both boys and girls with ASD had elevated AA/DHA ratios in CB. Elevations in AA/EPA ratios are consistent with inflammation. Elevations in AA/DHA ratios are implicated in reactive oxygen species (ROS)-induced inflammation in hepatoma cell lines (70). Reductions in AA/DHA ratios inhibit TNF-α-induced inflammation in alveolar cells (71). In a recent plasma cytokine profile analysis with the same cohort reported here (72), we found prominent TNF-α elevations in the MMG and CB of both ASD boys and girls.

[0095] The metabolomic analyses set forth herein provide evidence of dysregulation in fatty acid metabolism that is consistent with lipid-mediated gestational inflammation.

Disrupted membrane integrity

[0096] Reduced levels of analytes in the PC chemical cluster were found in MMG in both ASD girls and boys, and in CB of ASD boys, but not girls. PCs are abundant in the membranes of cells and organelles, and regulate membrane integrity, transport, and G protein-coupled receptor (GPCR) signaling (73, 74). PC depletion affects the stability of the membrane protein translocases, impairs membrane integrity, and induces apoptosis (75, 76). In MMG, levels of analytes in the PE chemical cluster were reduced in ASD girls, but not in boys. PEs are integral membrane components essential for membrane fusion and stability (74). PE depletion disrupts membrane dynamics with altered membrane transport and receptor signaling (77), and impairs autophagy (78). In CB, we found elevated levels of analytes in the carnitine chemical cluster in ASD boys, but not in girls. Carnitines facilitate the transport of fatty acids across mitochondrial membranes (79). Previous studies reported elevations of short chain acyl-carnitines in plasma of ASD human subjects (80) and in brain tissues of ASD rodents (81) that are consistent with impaired membrane transport in ASD.

[0097] Levels of analytes in the ceramide chemical cluster were elevated in CB of both ASD boys and girls. The levels of metabolites in the long-chain ceramide and unsaturated ceramide chemical clusters were increased in CB of ASD boys, but not in girls. Ceramides are bioactive sphingolipids of biological membranes that regulate membrane packing and membrane fusion (82). Apart from controlling cellular signaling, these sphingolipids control cell fate (83). Ceramide accrual has been implicated in gestational complications including preeclampsia (84) and neural tube defects (85). As key components of the neuronal membranes, ceramides are critical to neuronal health (86). These sphingolipids induce neuronal apoptosis (87), and promote neurodegeneration (88). Levels of metabolites in the SM chemical cluster were elevated in the CB of ASD girls, but not in boys. SMs are phosphocholine derivatives of ceramides (89) that are integral components of membrane lipid microdomains or "lipid rafts" (90). SMs regulate exocytotic processes and membrane fusion (90). Accumulation of SMs disrupt maturation and closure of autophagic membranes and leads to impaired autophagy in Niemann-Pick type A (NPA) disease (91).

[0098] Intake of selective SSRIs is reported to affect levels of plasma lipid profiles (92, 93). Only a small number of ABC subjects received SSRIs during pregnancy (n=1 in MMG girls, n=7 in MMG boys, n=1 in CB girls, n=8 in CB boys). Repeated analyses excluding these subjects had no impact of group-specific chemical clustering (data not shown).

[0099] Significant sex differences in ASD associations with lipid chemical clusters in MMG were found. Whereas ASD versus control girls had lower levels of analytes in the PC-PUFA, PI, and PE chemical clusters, and higher levels of analytes in the PC-ether and PC-ether-vlc chemical clusters, ASD boys had no alterations in these clusters. PEs, PIs and PC-PUFAs are non-ether phospholipids, whereas PC-ethers and PC-ether-vlcs are ether phospholipids of peroxisomal origin (94). Ether and non-ether phospholipids are critical to membrane integrity and dynamics (74, 94). As key components of neuronal membranes, they regulate function (94). ASD girls exhibit lower cognitive ability than ASD boys (95). Elevations in ether phospholipids with concomitant reductions in non-ether phospholipids in the MMG plasma of ASD girls may indicate lipid remodeling consistent with a shift in the plasma lipidome toward ether phospholipids from the non-ether equivalents. An ether/non-ether phospholipid imbalance, with a shift toward ether phospholipids, is reported with de novo variations in fatty acyl-CoA reductase 1 (FAR1), Sjögren-Larsson syndrome, and complex hereditary spastic paraplegia due to phosphate cytidylyltransferase 2, ethanolamine (PCYT2) mutations (54, 55, 56).

[0100] We obtained information on the cognitive dysfunctions as reflected in the diagnosis of intellectual deficiency (ID) in ASD cases. Compared to ASD girls without ID, ASD girls with ID had decreased levels in MMG plasma of analytes in the PE cluster (altered ratio 0.16, 100% decreased, p-value < 0.0001), and increased levels of analytes in basic amino acid cluster (altered ratio 0.36, 75% increased, p-value < 0.001).

Impaired neurotransmission and neurotoxicity

[0101] In MMG, ASD girls, but not boys, had elevations in the levels of Glu. Levels of Gln were reduced in ASD boys (BF= 3.52), but not in girls, though it did not meet the criterion for BF > 10. In CB, ASD boys, but not girls, had reduced levels of Gln. Dysregulations in Glu/Gln metabolism with elevated plasma Glu and reduced plasma Gln are reported in children with ASD (96, 97). However, to our knowledge, this is the first evidence of dysregulations in Glu/Gln metabolism during gestation. Glu is an excitatory neurotransmitter that is critical in neuronal migration and plasticity, and synaptogenesis (98). Nonetheless, excess Glu leads to dysfunction of glutamatergic neurotransmission (98), and promotes excitotoxicity-induced neuronal apoptosis (99, 100). In a recent plasma cytokine profile analysis with the same cohort reported here, we found prominent TNF-α elevations in MMG plasma of ASD subjects (72). TNF-α stimulates Glu production and

promotes excitoneurotoxicity in murine microglial cells (101). TNF-α also increases blood-brain barrier permeability in animal models (102). Although Glu does not readily cross an intact blood-brain barrier (103), the observation that TNF-α is elevated may explain the correlation between elevated levels of Glu in plasma and in brain in ASD (104).

[0102] Gln is a precursor to both excitatory (Glu, aspartate) and inhibitory (γ-aminobutyric acid) neurotransmitters (105). Gln also has a neuroprotective role (106, 107). Given studies reporting elevated plasma Glu/Gln ratios in ASD (40, 41), we compared Glu/Gln ratios between ASDs and controls in each of the sex- and sample type-stratified dataset (Table 5). Glu/Gln ratios were elevated in both ASD boys and girls in MMG plasma. In the CB plasma, ASD boys, but not ASD girls, had elevated Glu/Gln ratios.

[0103] We previously reported that prenatal folic acid supplementation is associated with decreased risk of ASD (43). Folic acid has also been shown to reduce glutamate-induced excitotoxicity (108). Accordingly, we tested whether folic acid supplementation confounded results. The elevations of Glu/Gln ratios in ASD were not confounded by maternal intake of folic acid supplements (data not shown).

[0104] The analyses disclosed herein provide evidence for dysregulations in the Glu/Gln metabolism in gestation that are consistent with the neuropathology of ASD.

Testing for Biomarkers

[0105] As shown herein, certain metabolite and chemical markers and chemical clusters and ratios are associated with ASD. By using these markers, important predictions and determinations can be made regarding a newborn. Tests for these biomarkers can be performed at birth or during pregnancy. Blood or plasma can be obtained from cord blood at birth or from a routine pregnancy checkup of the mother.

[0106] The presence or amount of the markers can be compared to a reference value. In some embodiments, the reference value is from a healthy control, *i.e.,* a newborn, infant or child not suffering from an ASD. In some embodiments, the reference value is from a newborn, infant or child with an ASD.

[0107] In certain embodiments, a sample of fluid from a subject is obtained. In some embodiments, the subject is a newborn. In some embodiments, the subject is a pregnant mother. In some embodiments, the fluid is blood. In some embodiments, the fluid is plasma.

[0108] Any method known in the art can be used to test for the biomarkers in the sample, including but not limited to targeted bioactive oxylipin assay, and chromatography/mass spectrometry-based assays (MS) including but not limited to gas chromatography/time-of-flight mass spectrometry (GC-TOF MS), hydrophilic interaction liquid chromatography/quadrupole time-of-flight mass spectrometry (HILIC-QTOF MS), and liquid chromatography (LC)/quadrupole time-of-flight mass spectrometry (CSH-QTOF MS).

Early Intervention for Autism Spectrum Disorders

[0109] One major advantage of the present disclosure is it allows the identification or diagnosis of newborns who have or will develop or are at high risk of developing an ASD. This allows intervention at the earliest possible stages of development, which increases the success of the intervention.

[0110] Using the biomarkers disclosed herein, if a newborn is identified or diagnosed with an ASD, interventions can be provided to the newborn or to the newborn as they become an infant and beyond including but not limited to applied behavioral analysis (ABA), Early Social Interaction (ESI), floortime, occupational therapy, and pivotal response therapy.

[0111] Additionally, there are medical and physical conditions associated with ASD including but not limited to gastrointestinal (GI) issues, epilepsy, feeding issues, and sleep issues that can also be addressed more efficiently if the newborn has been identified or diagnosed with ASD.

Examples

[0112] The present invention may be better understood by reference to the following non-limiting examples, which are presented in order to more fully illustrate the preferred embodiments of the invention. They should in no way be construed to limit the broad scope of the invention.

Example 1 – Materials and Methods

Study subjects

[0113] The ABC (11), comprising 804 ASD cases and 1,786 randomly selected pool of control subjects is nested within the Norwegian Mother, Father, and Child Cohort (MoBa) (12, 13), a population-based pregnancy cohort comprising 114,473 children born in 1999-2009, with 95,244 mothers and more than 75,000 fathers (12).

ASD case ascertainment

[0114] Children with ASD were identified through questionnaire screening of mothers at offspring of 3, 5, and 7 years of age; referrals of participants suspected of ASD; and annual linkage to the Norwegian Patient Register (NPR) (14). Cases were defined according to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision* (DSM-IV-TR) (15) criteria. Children diagnosed at the ABC Study Clinic were assessed by clinical psychologists and child psychiatrists using standardized diagnostic instruments including the *Autism Diagnostic Interview-Revised* (ADI-R) (16) and the *Autism Diagnostic Observation Schedule* (ADOS) (17) and tests of intellectual and adaptive functioning and language capacity. ASD cases among NPR-identified children not evaluated at the ABC Study Clinic were those assigned *International Classification of Diseases, Tenth Revision* (ICD-10) F84 diagnoses (18). NPR ASD diagnoses have high validity with a positive predictive value of 96.7% (14).

Study sample selection

[0115] Inclusion criteria included singleton birth, continued participation in the cohort, survival to at least age three years, and availability of 200 microliters or more of MMG and/or umbilical CB plasma. We prioritized subjects with both MMG and CB samples for the metabolomics analyses. For each sex- and sample type-stratified group, we included all available MMG and CB samples from ASD cases that satisfied the inclusion criteria, and the control samples were randomly selected from the pool of eligible controls (Fig. 4).

Covariates

[0116] Data on covariates were extracted from MoBa questionnaires. For MMG analyses, data from the questionnaires covered the time period during pregnancy up to the MMG blood draw; for CB analyses, data covered the entire pregnancy. Confounders included maternal age; maternal report of fever; respiratory or other infection; autoimmune/allergic disorders; emotional distress

ratings (Hopkins Symptom Check List) (SCL-5) (19); and use of antipyretics (ATC codes N02BE01, N02BA01, N02BA51, N02BB51). For MMG analyses, we adjusted for gestational age at the MMG blood draw. For both MMG and CB analyses, we conducted sensitivity analyses based on covariates of parental education, maternal exposure to folic acid supplements between 0-8 weeks of gestation, and maternal reports of intake of selective serotonin reuptake inhibitors (SSRIs).

MMG and CB plasma collection

[0117] MMG plasma samples were collected during the routine ultrasound examinations at 18 weeks of gestation. Umbilical CB plasma samples were collected at birth. Samples were stored at −80°C (13, 20).

Metabolomics assays

[0118] For MMG and CB analyses, samples were sequentially run in an order that alternated based on ASD outcome and sex. Untargeted metabolomics data were acquired using three chromatography/mass spectrometry-based assays (MS): (1) Primary metabolites such as mono- and disaccharides, hydroxyl- and amino acids were measured by gas chromatography/time-of-flight mass spectrometry (GC-TOF MS) (21) including data alignment and compound annotation using the BinBase database algorithm (22); (2) Biogenic amines including microbial compounds such as trimethylamine N-oxide (TMAO), methylated and acetylated amino acids, and short di- and tripeptides were measured by hydrophilic interaction liquid chromatography/quadrupole time-of-flight mass spectrometry (HILIC-QTOF MS); (3) Complex lipids including phosphoglycerolipids, triacylglycerides, sphingolipids, and free fatty acids were analyzed by liquid chromatography (LC)/quadrupole time-of-flight mass spectrometry (CSH-QTOF MS) (23). Targeted bioactive oxylipin assay included thromboxanes, prostaglandins, and hydroxy-, keto- and epoxy-lipins. All LC-MS/MS data included diverse sets of internal standards. LC-MS data were processed by MS-DIAL vs. 4.0 software (24), and compounds annotated based on accurate mass, retention time and MS/MS fragment matching using LipidBlast (25) and Massbank of North America libraries (26). MS-FLO was used to remove erroneous peaks and reduce the false discovery rate in LC datasets (27). A total of 1 208 metabolites were annotated, including 146 primary metabolites (PM), 416 biogenic amines (BA), 577 complex lipids (CL), and 69 oxylipins

(OL). Data were normalized by SERRF (28). Residual technical errors were assessed by coefficients of variation for known metabolites.

Statistical analyses

[0119] For each analyte, missing values reflecting measurements below the detection limit were replaced with 50% of the smallest available value. For each of the sex- and sample type-stratified study cohorts, we identified outliers in each of the four panels using principal component analysis (PCA) (Table 4). After eliminating outliers, data were natural log-transformed and divided by the standard deviation of the analyte within the control group.

[0120] Adjusted logistic regression models were used to test separately in boys and girls for associations between metabolites and ASD. In both MMG and CB analyses, models were adjusted for maternal age, illnesses (fever, infection, inflammatory, autoimmune, allergic disorders), emotional distress scores (SCL-5), and non-NSAID antipyretic medications, as well as gestational age. Multiple comparisons were corrected using the Benjamini-Hochberg procedure, controlling the overall false discovery rate (FDR) at the level of 0.05. Additionally, chemical enrichment analyses (ChemRICH) (29) were performed to determine chemical classes that were altered between groups. ChemRICH does not rely upon background databases for statistical calculations and provides enrichment analysis based upon chemical structure, as opposed to defined pathways that can be inherently flawed (29). For each metabolite, we conducted Bayesian analysis on the logistic regression models. The Bayesian alternatives to the null hypothesis significance testing (NHST) framework has been proven to improve the biological interpretability of metabolomics data from human cohorts (30). We then calculated the BF and 95% highest density credible intervals (HDI). We considered a metabolite to be associated with ASD if BF > 10 and the 95% credible interval did not overlap with zero (31) (see below Bayesian Analyses).

[0121] To examine sex differences in metabolomic profiles, we combined data from boys and girls and repeated the adjusted logistic regression models including sex as a covariate, together with the interaction term between sex and each metabolite. ChemRICH analyses were conducted using the estimates of the sex-metabolite interaction terms.

