

OPEN

Proteomic Investigations of Autism Brain Identify Known and Novel Pathogenetic Processes

Joseph R. Abraham¹, Nicholas Szoko¹, John Barnard², Robert A. Rubin³, Daniela Schlatter⁴, Kathleen Lundberg⁴, Xiaolin Li⁴ & Marvin R. Natowicz^{1,5}

Autism Spectrum Disorder (ASD) is a set of heterogeneous neurodevelopmental conditions defined by impairments in social communication and restricted, repetitive behaviors, interests or activities. Only a minority of ASD cases are determined to have a definitive etiology and the pathogenesis of most ASD is poorly understood. We hypothesized that a global analysis of the proteomes of human ASD vs. control brain, heretofore not done, would provide important data with which to better understand the underlying neurobiology of autism. In this study, we characterized the proteomes of two brain regions, Brodmann area 19 (BA19) and posterior inferior cerebellum (CB), from carefully selected idiopathic ASD cases and matched controls using label-free HPLC-tandem mass spectrometry. The data revealed marked differences between ASD and control brain proteomes for both brain regions. Unlike earlier transcriptomic analyses using frontal and temporal cortex, however, our proteomic analysis did not support ASD attenuating regional gene expression differences. Bioinformatic analyses of the differentially expressed proteins between cases and controls highlighted canonical pathways involving glutamate receptor signaling and glutathione-mediated detoxification in both BA19 and CB; other pathways such as Sertoli cell signaling and fatty acid oxidation were specifically enriched in BA19 or CB, respectively. Network analysis of both regions of ASD brain showed up-regulation of multiple pre- and post-synaptic membrane or scaffolding proteins including glutamatergic ion channels and related proteins, up-regulation of proteins involved in intracellular calcium signaling, and down-regulation of neurofilament proteins, with DLG4 and MAPT as major hub proteins in BA19 and CB protein interaction networks, respectively. Upstream regulator analysis suggests neurodegeneration-associated proteins drive the differential protein expression for ASD in both BA19 and CB. Overall, the proteomic data provide support for shared dysregulated pathways and upstream regulators for two brain regions in human ASD brain, suggesting a common ASD pathophysiology that has distinctive regional expression.

Autism Spectrum Disorder (ASD) is a clinically heterogeneous group of neurodevelopmental conditions characterized by early age-of-onset impairments in social communication and social interaction and restricted, repetitive patterns of behaviors, interests or activities^{1,2}. Persons with ASD commonly also have medical, neurological and/or psychiatric co-morbidities^{3,4} and, as ASD is a lifelong condition, there are many years lived with disability⁵. Consequently, ASD is associated with varied and considerable personal, family and societal challenges and costs⁶. ASD is also common, with an estimated global prevalence of about 1–1.5%^{7,8}. For these reasons, determination of the underlying pathobiology of ASD and efforts directed toward ASD prevention, early diagnosis, and effective treatment are public health priorities.

More than 100 mendelian and cytogenomic variants have been causally associated with ASD, all of which are uncommon or rare and many of which are incompletely characterized^{9–11}; environmental factors, while clearly relevant, are even less well understood¹². At this time, despite extensive diagnostic evaluation, the vast majority of persons with ASD do not receive an etiological diagnosis and receive a diagnosis of autism of unknown etiology

¹Cleveland Clinic Lerner College of Medicine of Case Western Reserve University, Cleveland, OH, 44195, USA.

²Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, 44195, USA. ³Department of Mathematics, Whittier College, Whittier, CA, 90602, USA. ⁴Center for Proteomics, Case Western Reserve University, Cleveland, OH, 44106, USA. ⁵Pathology and Laboratory Medicine, Genomic Medicine, Neurological and Pediatrics Institutes, Cleveland Clinic, Cleveland, OH, 44195, USA. Correspondence and requests for materials should be addressed to M.R.N. (email: natowim@ccf.org)