[0122] To explore the utility of the metabolomics assay as a biomarker for ASD, we employed four machine learning algorithms: least absolute shrinkage and selection operator (Lasso) (32), adaptive Lasso (AdaLasso) (33), Random Forests (RF) (34), and XGBoost (35). Models were built,

trained, and evaluated separately for each sample type (MMG, CB), and sex. Due to the high dimensionality in the metabolomics data, we applied a novel feature selection method called model-X (MX) knockoffs (36) that distinguish important variables from noise in high-dimensional datasets while controlling the FDR. The procedure was repeated for 500 iterations controlling the FDR at the level of 0.1, and the variable importance was measured as the number of iterations in which a metabolomic analyte was selected (see below, MX Knockoffs). For each of the machine learning algorithms, we considered four sets of predictors: ***Set 1***, all metabolites; ***Set 2***, metabolites with BF > 3; ***Set 3***, metabolites that were selected by MX knockoffs in more than one iteration; ***Set 4***, metabolites that were selected by MX knockoffs in at least one iteration and had BF > 1. The predictive performance was evaluated in random sampling cross-validation with 500 iterations. In each iteration, the sample set was randomly divided into an 80% training set and a 20% test set. We generated Receiver Operating Characteristic (ROC) curves and then calculated the Area under the Receiver Operating Characteristic (AUROC) curve.

[0123] Data analyses were conducted using Matlab (R2021a, The Mathworks Inc., MA) and R (version 4.1.0). All *p*-values were 2-tailed.

Batch Effect Correction

[0124] Using scatter plots of the principal components, we detected a batch effect in the complex lipid (CL) data from maternal mid-gestation (MMG) plasma samples that was linked to equipment failure resulting in a one-month delay in completing the lipidomic analyses (Fig. 5A). After encountering batch effects in transcriptomic studies, we adopted a strategy for correction whereby analyses of specimens from cases and controls alternate. In this instance, we assembled quartets comprising (in order) one autism spectrum disorder (ASD) boy, one control boy, one ASD girl, and one control girl. The individual quartets were then randomized and submitted for metabolomic analysis. We identified 3 metabolites for which the prevalence of values below the detection limit differed dramatically between batches: PC (34:0) - ESI (-) (0% in batch 1, 48.8% in batch 2), LPE (20:4) - ESI (+) (4.3% in batch 1, 48.4% in batch 2), and PC (p-40:7)/PC (o-40:8) (49.4% in batch 1, 29.5% in batch 2). The difference in prevalence did not exceed 5% for other metabolites. Accordingly, PC (34:0) - ESI (-), LPE (20:4) - ESI (+), and PC (p-40:7)/PC (o-40:8) were excluded from MMG analyses. Comparisons of variance between batches suggested that mere shifting the means or medians was not sufficient, therefore, we implemented ComBat [109] that uses a

Bayesian framework to correct for the batch effect. Fig. 5B shows that the correction was successful.

Bayesian Analyses

[0125] For each metabolite, we conducted Bayesian analysis on the logistic regression models using R packages "rstanarm" [110] and "bayestestR" [111]. The Bayesian alternatives to the null hypothesis significance testing (NHST) framework has been proven to improve the biological interpretability of metabolomics data from human cohorts [112]. Specifically, under the frequentist NHST framework, univariate metabolomic analyses using false discovery rate (FDR) or family-wise error rate (FWER) procedure only consider a metabolite to be associated with the outcome if its *p*-value falls below an arbitrary threshold. However, NHST cannot distinguish whether there is a true null effect or the data are insensitive. Moreover, a *p*-value reports the probability of the data given the hypothesis, rather than the probability of the hypothesis given the data. Thus, analysis under the frequentist framework cannot provide support for an alternative hypothesis. On the other hand, univariate metabolomic analyses under Bayesian framework can quantify the size of an effect and the strength of evidence in favor of one hypothesis over another through the estimation of BayesFactor (BF). This allows investigators to discriminate between an inconclusive finding and evidence in favor of the null hypothesis. In our Bayesian analyses, default (weakly informative) prior distributions from rstanarm were applied adjusting the scales of the priors internally. The default priors do not strongly affect the posterior distribution but help stabilize computation, while still allowing for extreme effect sizes if warranted by the data [113, 114]. We then calculated the BFs and 95% highest density credible intervals (HDIs). BFs are ratios that quantify the probability of the alternative hypothesis ($\beta \neq 0$) over the null hypothesis ($\beta = 0$) by estimating the strength of evidence [115]. Therefore, BF=1 indicates equal likelihoods of either hypothesis, given the data. Values larger than 1 imply stronger evidence in favor of the alternative hypothesis over the null hypothesis. We used Jeffreys' [116] guidelines to categorize these values as anecdotal (BF between 1 and 3), moderate (BF between 3 and 10), strong (BF between 10 and 30), very strong (BF between 30 and 100), or extreme (BF > 100) evidence for the alternative hypothesis. We considered a metabolite to be associated with ASD if BF > 10. The 95% credible intervals are a range of values within which the true effect falls at 95% confidence, given the data [117]. It has a different interpretation from the 95% confidence interval (CI) in the frequentist

framework which instead signifies that with a large number of repeated samples, 95% of such calculated CIs would include the true value of the parameter. Metabolites were considered to be altered if the 95% credible interval did not overlap with zero.

MX Knockoffs

[0126] To explore the utility of the metabolomics assay as a biomarker for ASD, we applied a novel feature selection method called "model-X" knockoffs or MX knockoffs [118] due to the high dimensionality in the metabolomics data. MX knockoffs teases apart the important variables from noise in high-dimensional datasets while controlling the FDR. The basic idea behind the approach is to generate a set of "control" variables—the eponymous knockoffs—against which to test the importance of the original input features. The method has mostly been employed in genetic association studies [119-120], but has recently drawn attention in metabolomics [121, 122]. The MX knockoff procedure can be conducted in the parametric setting using Lasso and in the nonparametric setting using Random Forests (RF). We used MX knockoffs with RF because the covariate distribution of the metabolomics data is unknown. The procedure was repeated for 500 iterations controlling the FDR at the level of 0.1, and the variable importance was measured as the number of iterations in which a metabolomic analyte was selected.

Example 2- Study population characteristics

[0127] MMG analyses included 408 samples from 158 ASD boys, 158 control boys, 45 ASD girls, and 47 control girls. CB analyses included 418 samples from 155 ASD boys, 164 control boys, 49 ASD girls, and 50 control girls. Demographic and clinical characteristics are illustrated in Table 1. Birth year was differently proportioned between ASD and control boys in the CB analysis (*chi-squaredp*=0.010). The distribution of birth season was different between ASD and control boys in both MMG (*chi-squaredp*=0.049) and CB (*chi-squaredp*=0.020) analyses. In CB analysis, ASD boys were more likely to be born outside of the 37-41 gestational week window (*chi-squaredp*=0.015). Other parameters did not significantly differ between cases and controls.

Example 3 - Metabolomic datasets

[0128] Targeted and untargeted mass spectrometry platforms yielded data for 1,208 metabolic analytes comprising 146 PM, 416 BA, 577 CL, and 69 OL. We detected a batch effect in the CL

data from MMG samples (Fig. 5A). Through quality control, we eliminated three metabolites (PC (34:0) - ESI (-), LPE (20:4) - ESI (+), and PC (p-40:7)/PC (o-40:8)) in the MMG analysis due to dramatic difference in the missing value prevalence between batches. Using ComBat (37), we corrected for the batch effect (Fig. 5B).

Example 4 - ASD is associated with altered metabolomic profiles

[0129] Table 2 shows the estimated adjusted odds ratios (aOR), associated 95% confidence intervals (CI), *p*-values, FDR adjusted *p*-values, and BFs (31) of the metabolites associated with ASD risk in each of the sex- and sample type-stratified datasets. Because we used weakly informative priors in Bayesian analysis, the maximum-a-posteriori estimates of the aORs (aOR_{MAP}) and the 95% HDIs were similar to the aORs and 95% CIs, respectively.

[0130] In MMG, no metabolite was associated with ASD risk under the NHST framework (all FDR adjusted *p*-values were above 0.05). Using Bayesian analysis, we found ASD associations (BF > 10) in girls with decreased levels of 2,6-dihydroxybenzoic acid and increased levels of glutamic acid, pyroglutamic acid, propionic acid, N-methylalanine, and pseudouridine in the BA panel. In the CL panel, ASD girls had higher levels of PC (p-40:3)/PC (o-40:4) and HexCer-NS (d34:1) than control girls; and the levels of TAG (56:9), PE (36:5)/PE (16:0-20:5), PC (40:8) - ESI (+), and PC (37:5) were lower in ASD girls compared to control girls.

[0131] In MMG boys, Bayesian analysis identified increased levels of FA (20:3) (homo-gamma-linolenic acid) and OxPC (34:2+1O)/OxPC (16:0-18:2+1O) in the CL panel to be associated with increased ASD risk. Levels of 17-hydroxy-4,7,10,13,15,19-docosahexaenoic acid in the OL panel were lower in ASD boys.

[0132] In CB, through Bayesian analysis, we found increased levels of orotic acid in the PM panel and 2'-O-methylguanosine in the BA panel in ASD girls. In the CL panel, ASD girls had higher levels of DAG (36:1), TAG (56:3), and FA (20:3) (eicosatrienoic acid). Levels of TAG (56:9), TAG (56:8) A, and TAG (58:10) were reduced. Compared to controls, ASD boys had higher levels of FA (20:4) (arachidonic acid, AA) and erythritol under both NHST and Bayesian frameworks (FDR adjusted $p < 0.05$ and BF > 10). Bayesian analysis additionally identified ASD association with increased levels of alanine, glucose-6-phosphate, pseudouridine, methionine, and succinic acid in the PM panel in cord blood samples from the male subjects. ASD boys had higher levels

of 2'-deoxyadenosine-5'-monophosphate and glycerophosphocholine, and lower levels of 1,4-cyclohexanedione and glutamine. Levels of Leukotriene B4 in the OL panel were significantly reduced in ASD boys.

[0133] We also compared ratios of AA/Docosahexaenoic acid (DHA), AA/Eicosapentaenoic acid (EPA) (38, 39), and glutamate (Glu)/glutamine (Gln) (40, 41) between ASD and control subjects in each of the sex- and sample type-stratified datasets (Table 5). In MMG, AA/EPA ratios were elevated in ASD girls, but not in boys. In CB, both boys and girls with ASD had elevated AA/DHA ratios. Glu/Gln ratios were elevated in both ASD boys and girls in MMG. In CB, ASD boys, but not girls, had elevated Glu/Gln ratios.

Example 5 - Set enrichment analysis revealed altered chemical clusters in ASD

[0134] Chemical enrichment analyses of the results from the logistic regression models were performed using ChemRICH (29) to determine chemical clusters that were altered between ASD and control groups (Fig. 1A – 1D, Table 3).

[0135] In MMG, ASD girls had reduced levels of polyunsaturated fatty acid-containing phosphatidylcholine (PC-PUFA), hydroxybenzoates, phosphatidylcholines (PC), and phosphatidylethanolamines (PE). Increased levels of galactosylceramides, very long-chain ether-linked phosphatidylcholine (PC-ether-vlc), alanine and derivatives, polyunsaturated fatty acid-containing ether-linked phosphatidylcholine (PC-ether-PUFA), and ether-linked phosphatidylcholine (PC-ether) were associated with increased risk of ASD (Fig. 1A).

ASD boys had lower levels of hydroxy eicosapentaenoic acid (HEPE), hydroxy fatty acid_22_6_1 (OH-FA_22_6_1), and basic amino acids. There were mixed directional alterations in the levels of PCs, 1-acyl-sn-glycero-3-phosphocholines, and saturated lysophosphatidylcholines (LPC). Levels of adipates and unsaturated fatty acids were increased in ASD boys (Fig. 1B).

[0136] In CB, levels of triacylglycerols and the majority of unsaturated triglycerides were reduced in ASD girls. ASD girls had increased levels of ceramides, purine nucleosides, and sphingomyelins (SM) (Fig. 1C). ChemRICH analysis of the CB from the boys revealed the highest number of altered chemical clusters (Fig.1D). Levels of nitro compounds, dihydroxyeicosatrienoic acid (DiHETrE), epoxyeicosatrienoic acid (EpETrE), epoxyoctadecadienoic acid (EpODE), PCs, and the majority of hexoses were lower in ASD boys. The majority of hexosephosphates, unsaturated

fatty acids, and dipeptides were up-regulated. ASD boys also had higher levels of long-chain ceramides, ceramides, aspartic acids, pyrimidinones, laurates, alanine and derivatives, malates, sugar alcohols, unsaturated ceramides, and carnitines.

[0137] ASD risk in children is associated with parental education (42) and periconceptional maternal intake of folic acid (43). Accordingly, we conducted sensitivity analyses with ChemRICH additionally adjusting for parental education and maternal exposure to folic acid supplements between 0-8 weeks of gestation. No substantive changes in the dysregulated chemical clusters were observed (data not shown).

Example 6 - ASD boys and girls have different metabolomic profiles

[0138] We combined metabolomic data from boys and girls and repeated the adjusted logistic regression models including sex as a covariate, together with the interaction term between sex and each metabolite. ChemRICH analyses were conducted using the estimates of the sex-metabolite interaction terms. Univariate analyses in boys and girls were used to elucidate the directions and magnitudes of sex differences (Table 2). The full ChemRICH results for these interaction analyses are reported in Table 6.

[0139] In MMG, the chemical clusters with significant sex differences in ASD association included PC-PUFA, galactosylceramides, PC-ether-vlc, PC-ether, phosphatidylinositols (PI), PEs, and glycosphingolipids. Compound estimates in boys and girls revealed that the sex-specific differences were driven by the female cohort; there were no differences in levels of the metabolites between ASD and control boys. In contrast, girls with ASD had lower levels of PC-PUFA, PIs, and PEs, and higher levels of galactosylceramides, PC-ether-vlc, PC-ether, and glycosphingolipids.

[0140] In CB, we found two chemical clusters with sex differences in ASD association: hexosephosphates and cyclohexanes. The ASD associations with these metabolites were in opposite directions between boys and girls (increased in ASD boys but decreased in ASD girls; or decreased in ASD boys but increased in ASD girls); estimates in the male cohort were more significant.

Example 7 - Assessment of the metabolomic assays as a biomarker for ASD

[0141] We considered four sets of predictors for each of the machine learning algorithms: ***Set 1***, all metabolites; ***Set 2***, metabolites with BF > 3; ***Set 3***, metabolites selected by MX knockoffs in more than one iteration (Table 7); and ***Set 4***, metabolites selected by MX knockoffs in at least one iteration with BF > 1. Fig. 2A–D show the ROC curves and the AUROC values, differentiating ASD cases from controls. The AUROC values and 95% CIs are reported in Table 8. The *p*-values comparing the predictive performances between machine learning algorithms within each predictor set and between predictor sets using each algorithm was also performed (results not shown).

[0142] In the MMG girls (Fig. 2A), machine learning models with ***Set 4*** predictors yielded highest average AUROC values (***Set 1***: 0.426; ***Set 2***: 0.697; ***Set 3***: 0.724; ***Set 4***: 0.742); the best performing classifier was RF with ***Set 3*** predictors that differentiated ASD girls from control girls with an AUROC value of 0.817 (95% CI: 0.711 ~ 0.890). In MMG boys (Fig. 2B), machine learning models with ***Set 2***, ***Set 3***, and ***Set 4*** predictors produced similar average AUROC values (***Set 1***: 0.531; ***Set 2***: 0.657; ***Set 3***: 0.663; ***Set 4***: 0.663). RF with ***Set 3*** predictors outperformed the other classifiers and yielded an AUROC value of 0.710 (95% CI: 0.647 ~ 0.766). In CB girls (Fig. 2C), machine learning models with ***Set 3*** predictors were more accurate than models based on the other predictors (***Set 1***: 0.492; ***Set 2***: 0.724; ***Set 3***: 0.766; ***Set 4***: 0.731). RF with ***Set 3*** predictors separated ASD girls from control girls with an AUROC value of 0.853 (95% CI: 0.754 ~ 0.917). In CB boys (Fig. 2D), ***Set 3*** was the most accurate predictors (***Set 1***: 0.617; ***Set 2***: 0.690; ***Set 3***: 0.720; ***Set 4***: 0.698); XGBoost with ***Set 3*** predictors produced an AUROC value of 0.766 (95% CI: 0.707 ~ 0.816).