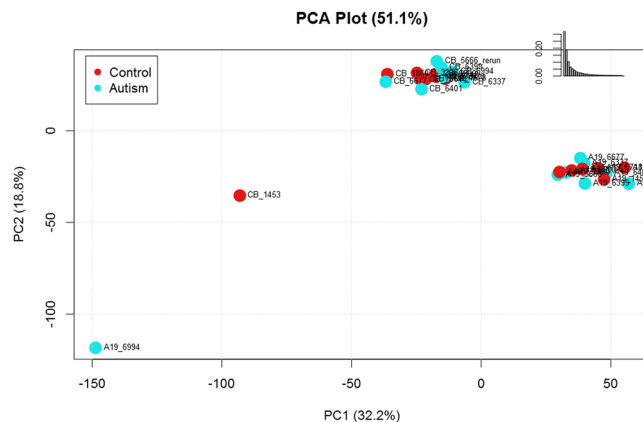


Figure 1. Principal component analysis of peptide expression data showing clear separation of samples by brain region and one outlier (BA19_6994).

or idiopathic autism; only about 10–25% of cases have defined pathogenic abnormalities that are determinative or significantly contributory to ASD causation¹⁰. The marked etiologic heterogeneity of ASD, the diversity of known and purported etiologies, and the predominance of idiopathic cases underscore the complexity and challenge inherent in uncovering the underlying pathogenesis of ASD¹³.

Diverse investigative approaches - from neuropathological evaluation to functional neuroimaging - suggest that persons with ASD have central nervous system (CNS) developmental dysconnectivity, and important theories of ASD pathogenesis center on abnormal synaptic homeostasis and the abnormal development or regulation of neuronal circuitry^{14–16}. Considerable efforts to unravel the mechanisms underlying the developmental dysconnectivity of ASD are ongoing. Molecular analyses have revealed a wide array of contributors to these processes including genes and proteins involved in diverse cellular processes including synapse biology, regulation of neuronal gene transcription and translation, Wnt signaling and mitochondrial bioenergetics, among others^{10,17–21}. There is, however, a notable gap in the pathophysiological characterization of ASD at the level of the proteome.

Proteomics, the large-scale study of protein expression in cells or tissues, enables an analysis of protein expression profiles and the determination of biologically relevant protein interaction networks and signaling pathways²². Different proteomic methodologies and experimental designs can be used for diverse purposes, from providing an unbiased relative quantification of a large number of proteins expressed in cells or tissues, often designated as label-free expression proteomics, to the quantification of particular targeted proteins or isoforms of proteins, or groups of proteins having a specific post-translational modification²³. Proteomic analysis affords a potentially complementary approach to genetic analyses in characterizing the pathophysiology of a phenotype and has an advantage in being ‘closer’ to the expression of the cellular or tissue phenotype than studies done at a genomic or transcriptomic level.

Despite the power of proteomics and its increasing utilization in the understanding of neurological and developmental disorders, there are few applications of proteomic approaches to the study of ASD^{24,25}. Moreover, most of the proteomic studies of ASD have examined non-neural tissue and no consistent findings have been apparent²⁵.

In this study, we sought to perform a global label-free proteomic analysis of stringently selected idiopathic ASD post-mortem brain samples and carefully matched controls as a means to better understand ASD pathobiology. Two brain regions were evaluated in cases and controls: Brodmann area 19 region of occipital cortex (BA19) and the left posterior inferior cerebellar hemispheric cortex (CB); both regions are affected in ASD and have been studied extensively^{26–29}. We hypothesized that brain proteomic analysis would show distinctive protein expression profiles between the two brain regions and also between ASD and controls for these regions and that analyses of these differences would provide insights regarding the underlying biology of ASD.

Results

The proteomes of BA19 and CB are distinct from each other and between ASD cases and controls. The proteomic analysis of ASD and control samples detected 9205 peptides belonging to 1807 proteins in BA19 occipital cortex and 8017 peptides from 1640 proteins in cerebellar (CB) tissue. For BA19, 63% of proteins were identified by more than one uniquely identifying peptide; similarly, 64% of CB proteins were identified by more than one peptide. Principal component analysis, based on the peptides noted in each brain region, revealed complete separation of the two regions, indicating the differences in protein expression between BA19 and cerebellum (Fig. 1).

Analysis of proteins identified by at least 2 unique peptides revealed 146 proteins from BA19 having significant differential expression between ASD and controls ($p < 0.05$) with 27 remaining significant after false discovery rate (FDR) adjustment (Benjamini-Hochberg (BH) adjusted $p < 0.05$); 191 proteins were found to be significantly differentially expressed in CB ($p < 0.05$) with 38 proteins remaining significant after FDR adjustment (BH adjusted $p < 0.05$). Supplementary Table 1 provides subject characteristics and Supplementary Table 2 notes all differentially expressed proteins identified by ≥ 2 peptides from each brain region with p-values, BH adjusted p-values, and associated percent expression changes.

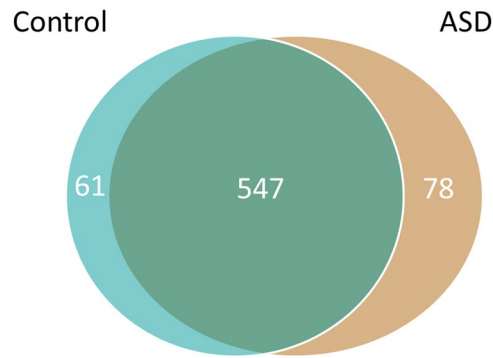


Figure 2. Differentially expressed proteins between cerebellar cortex and occipital cortex (BA19) comparing ASD and control brains. The differentially expressed proteins common to both types of brain as well as numbers unique to each are noted.

Canonical Pathways	Enrichment Score (BH-adjusted)
FXR/RXR Activation	1.61
Glutamate Receptor Signaling	1.61
Clathrin-mediated Endocytosis Signaling	1.61
Heme Degradation	1.43
Sertoli Cell-Cell Junction Signaling	1.41
Glutathione-mediated Detoxification	1.41

Table 1. BA19-enriched canonical pathways from differentially expressed proteins ($p < 0.05$), adjusted for multiple comparisons. Enrichment score of 1.3 equals a BH adjusted p-value of 0.05.

Brain regional protein differences are maintained in ASD. Previous transcriptomic work showed an attenuation of normal gene expression differences between frontal and temporal cortex in ASD compared to control brains^{30,31}. Analysis of proteins differentially expressed ($p < 0.05$) by brain region in our ASD and control data sets showed 625 differentially expressed proteins noted between the BA19 and CB regions in the ASD brains and 608 differentially expressed proteins noted between the BA19 and CB regions in the control brains. The vast majority of these differentially expressed proteins, 547 proteins, were shared between ASD and control brains; of these 544 showed identical directionality of the differential expression between the brain regions for both ASD and controls (Fig. 2; Supplementary Table 3). These data indicate insignificant differences comparing BA19 vs. CB regional differentiation in ASD vs. control brains.

Proteomic canonical pathway analysis reveals differentially enriched pathways in ASD vs. control brains, with BA19 and CB having both shared and distinctive pathway enrichments. To more fully understand the meanings of the differentially expressed proteins of each brain region, we used the Ingenuity Pathway Analysis[®] (IPA) bioinformatics tool to determine whether specific pathways and upstream regulators are enriched in ASD cases vs. controls. We found enrichment in ASD cases vs. controls for 58 and 86 canonical pathways for BA19 and CB, respectively, at p -value < 0.05 , with 27 canonical pathways enriched in both BA19 and CB. Adjusting for multiple comparisons using Benjamini-Hochberg showed enrichment for 6 canonical pathways in BA19 and for 36 pathways in CB (Tables 1, 2 and Supplementary Table 4). Pathways enriched in both brain regions include glutamate receptor signaling and glutathione-mediated detoxification/oxidative stress. Pathways specifically enriched in one of the two regions include, for example, enrichment for Sertoli cell signaling in BA19 and enrichment for pathways of fatty acid, organic acid and branched chain amino acid metabolism in CB. Canonical pathway analysis using an independent bioinformatics tool, Reactome, highlighted major themes of the IPA-based analysis including glutamate (NMDA) receptor signaling in both brain regions and mitochondrial fatty acid and branched chain amino acid catabolism in CB (Supplementary Table 5).