Example 8 - Post Hoc Power Analysis

[0143] Autism is approximately 4-fold less common in girls. Accordingly, our sample size was smaller for girls than boys. We conducted *post hoc* power analysis for every logistic regression model and compared power between boys and girls in the MMG and cord blood (CB) datasets, respectively. In the MMG analysis, the female cohort, despite smaller sample size, exhibited higher power than the male cohort (power in girls had a mean of 0.188 and a standard deviation of 0.176 ***vs.*** power in boys had a mean of 0.168 and a standard deviation of 0.166, *paired t-test* p=0.003), suggesting that the effect sizes in girls were larger than those in boys. In the CB analysis, the male

cohort had higher power than the female cohort (power in girls had a mean of 0.173 and a standard deviation of 0.162 ***vs.*** power in boys had a mean of 0.219 and a standard deviation of 0.219, *paired t-test* $p=5.11 \times 10^{-9}$) (results of post hoc powers for individual metabolites not shown).

Example 9 - Subsampling Analysis

[0144] To examine whether the difference in the altered chemical clusters between boys and girls resulted from the difference in the sample sizes, we conducted 1 000 iterations of random subsampling (without replacement) of 50 ASD boys and 50 control boys in the MMG and CB datasets, respectively. The logistic regression estimates from 1,000 subsamples were pooled using Rubin's rules [123] and used as inputs for ChemRICH. No chemical clusters, in either MMG or CB analyses, were significantly altered (data not shown).

[0145] REFERENCES

1. Maenner MJ, Shaw KA, Bakian AV, Bilder DA, Durkin MS, Esler A, et al. Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. MMWR Surveill Summ. 2021;70(11):1-16.

2. Towle PO, Patrick PA, Ridgard T, Pham S, Marrus J. Is Earlier Better? The Relationship between Age When Starting Early Intervention and Outcomes for Children with Autism Spectrum Disorder: A Selective Review. Autism Res Treat. 2020;2020:7605876.

3. Wallace KS, Rogers SJ. Intervening in infancy: implications for autism spectrum disorders. J Child Psychol Psychiatry. 2010;51(12):1300-20.

4. Zwaigenbaum L, Penner M. Autism spectrum disorder: advances in diagnosis and evaluation. BMJ. 2018;361:k1674.

5. Siu MT, Weksberg R. Epigenetics of Autism Spectrum Disorder. Adv Exp Med Biol. 2017;978:63-90.

6. Ritz B, Yan Q, Uppal K, Liew Z, Cui X, Ling C, et al. Untargeted Metabolomics Screen of Mid-pregnancy Maternal Serum and Autism in Offspring. Autism Res. 2020;13(8):1258-69.

7. Schmidt RJ, Liang D, Busgang SA, Curtin P, Giulivi C. Maternal Plasma Metabolic Profile Demarcates a Role for Neuroinflammation in Non-Typical Development of Children. Metabolites. 2021;11(8).

8. Panjwani AA, Ji Y, Fahey JW, Palmer A, Wang G, Hong X, et al. Maternal Dyslipidemia, Plasma Branched-Chain Amino Acids, and the Risk of Child Autism Spectrum Disorder: Evidence of Sex Difference. J Autism Dev Disord. 2020;50(2):540-50.

9. Courraud J, Ernst M, Svane Laursen S, Hougaard DM, Cohen AS. Studying Autism Using Untargeted Metabolomics in Newborn Screening Samples. J Mol Neurosci. 2021;71(7):1378-93.

10. Barone R, Alaimo S, Messina M, Pulvirenti A, Bastin J, Group MI-A, et al. A Subset of Patients With Autism Spectrum Disorders Show a Distinctive Metabolic Profile by Dried Blood Spot Analyses. Front Psychiatry. 2018;9:636.

11. Stoltenberg C, Schjolberg S, Bresnahan M, Hornig M, Hirtz D, Dahl C, et al. The Autism Birth Cohort: a paradigm for gene-environment-timing research. Mol Psychiatry. 2010;15(7):676-80.

12. Magnus P, Irgens LM, Haug K, Nystad W, Skjaerven R, Stoltenberg C, et al. Cohort profile: the Norwegian Mother and Child Cohort Study (MoBa). Int J Epidemiol. 2006;35(5):1146-50.

13. Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, et al. Cohort Profile Update: The Norwegian Mother and Child Cohort Study (MoBa). Int J Epidemiol. 2016;45(2):382-8.

14. Bresnahan M, Hornig M, Schultz AF, Gunnes N, Hirtz D, Lie KK, et al. Association of maternal report of infant and toddler gastrointestinal symptoms with autism: evidence from a prospective birth cohort. JAMA Psychiatry. 2015;72(5):466-74.

15. Association AP. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR). . American Psychiatric Association: Arlington. 2000.

16. Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord. 1994;24(5):659-85.

17. Lord C, Risi S, Lambrecht L, Cook EH, Jr., Leventhal BL, DiLavore PC, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord. 2000;30(3):205-23.

18. Organization WH. International classification of diseases and related health problems, 10th revision. Geneva: World Health Organization. 1992.

19. Strand BH, Dalgard OS, Tambs K, Rognerud M. Measuring the mental health status of the Norwegian population: a comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36). Nord J Psychiatry. 2003;57(2):113-8.

20. Ronningen KS, Paltiel L, Meltzer HM, Nordhagen R, Lie KK, Hovengen R, et al. The biobank of the Norwegian Mother and Child Cohort Study: a resource for the next 100 years. Eur J Epidemiol. 2006;21(8):619-25.

21. Fiehn O. Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling. Curr Protoc Mol Biol. 2016;114:30 4 1- 4 2.

22. Kind T, Wohlgemuth G, Lee DY, Lu Y, Palazoglu M, Shahbaz S, et al. FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. Anal Chem. 2009;81(24):10038-48.

23. Cajka T, Smilowitz JT, Fiehn O. Validating Quantitative Untargeted Lipidomics Across Nine Liquid Chromatography-High-Resolution Mass Spectrometry Platforms. Anal Chem. 2017;89(22):12360-8.

24. Tsugawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K, et al. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat Methods. 2015;12(6):523-6.

25. Tsugawa H, Ikeda K, Takahashi M, Satoh A, Mori Y, Uchino H, et al. A lipidome atlas in MS-DIAL 4. Nat Biotechnol. 2020;38(10):1159-63.

26. Wohlgemuth G, Mehta SS, Mejia RF, Neumann S, Pedrosa D, Pluskal T, et al. SPLASH, a hashed identifier for mass spectra. Nat Biotechnol. 2016;34(11):1099-101.

27. DeFelice BC, Mehta SS, Samra S, Cajka T, Wancewicz B, Fahrmann JF, et al. Mass Spectral Feature List Optimizer (MS-FLO): A Tool To Minimize False Positive Peak Reports in Untargeted Liquid Chromatography-Mass Spectroscopy (LC-MS) Data Processing. Anal Chem. 2017;89(6):3250-5.

28. Fan S, Kind T, Cajka T, Hazen SL, Tang WHW, Kaddurah-Daouk R, et al. Systematic Error Removal Using Random Forest for Normalizing Large-Scale Untargeted Lipidomics Data. Anal Chem. 2019;91(5):3590-6.

29. Barupal DK, Fiehn O. Chemical Similarity Enrichment Analysis (ChemRICH) as alternative to biochemical pathway mapping for metabolomic datasets. Sci Rep. 2017;7(1):14567.

30. Brydges C, Che X, Lipkin WI, Fiehn O. Bayesian statistics improves biological interpretability of metabolomics data from human cohorts. bioRxiv. 2022.

31. Jeffreys H. Theory of Probability. 3rd ed. Oxford: Clarendon Press; 1961.

32. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological). 1996;58(1):267-88.

33. Zou H. The Adaptive Lasso and Its Oracle Properties. Journal of the American Statistical Association. 2012;101(476):1418-29.

34. Breiman L. Random Forests. Machine Learning. 2001;45:5-32.

35. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016; San Francisco, CA2016. p. 785-94.

36. Candes E, Fan Y, Janson L, Lv J. Panning for gold:'model-X'knockoffs for high dimensional controlled variable selection. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2018;80(3):551-77.

37. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118-27.

38. Cappelletti M, Della Bella S, Ferrazzi E, Mavilio D, Divanovic S. Inflammation and preterm birth. J Leukoc Biol. 2016;99(1):67-78.

39. Parletta N, Niyonsenga T, Duff J. Omega-3 and Omega-6 Polyunsaturated Fatty Acid Levels and Correlations with Symptoms in Children with Attention Deficit Hyperactivity Disorder, Autistic Spectrum Disorder and Typically Developing Controls. PLoS One. 2016;11(5):e0156432.

40. Al-Otaish H, Al-Ayadhi L, Bjorklund G, Chirumbolo S, Urbina MA, El-Ansary A. Relationship between absolute and relative ratios of glutamate, glutamine and GABA and severity of autism spectrum disorder. Metab Brain Dis. 2018;33(3):843-54.

41. El-Ansary A. Data of multiple regressions analysis between selected biomarkers related to glutamate excitotoxicity and oxidative stress in Saudi autistic patients. Data Brief. 2016;7:111-6.

42. Van Meter KC, Christiansen LE, Delwiche LD, Azari R, Carpenter TE, Hertz-Picciotto I. Geographic distribution of autism in California: a retrospective birth cohort analysis. Autism Res. 2010;3(1):19-29.

43. Suren P, Roth C, Bresnahan M, Haugen M, Hornig M, Hirtz D, et al. Association between maternal use of folic acid supplements and risk of autism spectrum disorders in children. JAMA. 2013;309(6):570-7.

44. Wiggs KK, Rickert ME, Sujan AC, Quinn PD, Larsson H, Lichtenstein P, et al. Antiseizure medication use during pregnancy and risk of ASD and ADHD in children. Neurology. 2020;95(24):e3232-e40.

45. Stromland K, Nordin V, Miller M, Akerstrom B, Gillberg C. Autism in thalidomide embryopathy: a population study. Dev Med Child Neurol. 1994;36(4):351-6.

46. Smith SE, Li J, Garbett K, Mirnics K, Patterson PH. Maternal immune activation alters fetal brain development through interleukin-6. J Neurosci. 2007;27(40):10695-702.

47. Nicolini C, Fahnestock M. The valproic acid-induced rodent model of autism. Exp Neurol. 2018;299(Pt A):217-27.

48. Patterson PH. Maternal infection and immune involvement in autism. Trends Mol Med. 2011;17(7):389-94.

49. Parenti M, Schmidt RJ, Ozonoff S, Shin HM, Tancredi DJ, Krakowiak P, et al. Maternal Serum and Placental Metabolomes in Association with Prenatal Phthalate Exposure and Neurodevelopmental Outcomes in the MARBLES Cohort. Metabolites. 2022;12(9).

50. Hollowood K, Melnyk S, Pavliv O, Evans T, Sides A, Schmidt RJ, et al. Maternal metabolic profile predicts high or low risk of an autism pregnancy outcome. Res Autism Spectr Disord. 2018;56:72-82.

51. Girchenko P, Lahti-Pulkkinen M, Lipsanen J, Heinonen K, Lahti J, Rantalainen V, et al. Maternal early-pregnancy body mass index-associated metabolomic component and mental and behavioral disorders in children. Mol Psychiatry. 2022;27(11):4653-61.

52. Lyall K, Windham GC, Snyder NW, Kuskovsky R, Xu P, Bostwick A, et al. Association Between Midpregnancy Polyunsaturated Fatty Acid Levels and Offspring Autism Spectrum

Disorder in a California Population-Based Case-Control Study. Am J Epidemiol. 2021;190(2):265-76.

53. Kim JH, Yan Q, Uppal K, Cui X, Ling C, Walker DI, et al. Metabolomics analysis of maternal serum exposed to high air pollution during pregnancy and risk of autism spectrum disorder in offspring. Environ Res. 2021;196:110823.

54. Ferdinandusse S, McWalter K, Te Brinke H, L IJ, Mooijer PM, Ruiter JPN, et al. An autosomal dominant neurological disorder caused by de novo variants in FAR1 resulting in uncontrolled synthesis of ether lipids. Genet Med. 2021;23(4):740-50.

55. Staps P, Rizzo WB, Vaz FM, Bugiani M, Giera M, Heijs B, et al. Disturbed brain ether lipid metabolism and histology in Sjogren-Larsson syndrome. J Inherit Metab Dis. 2020;43(6):1265-78.

56. Vaz FM, McDermott JH, Alders M, Wortmann SB, Kolker S, Pras-Raves ML, et al. Mutations in PCYT2 disrupt etherlipid biosynthesis and cause a complex hereditary spastic paraplegia. Brain. 2019;142(11):3382-97.

57. Teruya T, Chen YJ, Kondoh H, Fukuji Y, Yanagida M. Whole-blood metabolomics of dementia patients reveal classes of disease-linked metabolites. Proc Natl Acad Sci U S A. 2021;118(37).

58. Bjornevik K, Zhang Z, O'Reilly EJ, Berry JD, Clish CB, Deik A, et al. Prediagnostic plasma metabolomics and the risk of amyotrophic lateral sclerosis. Neurology. 2019;92(18):e2089-e100.

59. Yang F, Wu SC, Ling ZX, Chao S, Zhang LJ, Yan XM, et al. Altered Plasma Metabolic Profiles in Chinese Patients With Multiple Sclerosis. Front Immunol. 2021;12:792711.

60. Liao X, Yang J, Wang H, Li Y. Microglia mediated neuroinflammation in autism spectrum disorder. Journal of psychiatric research. 2020;130:167-76.

61. Gupta S, Ellis SE, Ashar FN, Moes A, Bader JS, Zhan J, et al. Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. Nat Commun. 2014;5:5748.

62. Balic A, Vlasic D, Zuzul K, Marinovic B, Bukvic Mokos Z. Omega-3 Versus Omega-6 Polyunsaturated Fatty Acids in the Prevention and Treatment of Inflammatory Skin Diseases. Int J Mol Sci. 2020;21(3).

63. So J, Wu D, Lichtenstein AH, Tai AK, Matthan NR, Maddipati KR, et al. EPA and DHA differentially modulate monocyte inflammatory response in subjects with chronic inflammation in part via plasma specialized pro-resolving lipid mediators: A randomized, double-blind, crossover study. Atherosclerosis. 2021;316:90-8.

64. Onodera T, Fukuhara A, Shin J, Hayakawa T, Otsuki M, Shimomura I. Eicosapentaenoic acid and 5-HEPE enhance macrophage-mediated Treg induction in mice. Sci Rep. 2017;7(1):4560.

65. Kishikawa A, Kitaura H, Kimura K, Ogawa S, Qi J, Shen WR, et al. Docosahexaenoic Acid Inhibits Inflammation-Induced Osteoclast Formation and Bone Resorption in vivo Through GPR120 by Inhibiting TNF-alpha Production in Macrophages and Directly Inhibiting Osteoclast Formation. Front Endocrinol (Lausanne). 2019;10:157.

66. Kikut J, Komorniak N, Zietek M, Palma J, Szczuko M. Inflammation with the participation of arachidonic (AA) and linoleic acid (LA) derivatives (HETEs and HODEs) is necessary in the course of a normal reproductive cycle and pregnancy. J Reprod Immunol. 2020;141:103177.