Protein interaction network analyses differ between BA19 and CB. We also sought to determine any protein-protein interaction networks that might also be differentially expressed between ASD cases and controls. Two significant interaction networks as determined by IPA, one for each brain region, are noted in Fig. 3. The BA19 network highlights proteins of synaptic function, cytoskeletal infrastructure of axon/neuron morphology, intracellular vesicle/organelle transport, and glutamatergic neurotransmission, with DLG4 as the major hub protein. The CB network highlights cytoskeletal and intracellular vesicular transport-related proteins, especially tubulins and tubulin-associated proteins, with MAPT as the major hub protein.

Networks in both brain regions share common features including: (1) up-regulation of multiple pre- and post-synaptic membrane or scaffolding proteins (eg. FLOT1, DLG2, DLG4, SHANK3 in BA19 and SYP, VAMP2, BSN, DLG2 in CB); (2) up-regulation of glutamatergic ion channels and related proteins (in BA19: FLOT1, CASK,

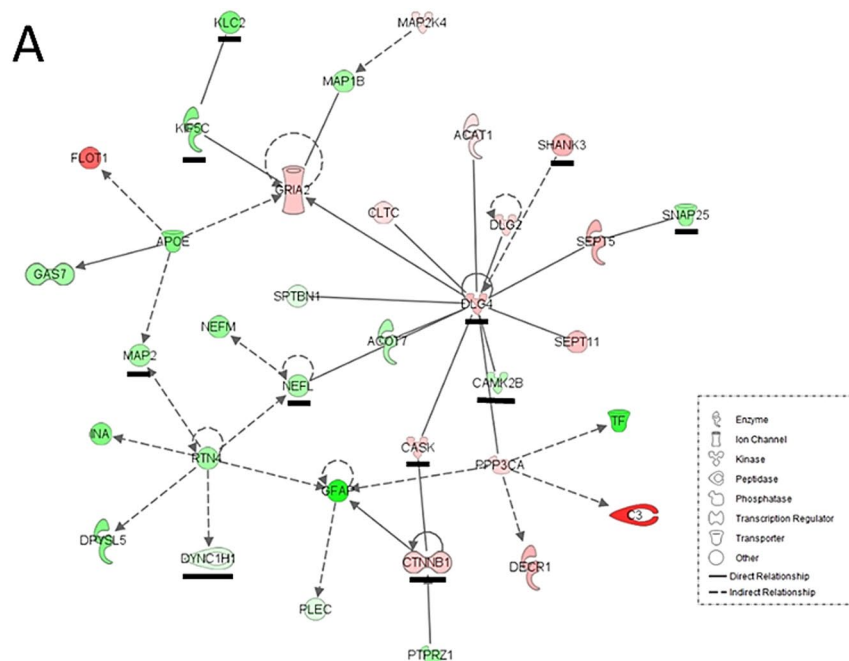
Canonical Pathways	Enrichment Score (BH-adjusted)
Fatty Acid β -oxidation I	6.96
Glutaryl-CoA Degradation	6.14
Tryptophan Degradation III (Eukaryotic)	5.48
Isoleucine Degradation I	4.85
Valine Degradation I	4.43
Ketolysis	4.43
Aryl Hydrocarbon Receptor Signaling	2.93
Calcium Signaling	2.36
Xenobiotic Metabolism Signaling	2.36
Calcium Transport I	2.36
Phagosome Maturation	2.36
Ketogenesis	2.28
Amyotrophic Lateral Sclerosis Signaling	2.17
NRF2-mediated Oxidative Stress Response	2.17
Ethanol Degradation II	2.15
Mevalonate Pathway I	2.15
Noradrenaline and Adrenaline Degradation	2.03
14-3-3-mediated Signaling	1.96
Superpathway of Geranylgeranyldiphosphate Biosynthesis I (via Mevalonate)	1.94
LXR/RXR Activation	1.94
Parkinson's Signaling	1.88
LPS/IL-1 Mediated Inhibition of RXR Function	1.79
Melatonin Signaling	1.78
Pyrimidine Deoxyribonucleotides <i>De Novo</i> Biosynthesis I	1.68
Tryptophan Degradation X (Mammalian, via Tryptamine)	1.68
Glutathione-mediated Detoxification	1.68
nNOS Signaling in Neurons	1.65
Ethanol Degradation IV	1.65
Serotonin Degradation	1.64
Synaptic Long Term Potentiation	1.6
PI3K Signaling in B Lymphocytes	1.54
PPAR α /RXR α Activation	1.5
Calcium-induced T Lymphocyte Apoptosis	1.5
Superpathway of Cholesterol Biosynthesis	1.41
Glutamate Receptor Signaling	1.35
CREB Signaling in Neurons	1.3