67. Rizzo MT, Carlo-Stella C. Arachidonic acid mediates interleukin-1 and tumor necrosis factor-alpha-induced activation of the c-jun amino-terminal kinases in stromal cells. Blood. 1996;88(10):3792-800.

68. Hughes-Fulford M, Li CF, Boonyaratanakornkit J, Sayyah S. Arachidonic acid activates phosphatidylinositol 3-kinase signaling and induces gene expression in prostate cancer. Cancer Res. 2006;66(3):1427-33.

69. Node K, Huo Y, Ruan X, Yang B, Spiecker M, Ley K, et al. Anti-inflammatory properties of cytochrome P450 epoxygenase-derived eicosanoids. Science. 1999;285(5431):1276-9.

70. Ghazali R, Mehta KJ, Bligh SA, Tewfik I, Clemens D, Patel VB. High omega arachidonic acid/docosahexaenoic acid ratio induces mitochondrial dysfunction and altered lipid metabolism in human hepatoma cells. World J Hepatol. 2020;12(3):84-98.

71. Cotogni P, Muzio G, Trombetta A, Ranieri VM, Canuto RA. Impact of the omega-3 to omega-6 polyunsaturated fatty acid ratio on cytokine release in human alveolar cells. JPEN J Parenter Enteral Nutr. 2011;35(1):114-21.

72. Che X, Hornig M, Bresnahan M, Stoltenberg C, Magnus P, Suren P, et al. Maternal mid-gestational and child cord blood immune signatures are strongly associated with offspring risk of ASD. Mol Psychiatry. 2022;27(3):1527-41.

73. Dawaliby R, Trubbia C, Delporte C, Masureel M, Van Antwerpen P, Kobilka BK, et al. Allosteric regulation of G protein-coupled receptor activity by phospholipids. Nat Chem Biol. 2016;12(1):35-9.

74. van der Veen JN, Kennelly JP, Wan S, Vance JE, Vance DE, Jacobs RL. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. Biochim Biophys Acta Biomembr. 2017;1859(9 Pt B):1558-72.

75. Schuler MH, Di Bartolomeo F, Bottinger L, Horvath SE, Wenz LS, Daum G, et al. Phosphatidylcholine affects the role of the sorting and assembly machinery in the biogenesis of mitochondrial beta-barrel proteins. J Biol Chem. 2015;290(44):26523-32.

76. Niebergall LJ, Vance DE. The ratio of phosphatidylcholine to phosphatidylethanolamine does not predict integrity of growing MT58 Chinese hamster ovary cells. Biochim Biophys Acta. 2012;1821(2):324-34.

77. Kano-Sueoka T, Nicks ME. Abnormal function of protein kinase C in cells having phosphatidylethanolamine-deficient and phosphatidylcholine-excess membranes. Cell Growth Differ. 1993;4(7):533-7.

78. Rockenfeller P, Koska M, Pietrocola F, Minois N, Knittelfelder O, Sica V, et al. Phosphatidylethanolamine positively regulates autophagy and longevity. Cell Death Differ. 2015;22(3):499-508.

79. Longo N, Frigeni M, Pasquali M. Carnitine transport and fatty acid oxidation. Biochim Biophys Acta. 2016;1863(10):2422-35.

80. Needham BD, Adame MD, Serena G, Rose DR, Preston GM, Conrad MC, et al. Plasma and Fecal Metabolite Profiles in Autism Spectrum Disorder. Biol Psychiatry. 2021;89(5):451-62.

81. Thomas RH, Foley KA, Mepham JR, Tichenoff LJ, Possmayer F, MacFabe DF. Altered brain phospholipid and acylcarnitine profiles in propionic acid infused rodents: further development of a potential model of autism spectrum disorders. J Neurochem. 2010;113(2):515-29.

82. Uche LE, Gooris GS, Bouwstra JA, Beddoes CM. Increased Levels of Short-Chain Ceramides Modify the Lipid Organization and Reduce the Lipid Barrier of Skin Model Membranes. Langmuir. 2021;37(31):9478-89.

83. Hussain MM, Jin W, Jiang XC. Mechanisms involved in cellular ceramide homeostasis. Nutr Metab (Lond). 2012;9(1):71.

84. Melland-Smith M, Ermini L, Chauvin S, Craig-Barnes H, Tagliaferro A, Todros T, et al. Disruption of sphingolipid metabolism augments ceramide-induced autophagy in preeclampsia. Autophagy. 2015;11(4):653-69.

85. Ross MM, Piorczynski TB, Harvey J, Burnham TS, Francis M, Larsen MW, et al. Ceramide: a novel inducer for neural tube defects. Dev Dyn. 2019;248(10):979-96.

86. Ariga T, Jarvis WD, Yu RK. Role of sphingolipid-mediated cell death in neurodegenerative diseases. J Lipid Res. 1998;39(1):1-16.

87. Stoica BA, Movsesyan VA, Lea PMt, Faden AI. Ceramide-induced neuronal apoptosis is associated with dephosphorylation of Akt, BAD, FKHR, GSK-3beta, and induction of the mitochondrial-dependent intrinsic caspase pathway. Mol Cell Neurosci. 2003;22(3):365-82.

88. France-Lanord V, Brugg B, Michel PP, Agid Y, Ruberg M. Mitochondrial free radical signal in ceramide-dependent apoptosis: a putative mechanism for neuronal death in Parkinson's disease. J Neurochem. 1997;69(4):1612-21.

89. Fanani ML, Maggio B. The many faces (and phases) of ceramide and sphingomyelin I - single lipids. Biophys Rev. 2017;9(5):589-600.

90. Rogasevskaia T, Coorssen JR. Sphingomyelin-enriched microdomains define the efficiency of native Ca(2+)-triggered membrane fusion. J Cell Sci. 2006;119(Pt 13):2688-94.

91. Corcelle-Termeau E, Vindelov SD, Hamalisto S, Mograbi B, Keldsbo A, Brasen JH, et al. Excess sphingomyelin disturbs ATG9A trafficking and autophagosome closure. Autophagy. 2016;12(5):833-49.

92. MahmoudianDehkordi S, Ahmed AT, Bhattacharyya S, Han X, Baillie RA, Arnold M, et al. Alterations in acylcarnitines, amines, and lipids inform about the mechanism of action of citalopram/escitalopram in major depression. Transl Psychiatry. 2021;11(1):153.

93. Bhattacharyya S, Ahmed AT, Arnold M, Liu D, Luo C, Zhu H, et al. Metabolomic signature of exposure and response to citalopram/escitalopram in depressed outpatients. Transl Psychiatry. 2019;9(1):173.

94. Dorninger F, Forss-Petter S, Berger J. From peroxisomal disorders to common neurodegenerative diseases - the role of ether phospholipids in the nervous system. FEBS Lett. 2017;591(18):2761-88.

95. Frazier TW, Georgiades S, Bishop SL, Hardan AY. Behavioral and cognitive characteristics of females and males with autism in the Simons Simplex Collection. J Am Acad Child Adolesc Psychiatry. 2014;53(3):329-40 e1-3.

96. Aldred S, Moore KM, Fitzgerald M, Waring RH. Plasma amino acid levels in children with autism and their families. J Autism Dev Disord. 2003;33(1):93-7.

97. Shimmura C, Suda S, Tsuchiya KJ, Hashimoto K, Ohno K, Matsuzaki H, et al. Alteration of plasma glutamate and glutamine levels in children with high-functioning autism. PLoS One. 2011;6(10):e25340.

98. Belov Kirdajova D, Kriska J, Tureckova J, Anderova M. Ischemia-Triggered Glutamate Excitotoxicity From the Perspective of Glial Cells. Front Cell Neurosci. 2020;14:51.

99. Tseng EE, Brock MV, Lange MS, Troncoso JC, Blue ME, Lowenstein CJ, et al. Glutamate excitotoxicity mediates neuronal apoptosis after hypothermic circulatory arrest. Ann Thorac Surg. 2010;89(2):440-5.

100. Ankarcrona M, Dypbukt JM, Bonfoco E, Zhivotovsky B, Orrenius S, Lipton SA, et al. Glutamate-induced neuronal death: a succession of necrosis or apoptosis depending on mitochondrial function. Neuron. 1995;15(4):961-73.

101. Takeuchi H, Jin S, Wang J, Zhang G, Kawanokuchi J, Kuno R, et al. Tumor necrosis factor-alpha induces neurotoxicity via glutamate release from hemichannels of activated microglia in an autocrine manner. J Biol Chem. 2006;281(30):21362-8.

102. Yang GY, Gong C, Qin Z, Liu XH, Lorris Betz A. Tumor necrosis factor alpha expression produces increased blood-brain barrier permeability following temporary focal cerebral ischemia in mice. Brain Res Mol Brain Res. 1999;69(1):135-43.

103. Zhou Y, Danbolt NC. Glutamate as a neurotransmitter in the healthy brain. J Neural Transm (Vienna). 2014;121(8):799-817.

104. Hassan TH, Abdelrahman HM, Fattah NRA, El-Masry NM, Hashim HM, El-Gerby KM, et al. Blood and brain glutamate levels in children with autistic disorder. Research in Autism Spectrum Disorders. 2013;7(4):541-8.

105. Albrecht J, Sidoryk-Wegrzynowicz M, Zielinska M, Aschner M. Roles of glutamine in neurotransmission. Neuron Glia Biol. 2010;6(4):263-76.

106. Chen J, Herrup K. Glutamine acts as a neuroprotectant against DNA damage, beta-amyloid and H2O2-induced stress. PLoS One. 2012;7(3):e33177.

107. Stelmashook EV, Lozier ER, Goryacheva ES, Mergenthaler P, Novikova SV, Zorov DB, et al. Glutamine-mediated protection from neuronal cell death depends on mitochondrial activity. Neurosci Lett. 2010;482(2):151-5.

108. Budni J, Molz S, Dal-Cim T, Martin-de-Saavedra MD, Egea J, Lopez MG, et al. Folic Acid Protects Against Glutamate-Induced Excitotoxicity in Hippocampal Slices Through a Mechanism that Implicates Inhibition of GSK-3beta and iNOS. Mol Neurobiol. 2018;55(2):1580-9.

109. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8(1):118-127.

110. Goodrich B, Gabry J, Ali I, Brilleman S. rstanarm: Bayesian applied regression modeling via Stan. R package version. 2020;2(1).

111. Makowski D, Ben-Shachar MS, Lüdecke D. bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. J Open Source Softw. 2019;4(40):1541.

112. Brydges C, Che X, Lipkin WI, Fiehn O. Bayesian statistics improvesbiological interpretability of metabolomics data from human cohorts. bioRxiv. 2022.

113. Gelman A, Jakulin A, Pittau MG, Su Y-S. A weakly informative default prior distribution for logistic and other regression models. Ann Appl Stat. 2008;2(4):1360-1383.

114. Muth C, Oravecz Z, Gabry J. User-friendly Bayesian regression modeling: A tutorial with rstanarm and shinystan. Quant Meth Psych. 2018;14(2):99-119.

115. Lee MD, Wagenmakers E-J. Bayesian cognitive modeling: A practical course. Cambridge university Press; 2014.

116. Jeffreys H. Theory of Probability. 3rd ed. Clarendon Press: Oxford; 1961.

117. Candes E, Fan Y, Janson L, Lv J. Panning for gold:'model-X'knockoffs for high dimensional controlled variable selection. J R Stat Soc Series B Stat Methodol. 2018;80(3):551-577.

118. He T, Baik JM, Kato C, Yang H, Fan Z, Cham J, et al. Novel Ensemble Feature Selection Approach and Application in Repertoire Sequencing Data. Front Genet. 2022;13:821832.

119. Chen D, Tashman K, Palmer DS, Neale B, Roeder K, Bloemendal A, et al. A data harmonization pipeline to leverage external controls and boost power in GWAS. Hum Mol Genet. 2022;31(3):481-489.

120. Watson DS. Interpretable machine learning for genomics. Hum Genet. 2022;141(9):1499-1513.

121. Ma R, Cai TT, Li H. Global and Simultaneous Hypothesis Testing for High-Dimensional Logistic Regression Models. J Am Stat Assoc. 2021;116(534):984-998.

122. Fu G-H, Wu Y-J, Zong M-J, Yi L-Z. Feature selection and classification by minimizing overlap degree for class-imbalanced data in metabolomics. Chemom Intell Lab Syst. 2020;196:103906.

123. Rubin DB. Multiple imputation for nonresponse in surveys, vol. 81. John Wiley & Sons; 2004.

124. Guthrie W. et al. The earlier the better: an RCT of treatment timing effects for toddlers on the autism spectrum, Autism, 2023; https://doi.org/10.1177/13623613231159153

CLAIMS

1. A method for providing early autism spectrum disorder (ASD) intervention to a newborn comprising the steps of:
   a. identifying the newborn as having ASD or as being at risk of developing ASD by:
      i. performing or having performed an ASD association assay on a sample from the newborn or a mother of the newborn to identify if the newborn has ASD or is at risk of developing ASD;
      ii. identifying the newborn as having ASD or being at risk for developing ASD when the ASD association assay indicates that there is a dysregulation of the newborn's metabolomic profile; and
   b. providing an early ASD intervention to the newborn so-identified.

2. The method of claim 1, wherein the newborn's metabolomic profile is compared to a control metabolomic profile.

3. The method of claim 1, wherein the newborn's dysregulated metabolomic profile is consistent with inflammation.

4. The method of claim 3, wherein the newborn is a male and the sample is blood or blood product from the mother at mid-gestation and wherein a decreased or lower level of the compound 17-hydroxy-4,7,10,13,15,19-docosahexaenoic acid or chemical cluster hydroxy eicosapentaenoic acid (HEPE) as compared to a control level identifies the newborn as having a dysregulated metabolomic profile consistent with inflammation and as having ASD or at risk for developing ASD.

5. The method of claim 3, wherein the newborn is a female and the sample is blood from the mother at mid-gestation and wherein an increased or higher ratio of arachidonic acid (AA)/eicosapentaenoic acid (EPA) as compared to a control level identifies the newborn

as having a dysregulated metabolomic profile consistent with inflammation and as having ASD or at risk for developing ASD.

6. The method of claim 3, wherein the sample is cord blood and wherein an increased or higher level of the compound arachidonic acid (AA) or ratio of arachidonic acid (AA)/ docosahexaenoic acid (DHA) as compared to a control level identifies the newborn as having a dysregulated metabolomic profile consistent with inflammation and as having ASD or at risk for developing ASD.

7. The method of claim 3, wherein the newborn is a male and the sample is cord blood and wherein a decreased or lower level of chemical cluster epoxyeicosatrienoic acid (EpETrE) as compared to a control level identifies the newborn as having a dysregulated metabolomic profile consistent with inflammation and as having ASD or at risk for developing ASD.

8. The method of claim 1, wherein the newborn's dysregulated metabolomic profile is consistent with disrupted membrane integrity.

9. The method of claim 8, wherein the sample is blood or a blood product from the mother at mid-gestation and wherein a decreased or lower level of chemical cluster phosphatidylcholines (PC) as compared to a control level, identifies the newborn as having a dysregulated metabolomic profile consistent with disrupted membrane integrity and as having ASD or at risk for developing ASD.

10. The method of claim 8, wherein the newborn is a female and the sample is blood or blood product from the mother at mid-gestation and wherein a decreased or lower level of compounds phosphatidylcholines (PC) or phosphatidylethanolamines (PE) or chemical cluster PE as compared to a control level, or an increased or higher level of chemical clusters PC-ether or PC ether-vlc as compared to a control level, identifies the newborn as having a dysregulated metabolomic profile consistent with disrupted membrane integrity and as having ASD or at risk for developing ASD.