Table 2. CB-enriched canonical pathways from differentially expressed proteins ($p < 0.05$), adjusted for multiple comparisons. Enrichment score of 1.3 equals a BH adjusted p-value of 0.05.

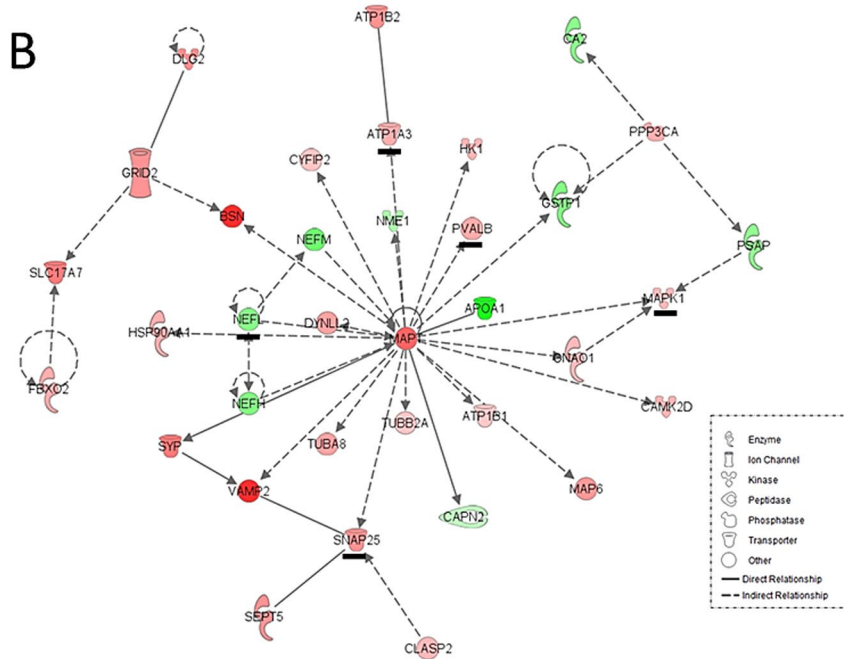
GRIA2, DLG4 and DLG2; in CB: GRID2, SLC17A7 and DLG2); (3) up-regulation of proteins involved in intracellular calcium signaling (eg, PPP3CA, CASK in BA19 and CAMK2D, PPP3CA in CB); and (4) down-regulation of multiple neurofilament proteins (eg, NEFL, NEFM, INA in BA19 and NEFL, NEFM, NEFH in CB). A notable protein interaction difference between the networks is the up-regulation in CB of proteins comprising microtubules or involved in microtubule assembly or stability (eg, TUBA8, TUBB2A, MAP6, CLASP2, MAPT).