11. The method of claim 8, wherein the sample is cord blood and wherein an increased or higher level of chemical cluster ceramide as compared to a control level identifies the newborn as having a dysregulated metabolomic profile consistent with disrupted membrane integrity and as having ASD or at risk for developing ASD.

12. The method of claim 8, wherein the newborn is a male and the sample is cord blood and wherein a decreased or lower level of chemical cluster phosphatidylcholines (PC) as compared to a control level, or an increased or higher level of chemical clusters carnitine, long chain ceramides or unsaturated ceramides as compared to a control level identifies the newborn as having a dysregulated metabolomic profile consistent with disrupted membrane integrity and as having ASD or at risk for developing ASD.

13. The method of claim 8, wherein the newborn is a female and the sample is cord blood and wherein an increased or higher level of chemical cluster sphingomyelin as compared to a control level, identifies the newborn as having a dysregulated metabolomic profile consistent with disrupted membrane integrity and as having ASD or at risk for developing ASD.

14. The method of claim 1, wherein the newborn's dysregulated metabolomic profile is consistent with impaired neurotransmission and neurotoxicity.

15. The method of claim 14, wherein the sample is blood from the mother at mid-gestation and wherein an increased or higher ratio of Glu/Gln as compared to a control level, identifies the newborn as having a dysregulated metabolomic profile consistent with impaired neurotransmission and neurotoxicity and as having ASD or at risk for developing ASD.

16. The method of claim 14, wherein the newborn is a female and the sample is blood from the mother at mid-gestation and wherein an increased or higher level of compound Glu as compared to a control level, identifies the newborn as having a dysregulated metabolomic

profile consistent with impaired neurotransmission and neurotoxicity and as having ASD or at risk for developing ASD.

17. The method of claim 14, wherein the newborn is a male and the sample is blood from the mother at mid-gestation and wherein a decreased or lower level of compound Gln as compared to a control level, identifies the newborn as having a dysregulated metabolomic profile consistent with impaired neurotransmission and neurotoxicity and as having ASD or at risk for developing ASD.

18. The method of claim 14, wherein the newborn is a male and the sample is cord blood and wherein a decreased or lower level of Gln as compared to a control level, or an increased or higher ratio of Glu/Gln as compared to a control level, or chemical clusters long chain ceramides or unsaturated ceramides, identifies the newborn as having a dysregulated metabolomic profile consistent with impaired neurotransmission and neurotoxicity and as having ASD or at risk for developing ASD.

19. A method for providing early autism spectrum disorder (ASD) intervention to a newborn comprising the steps of:
    a. identifying that the newborn has ASD or is at risk of developing ASD by:
        i. performing or having performed an ASD association assay on a sample from a mother of the newborn to identify if the newborn has ASD or is at risk of developing ASD, wherein the sample is cord blood;
        ii. identifying the newborn as having ASD or being at risk for developing ASD when there is high number of dysregulated chemical clusters in said sample as compared to a control number;
    b. providing early ASD intervention to the newborn to the newborn so-identified.

20. A method for providing early autism spectrum disorder (ASD) intervention to a female newborn comprising the steps of:

a. identifying that the newborn has ASD or is at risk of developing ASD with increased impaired cognitive development by:
    i. performing or having performed an ASD association assay on a sample to identify if the newborn has ASD or is at risk of developing ASD with increased impaired cognitive development, wherein the sample is blood from the mother at mid-gestation;
    ii. identifying the newborn as having ASD or being at risk for developing ASD with increased impaired cognitive development when there is an imbalance of ether/non-ether phospholipids in said sample; and
b. providing early ASD intervention to the newborn identified with autism spectrum disorder (ASD) with increased impaired cognitive development.

21. A method for providing early intervention to newborn with autism spectrum disorder (ASD) comprising the steps of:
    a. identifying that the newborn has ASD or is at risk of developing ASD by:
        i. performing or having performed an ASD association assay on a sample to identify if the newborn has ASD or is at risk of developing ASD;
        ii. identifying the newborn as having ASD or being at risk for developing ASD when there is a dysregulation of the newborn's metabolomic profile, wherein the metabolomic profile is determined using one or more of the metabolites listed in Table 7; and
    b. providing early ASD intervention to the newborn.

22. The method of any of claims 1-21, wherein the early ASD intervention comprises Early Social Interaction (ESI) model intervention.

23. The method of any of claims 1-21, wherein the early ASD intervention comprises speech therapy, and/or hearing impairment therapy/treatment, and/or physical therapy.

24. The method of any of claims 1-21, wherein the early ASD intervention is administered prior to 20 months of the age of the newborn or prior to 10 months of the age of the newborn.

Fig. 1A

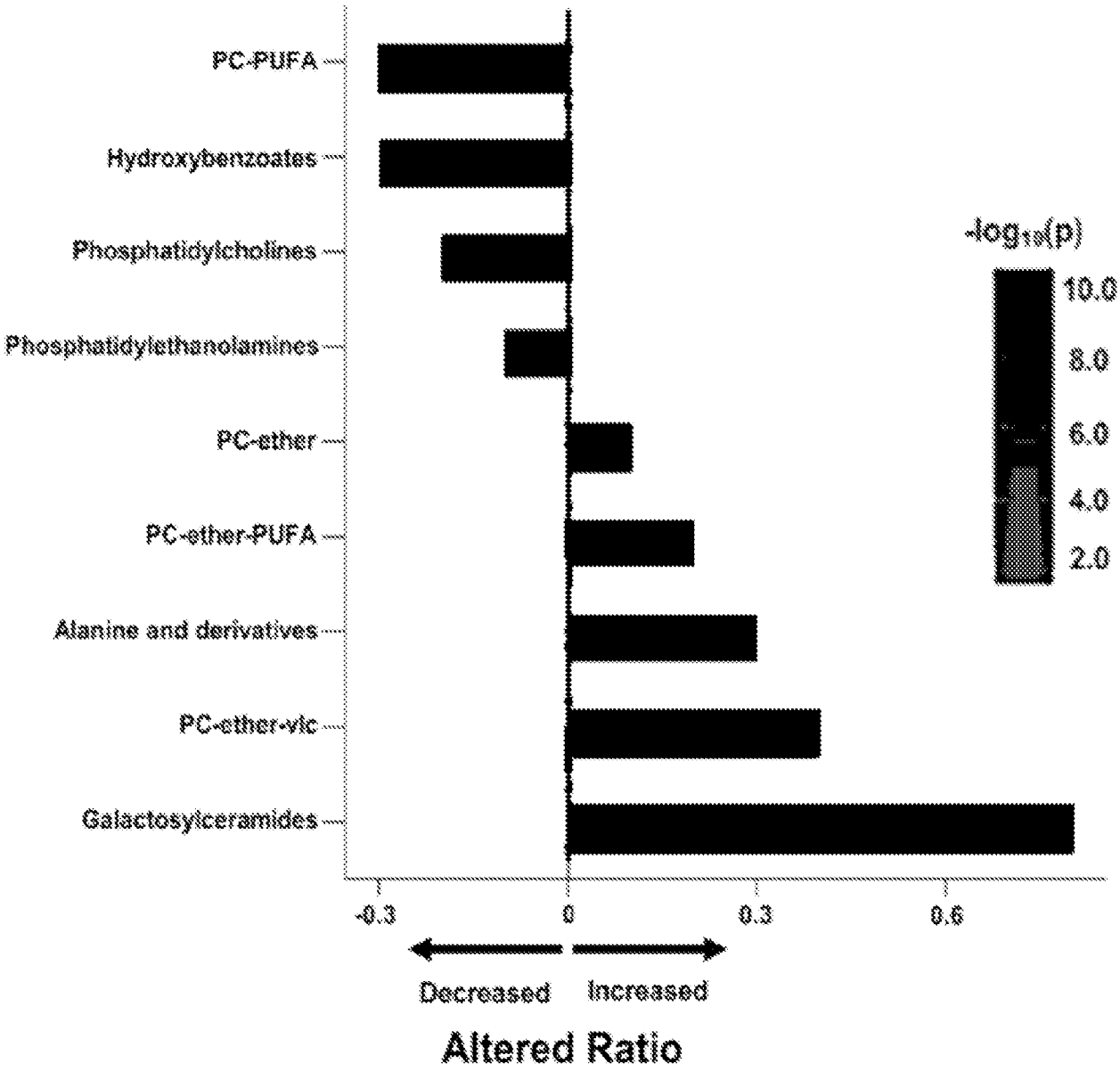

PC-PUFA
Hydroxybenzoates
Phosphatidylcholines
Phosphatidylethanolamines
PC-ether
PC-ether-PUFA
Alanine and derivatives
PC-ether-vlc
Galactosylceramides
-log10(p)
10.0
8.0
6.0
4.0
2.0
-0.3
0
0.3
0.6
Decreased
Increased
Altered Ratio

Fig. 1B

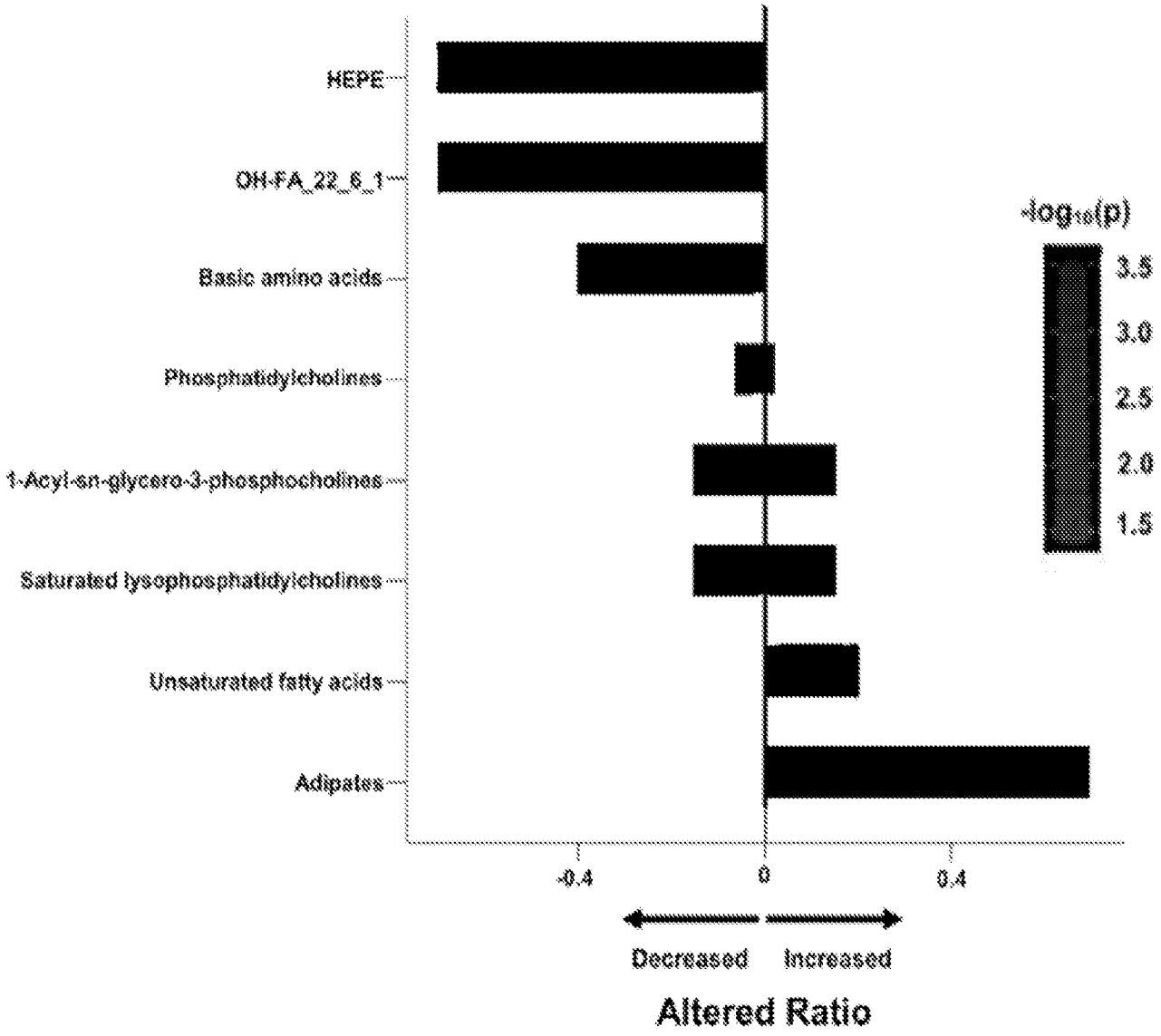

HEPE
OH-FA_22_6_1
Basic amino acids
Phosphatidylcholines
1-Acyl-sn-glycero-3-phosphocholines
Saturated lysophosphatidylcholines
Unsaturated fatty acids
Adipates
-log10(p)
3.5
3.0
2.5
2.0
1.5
-0.4
0
0.4
Decreased
Increased
Altered Ratio

Fig. 1C

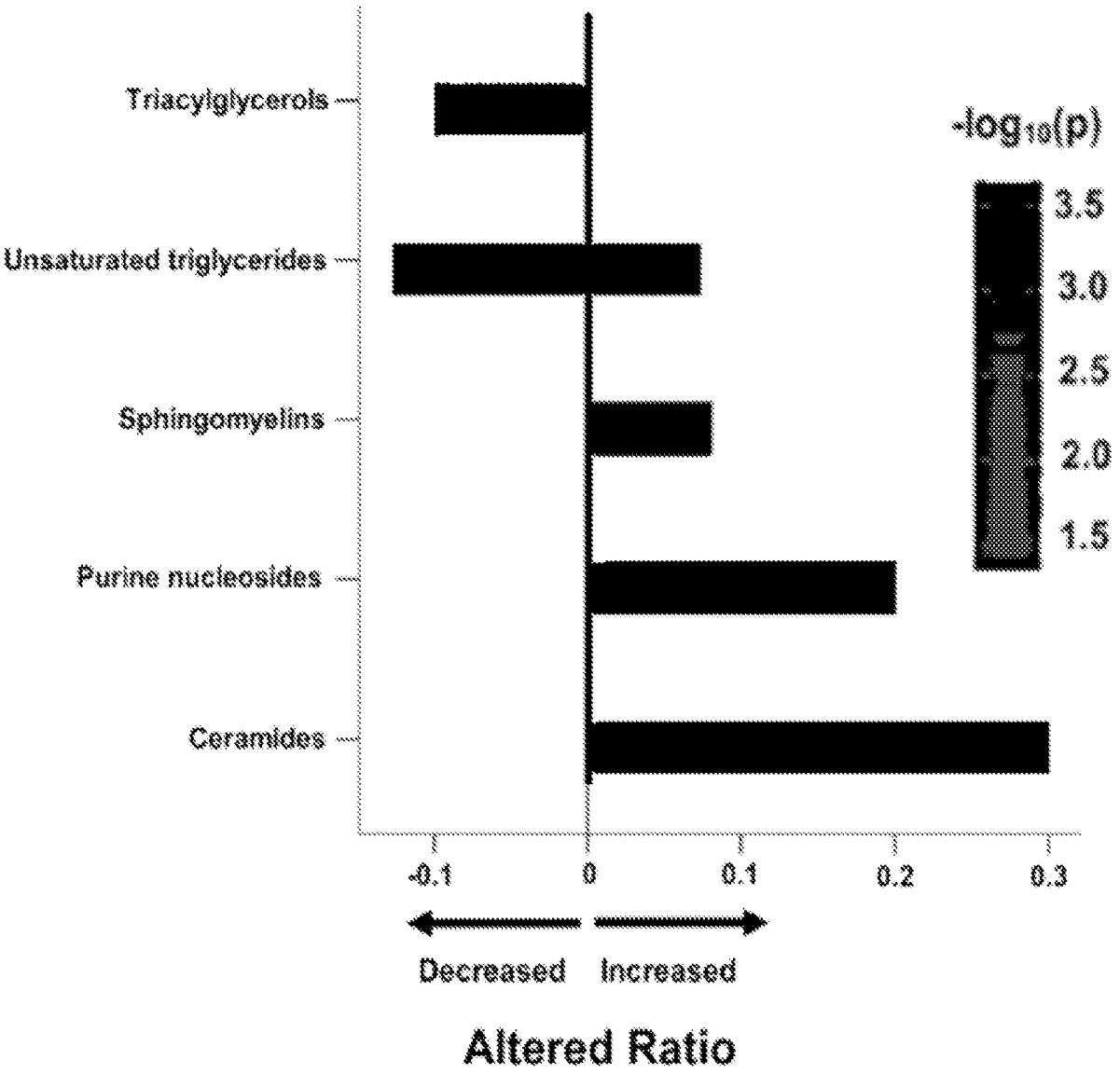

Triacylglycerols
Unsaturated triglycerides
Sphingomyelins
Purine nucleosides
Ceramides
-log10(p)
3.5
3.0
2.5
2.0
1.5
-0.1
0
0.1
0.2
0.3
Decreased
Increased
Altered Ratio

Fig. 1D

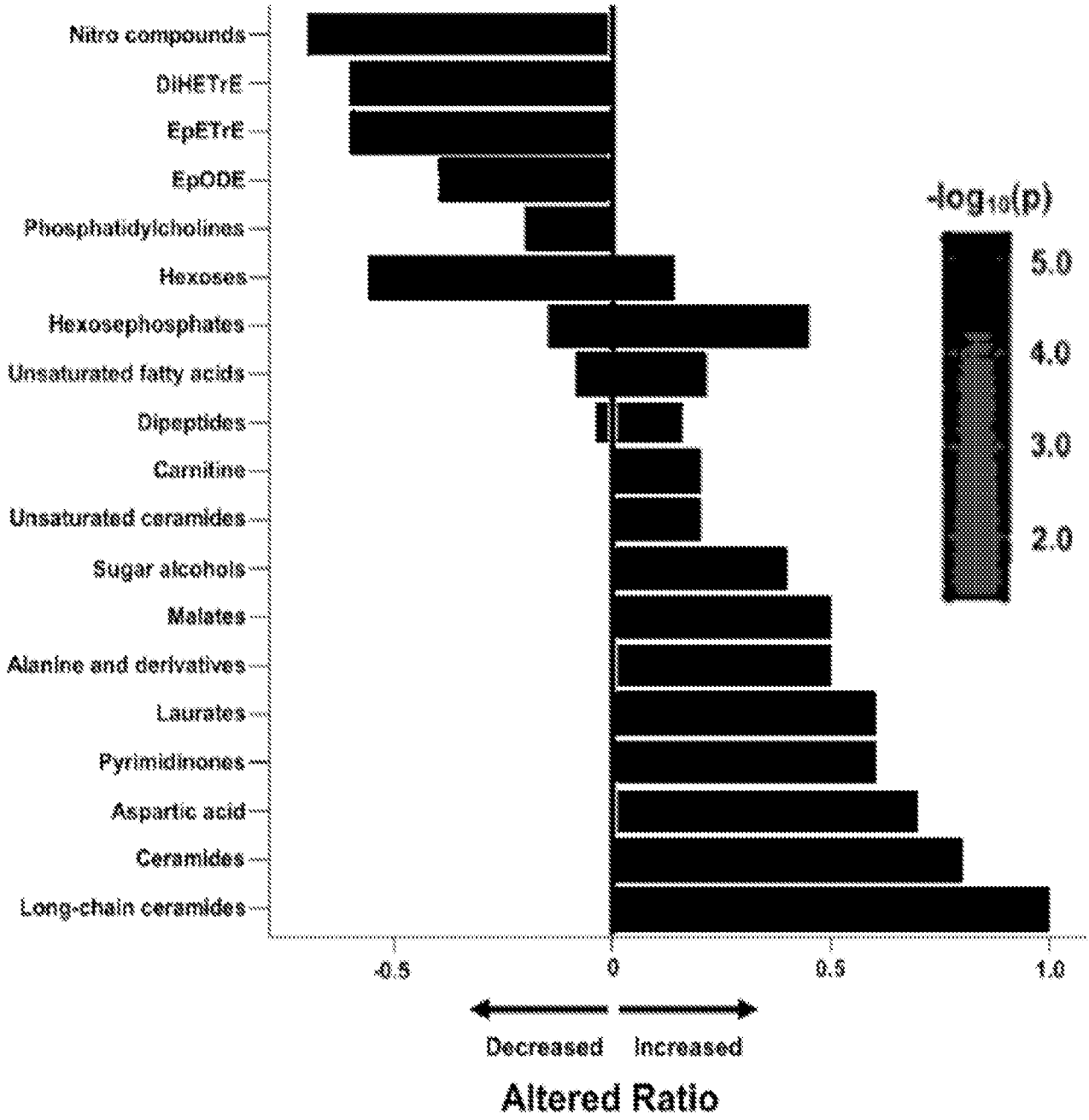

Nitro compounds
DiHETrE
EpETrE
EpODE
Phosphatidylcholines
Hexoses
Hexosephosphates
Unsaturated fatty acids
Dipeptides
Carnitine
Unsaturated ceramides
Sugar alcohols
Malates
Alanine and derivatives
Laurates
Pyrimidinones
Aspartic acid
Ceramides
Long-chain ceramides
-log10(p)
5.0
4.0
3.0
2.0
-0.5
0
0.5
1.0
Decreased
Increased
Altered Ratio

Fig. 2A

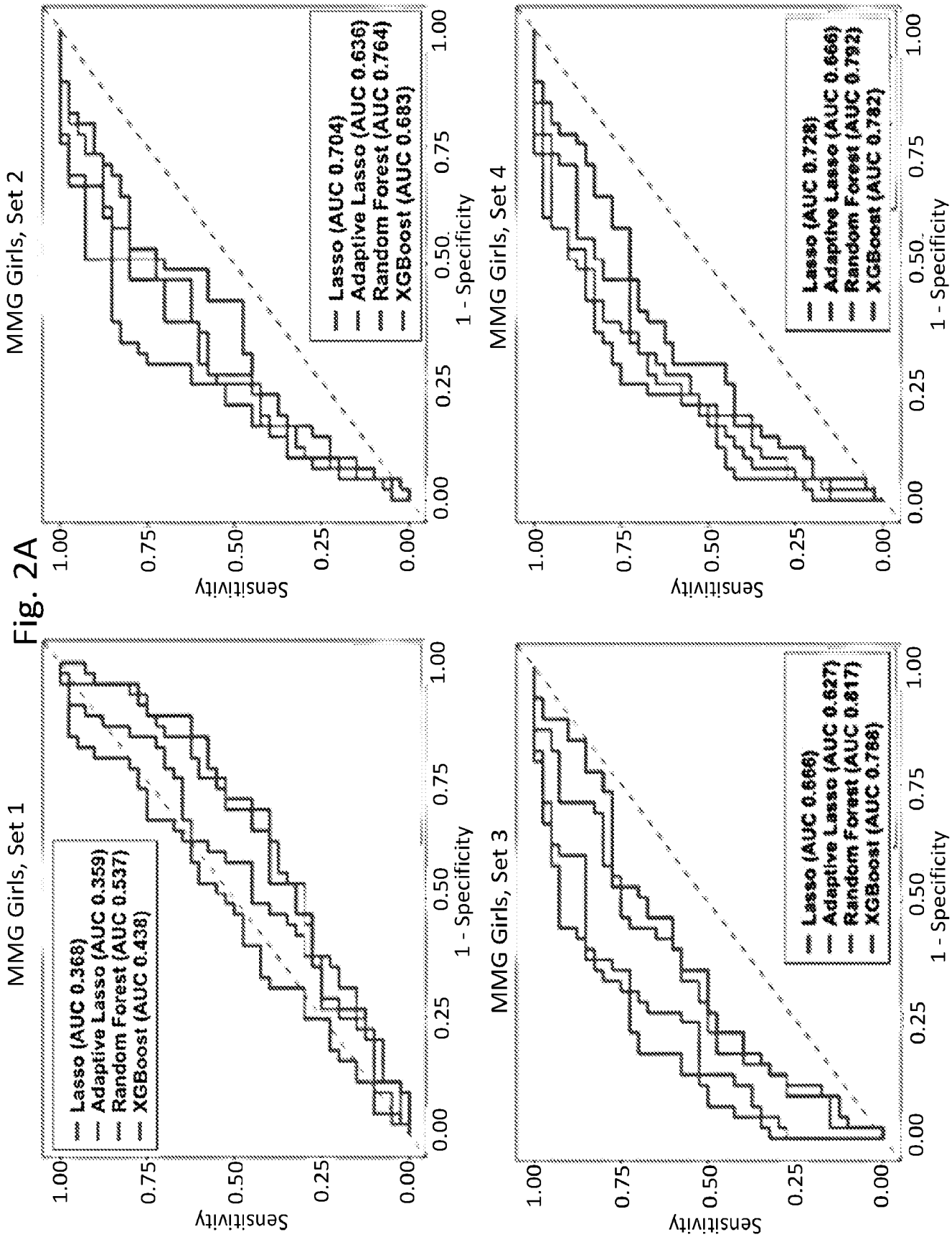

MMG Girls, Set 1
Lasso (AUC 0.368)
Adaptive Lasso (AUC 0.359)
Random Forest (AUC 0.537)
XGBoost (AUC 0.438)
MMG Girls, Set 2
Lasso (AUC 0.704)
Adaptive Lasso (AUC 0.636)
Random Forest (AUC 0.764)
XGBoost (AUC 0.683)
MMG Girls, Set 3
Lasso (AUC 0.666)
Adaptive Lasso (AUC 0.627)
Random Forest (AUC 0.817)
XGBoost (AUC 0.788)
MMG Girls, Set 4
Lasso (AUC 0.728)
Adaptive Lasso (AUC 0.666)
Random Forest (AUC 0.792)
XGBoost (AUC 0.782)
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00

Fig. 2B

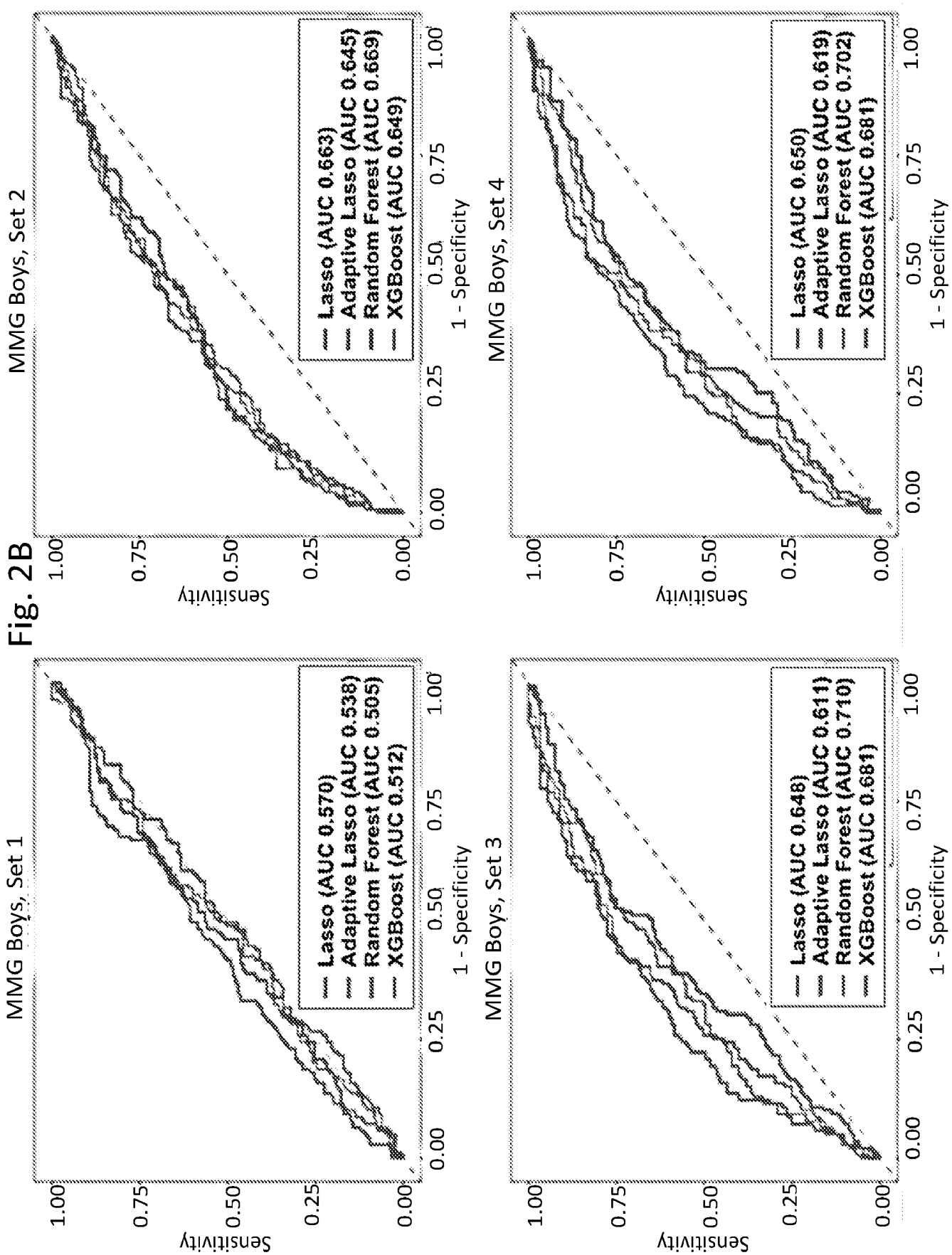

MMG Boys, Set 1
Lasso (AUC 0.570)
Adaptive Lasso (AUC 0.538)
Random Forest (AUC 0.505)
XGBoost (AUC 0.512)
MMG Boys, Set 2
Lasso (AUC 0.663)
Adaptive Lasso (AUC 0.645)
Random Forest (AUC 0.669)
XGBoost (AUC 0.649)
MMG Boys, Set 3
Lasso (AUC 0.648)
Adaptive Lasso (AUC 0.611)
Random Forest (AUC 0.710)
XGBoost (AUC 0.681)
MMG Boys, Set 4
Lasso (AUC 0.650)
Adaptive Lasso (AUC 0.619)
Random Forest (AUC 0.702)
XGBoost (AUC 0.681)
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00

Fig. 2C

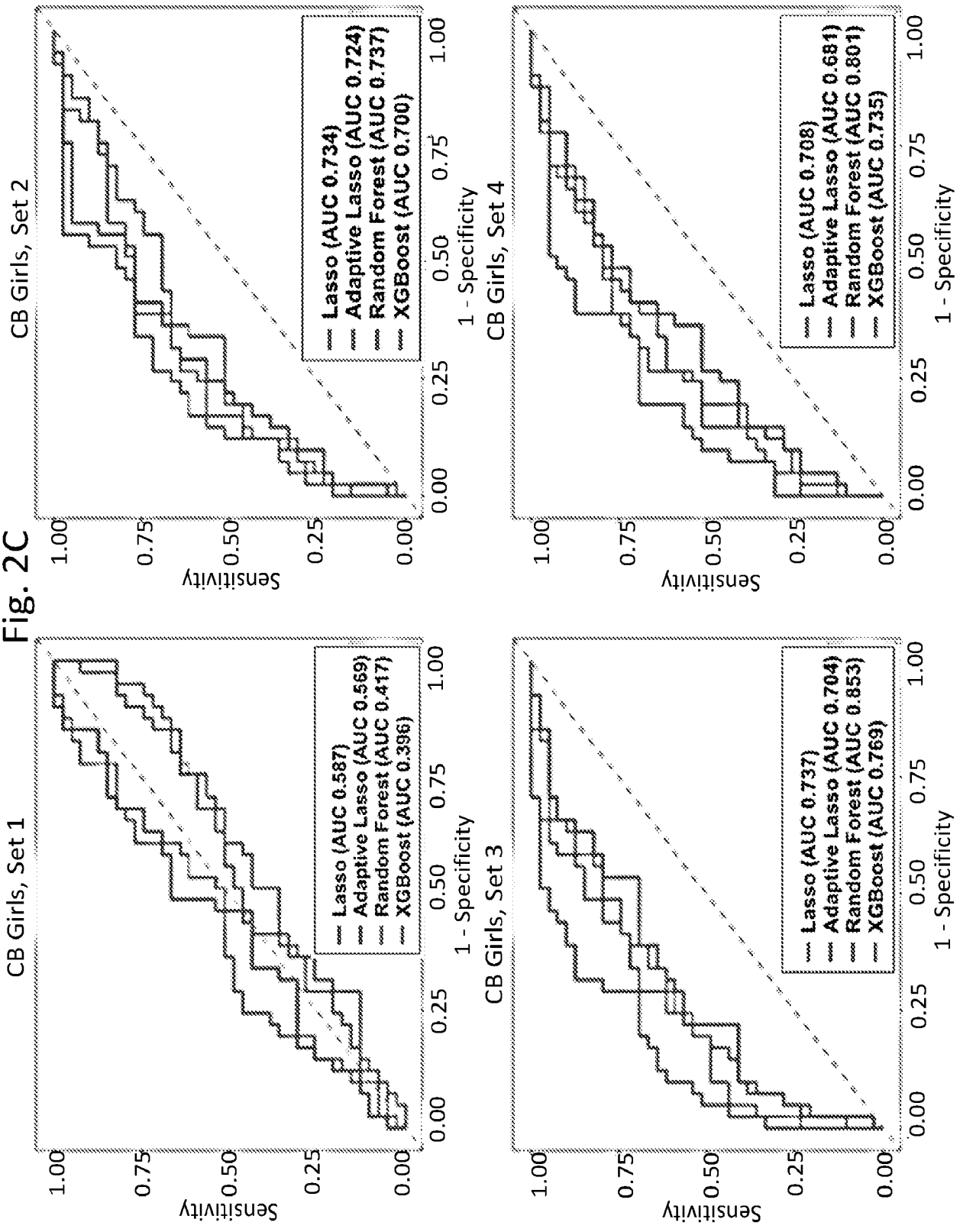

CB Girls, Set 1
Lasso (AUC 0.587)
Adaptive Lasso (AUC 0.569)
Random Forest (AUC 0.417)
XGBoost (AUC 0.396)
CB Girls, Set 2
Lasso (AUC 0.734)
Adaptive Lasso (AUC 0.724)
Random Forest (AUC 0.737)
XGBoost (AUC 0.700)
CB Girls, Set 3
Lasso (AUC 0.737)
Adaptive Lasso (AUC 0.704)
Random Forest (AUC 0.853)
XGBoost (AUC 0.769)
CB Girls, Set 4
Lasso (AUC 0.708)
Adaptive Lasso (AUC 0.681)
Random Forest (AUC 0.801)
XGBoost (AUC 0.735)
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00

Fig. 2D

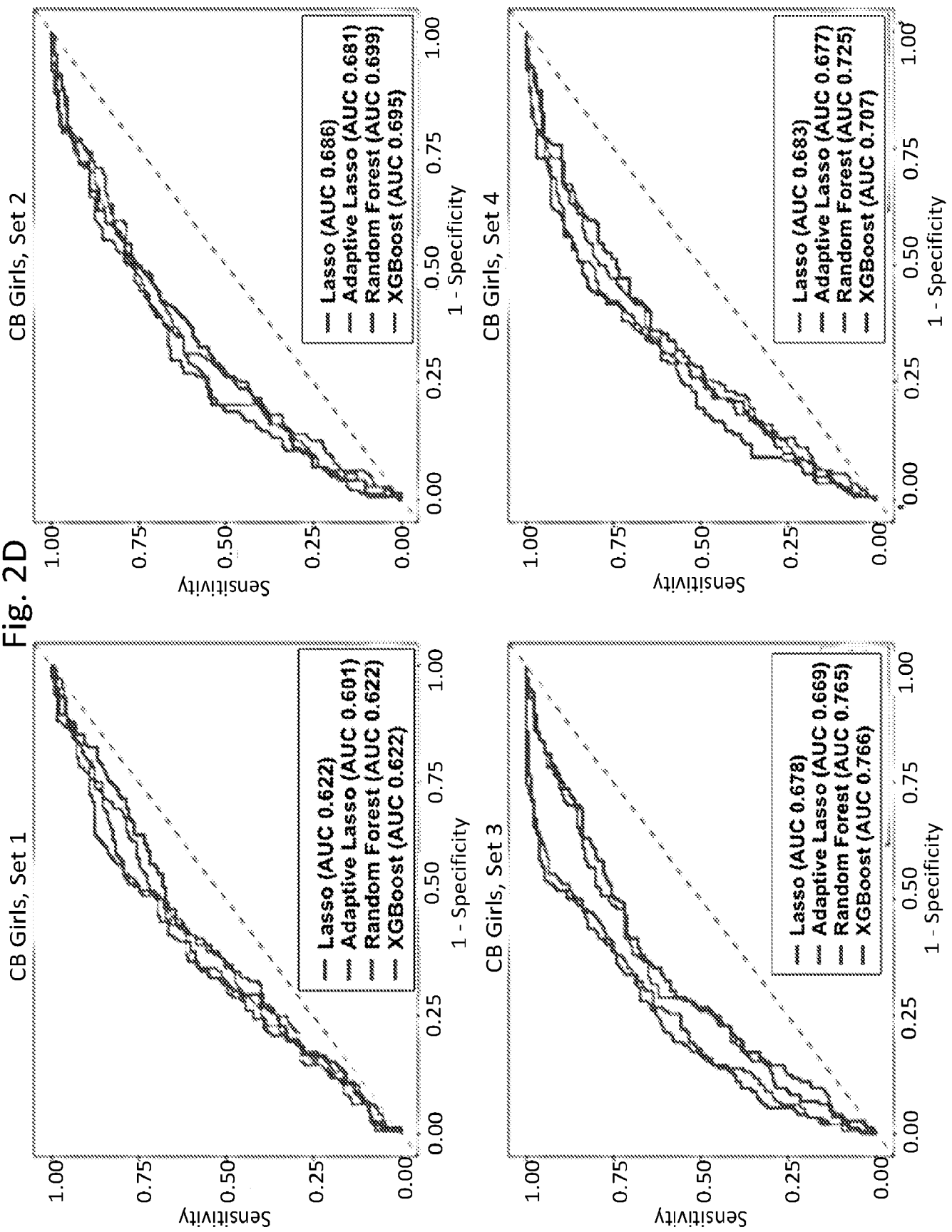

CB Girls, Set 1
Lasso (AUC 0.622)
Adaptive Lasso (AUC 0.601)
Random Forest (AUC 0.622)
XGBoost (AUC 0.622)
CB Girls, Set 2
Lasso (AUC 0.686)
Adaptive Lasso (AUC 0.681)
Random Forest (AUC 0.699)
XGBoost (AUC 0.695)
CB Girls, Set 3
Lasso (AUC 0.678)
Adaptive Lasso (AUC 0.669)
Random Forest (AUC 0.765)
XGBoost (AUC 0.766)
CB Girls, Set 4
Lasso (AUC 0.683)
Adaptive Lasso (AUC 0.677)
Random Forest (AUC 0.725)
XGBoost (AUC 0.707)
Sensitivity
1 - Specificity
0.00
0.25
0.50
0.75
1.00

Fig. 3A

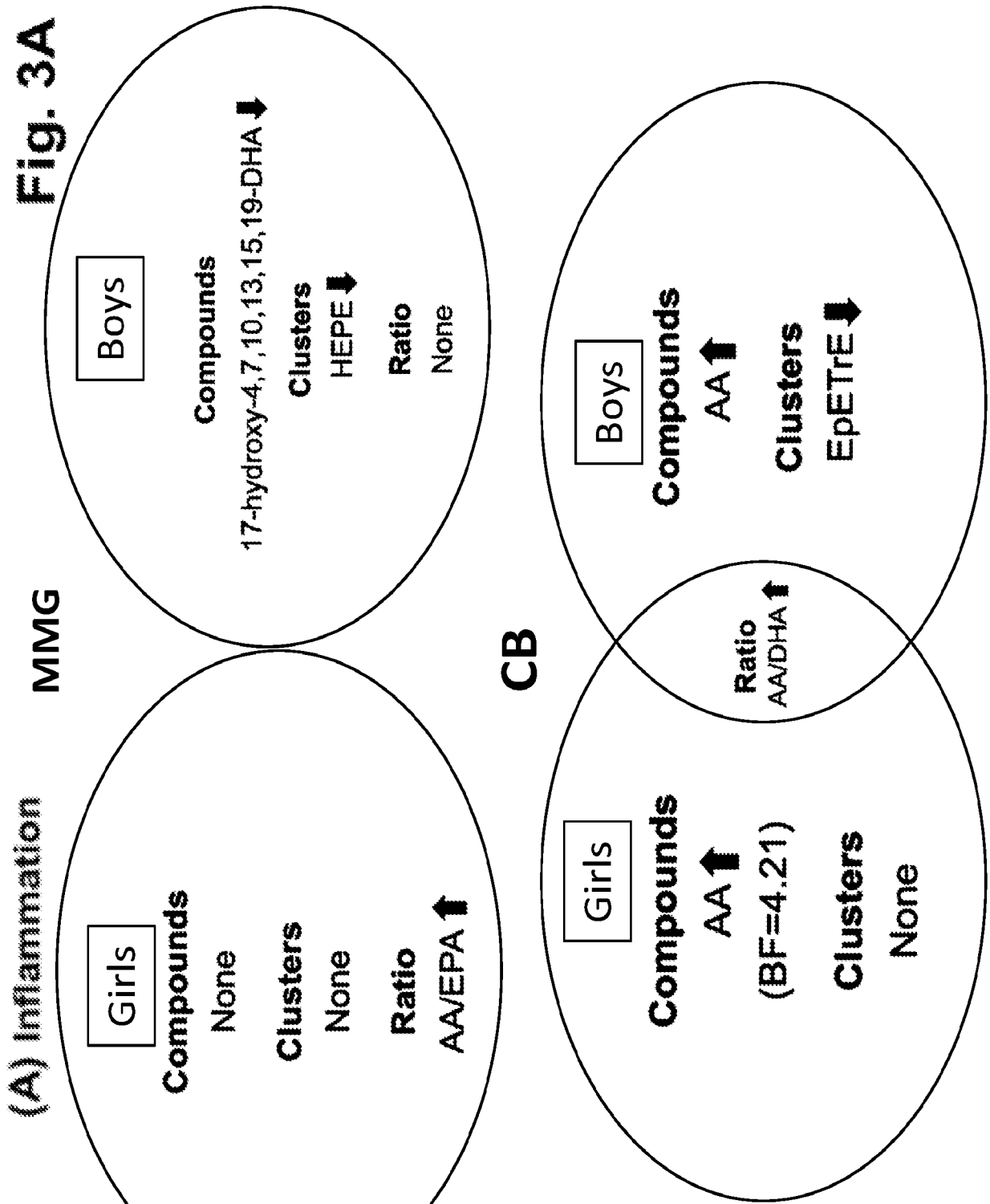

(A) Inflammation
MMG
Girls
Compounds
None
Clusters
None
Ratio
AA/EPA ⬆
Boys
Compounds
17-hydroxy-4,7,10,13,15,19-DHA ⬇
Clusters
HEPE ⬇
Ratio
None
CB
Girls
Compounds
AA ⬆
(BF=4.21)
Clusters
None
Ratio
AA/DHA ⬆
Boys
Compounds
AA ⬆
Clusters
EpETrE ⬇

Fig. 3B

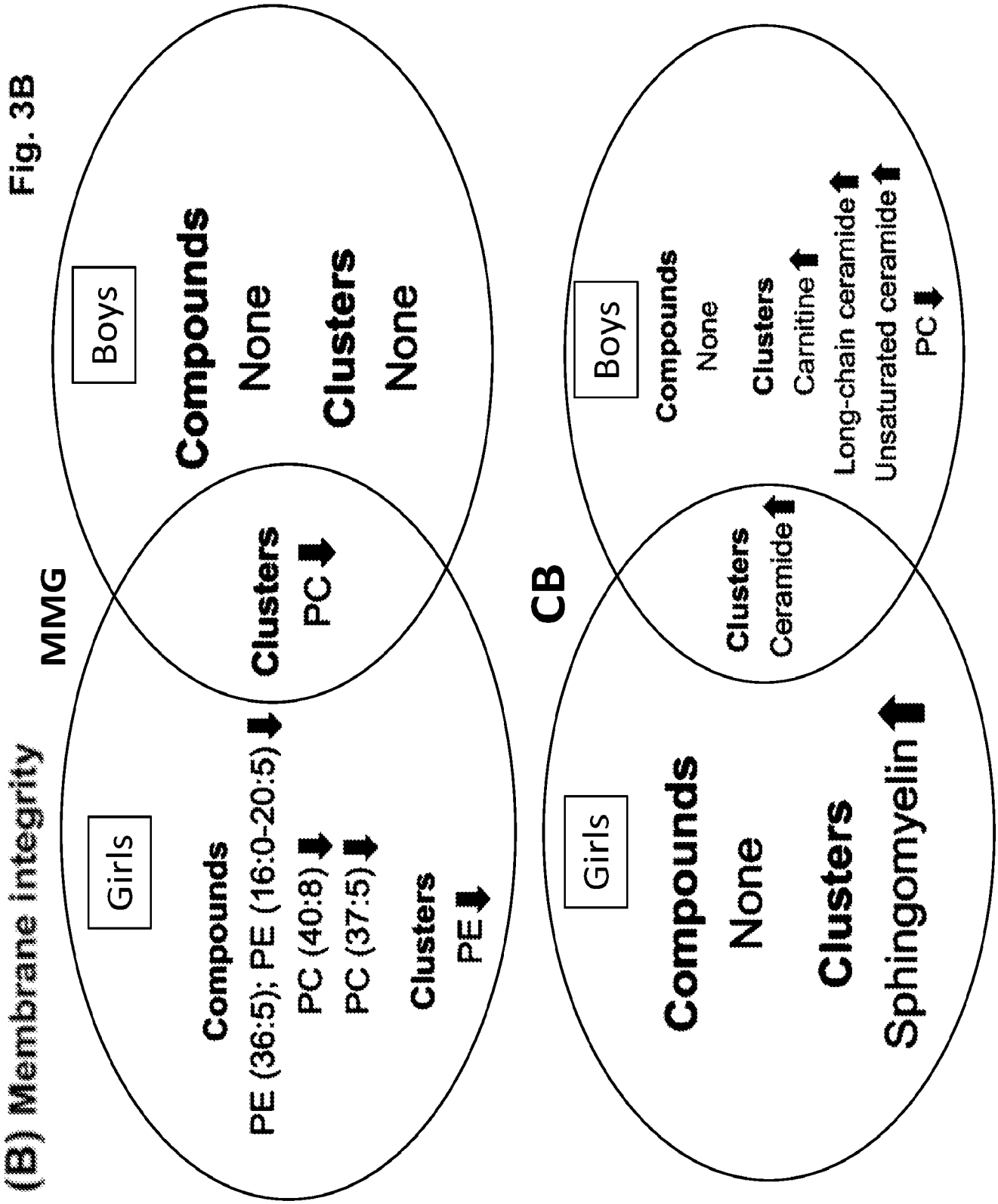
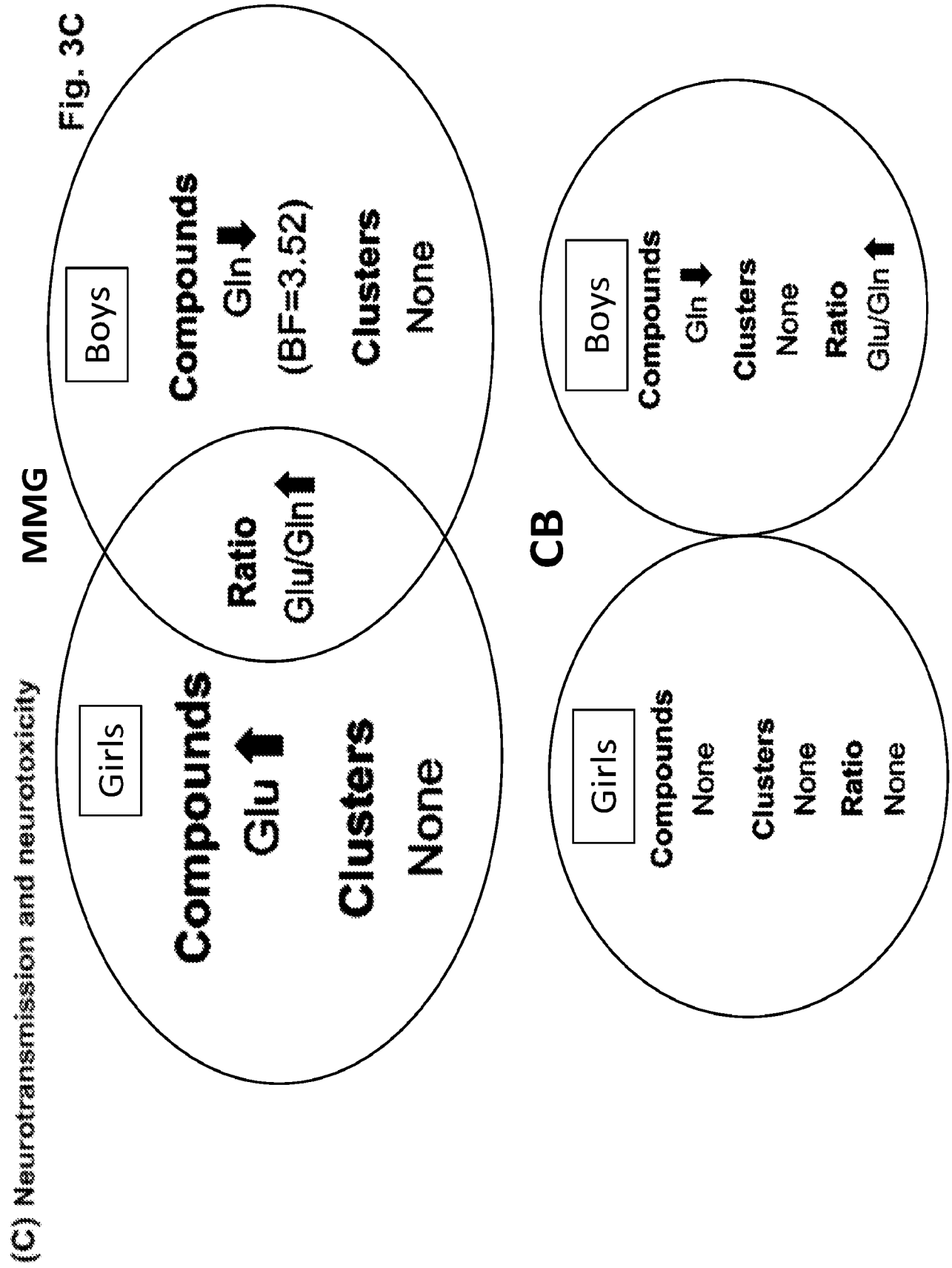

(B) Membrane integrity
MMG
Girls
Compounds
PE (36:5); PE (16:0-20:5)
PC (40:8)
PC (37:5)
Clusters
PE
Clusters
PC
Boys
Compounds
None
Clusters
None
CB
Girls
Compounds
None
Clusters
Sphingomyelin
Clusters
Ceramide
Boys
Compounds
None
Clusters
Carnitine
Long-chain ceramide
Unsaturated ceramide
PC

Fig. 3C


(C) Neurotransmission and neurotoxicity
MMG
Girls
Compounds
Glu ⬆
Clusters
None
Ratio
Glu/Gln ⬆
Boys
Compounds
Gln ⬇
(BF=3.52)
Clusters
None
CB
Girls
Compounds
None
Clusters
None
Ratio
None
Boys
Compounds
Gln ⬇
Clusters
None
Ratio
Glu/Gln ⬆

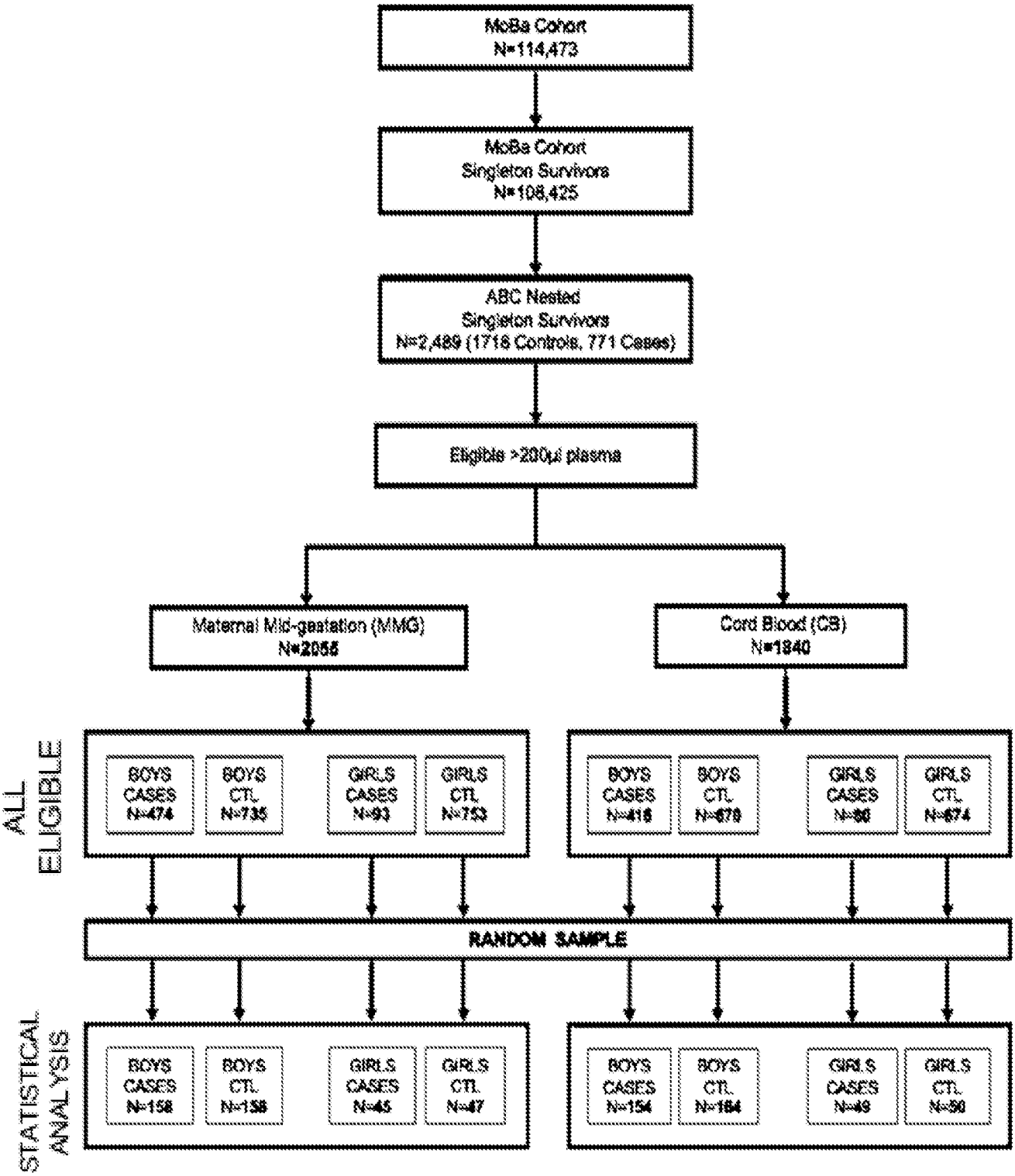

MoBa Cohort
N=114,473
MoBa Cohort
Singleton Survivors
N=108,425
ABC Nested
Singleton Survivors
N=2,489 (1716 Controls, 771 Cases)
Eligible >200µl plasma
Maternal Mid-gestation (MMG)
N=2055
Cord Blood (CB)
N=1940
ALL ELIGIBLE
BOYS CASES N=474
BOYS CTL N=735
GIRLS CASES N=93
GIRLS CTL N=753
BOYS CASES N=418
BOYS CTL N=678
GIRLS CASES N=86
GIRLS CTL N=674
RANDOM SAMPLE
STATISTICAL ANALYSIS
BOYS CASES N=158
BOYS CTL N=158
GIRLS CASES N=45
GIRLS CTL N=47
BOYS CASES N=154
BOYS CTL N=184
GIRLS CASES N=49
GIRLS CTL N=50

Fig. 4

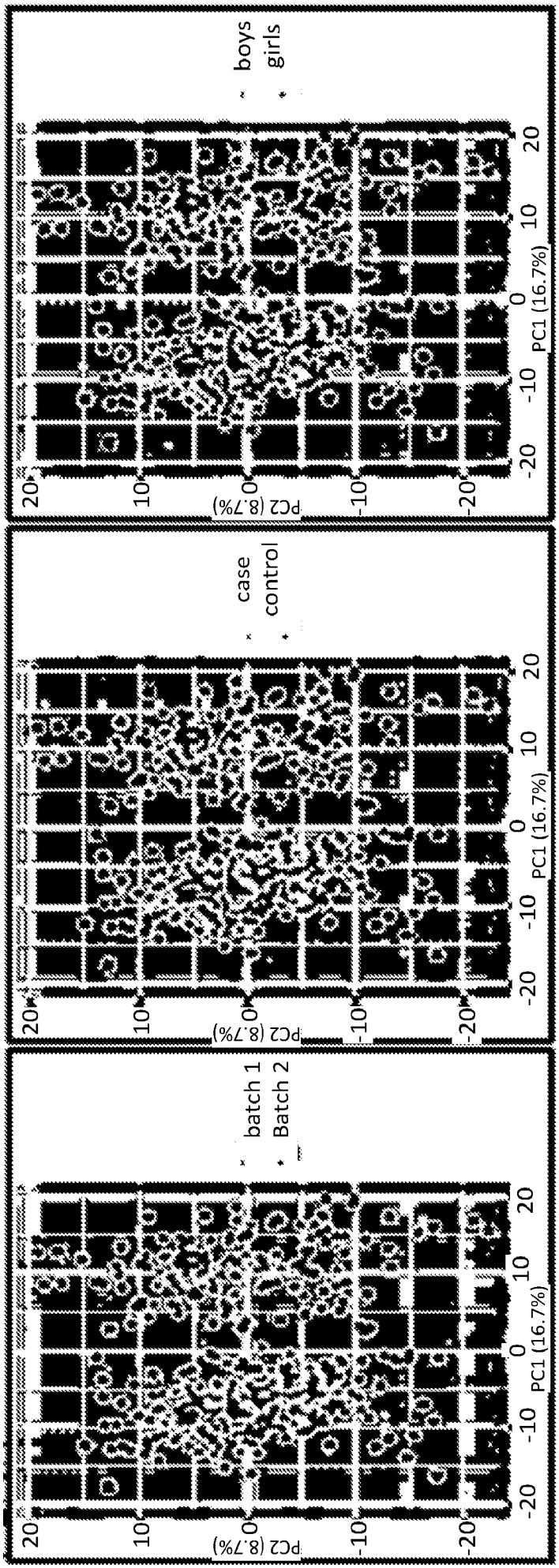

batch 1
Batch 2
case
control
boys
girls
PC1 (16.7%)
PC2 (8.7%)
20
10
0
-10
-20

Fig. 5A

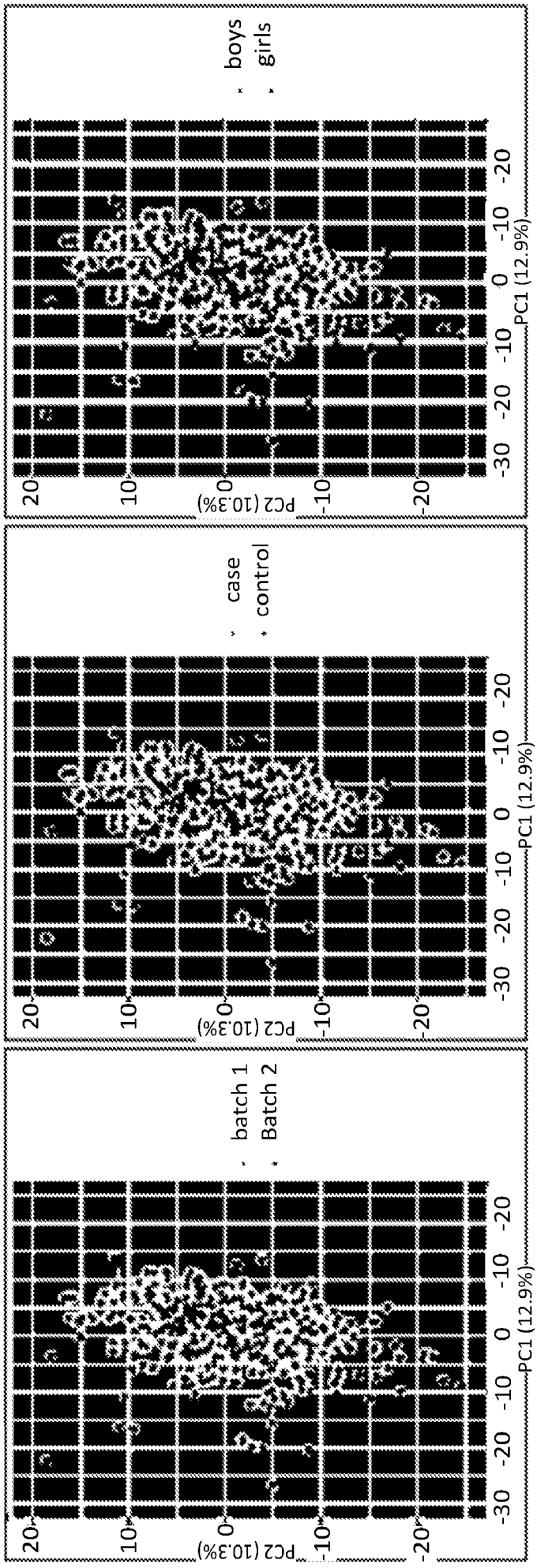

boys
girls
case
control
batch 1
Batch 2
PC1 (12.9%)
PC2 (10.3%)
20
10
0
-10
-20
-30

Fig. 5B