Upstream regulator analysis suggests shared molecular dysregulation between ASD and neurodegeneration-associated proteins. Upstream regulators are molecules that are predicted to regulate the expression of the differentially expressed proteins. Using IPA, we noted 134 and 114 upstream regulators in BA19 and in CB, respectively, at p -value < 0.05 ; 70 of these were in common for both brain regions (Table 3, Supplementary Table 6). Functional assessment of these regulators revealed many that are involved in intracellular vesicular biology, synaptic biology, transcription, translation, and various metabolic functions. The most statistically significant upstream regulators are identical for both brain regions and include proteins associated with CNS neurodegenerative processes: MAPT, APP, PSEN1 and HTT.

Selective reaction monitoring (SRM) validates the global, label-free proteomic analysis. To independently validate the global, label-free proteomic findings, we performed a targeted quantitative MS



© 2000-2018 QIAGEN. All rights reserved.



© 2000-2018 QIAGEN. All rights reserved.

Figure 3. (A) BA19 empirical protein-protein interaction network generated from differentially expressed proteins. (B) CB empirical protein-protein interaction network generated from differentially expressed proteins. Green molecules are down-regulated in ASD and red up-regulated compared to controls. Underlined proteins have reported associations with ASD in SFARI Gene database (<https://gene.sfari.org/>).

analysis of select peptides that are proxies for the 11 proteins of biological interest that are represented by 2 peptides, comparing both the directionality (ie, increased or decreased in ASD vs. controls) of the peptide noted in the unbiased, label-free analysis compared to the targeted analysis and whether the proteins were significantly differentially expressed by both label-free and SRM-based determinations. Of the 10 peptides selected for SRM validation analysis in BA19 that represented 5 proteins, 9 peptides validated in terms of directionality of expression and 3 of the 5 proteins validated in terms of significant differential expression for both types of analyses ($p < 0.05$). In CB samples, 12 peptides represented 6 proteins of interest; all 12 peptides validated in terms of directionality of expression and 5 of the 6 proteins validated in terms of significant differential expression for both of the two

Protein ID	Name	p-value	Associated Functions
Upstream Regulators for BA19			
MAPT	Microtubule-associated protein tau	3.39E-15	Microtubule assembly and stability
APP	Amyloid-beta A4 protein	7.45E-13	Neurite interactions, neuron adhesion, axonogenesis
PSEN1	Presenilin-1	3.36E-12	γ -secretase complex subunit
HTT	Huntington	5.17E-12	Microtubule-mediated transport or vesicle function
HDAC4	Histone deacetylase 4	1.12E-11	Histone deacetylation
MKNK1	MAP kinase-interacting serine/threonine-protein kinase 1	1.2E-09	Translation regulation, environmental stress response
BDNF	Brain-derived neurotrophic factor	3.18E-09	Peripheral and central nervous system development, synaptic function
ATN1	Atrophin-1	2.14E-07	Transcriptional corepressor
PPP3CA	Serine/threonine-protein phosphatase 2B catalytic subunit alpha isoform	2.86E-07	Calcineurin activity
PPARGC1A	Peroxisome proliferator-activated receptor gamma coactivator 1-alpha	6.29E-07	Transcriptional coactivator for PPARG, and TSH, metabolic responses to varying nutritional availability
Upstream Regulators for CB			
MAPT	Microtubule-associated protein tau	6.54E-09	Microtubule assembly and stability
APP	Amyloid-beta A4 protein	8.85E-08	Neurite interactions, neuron adhesion, axonogenesis
RTN4	Reticulon-4	1.10E-07	Neurite growth regulatory factor
PSEN1	Presenilin-1	2.35E-07	γ -secretase complex subunit
HTT	Huntington	2.22E-06	Microtubule-mediated transport or vesicle function
MTOR	Mechanistic Target Of Rapamycin	5.72E-06	Cellular metabolism, growth and survival
ROR1	Receptor Tyrosine Kinase Like Orphan Receptor 1	8.68E-05	Wnt signaling
ROR2	Receptor Tyrosine Kinase Like Orphan Receptor 2	8.68E-05	Osteoblast differentiation, bone formation, Wnt signaling
AGRN	Agrin	2.59E-04	Maintenance of the neuromuscular junction, post-synaptic differentiation
LPIN1	Phosphatidate phosphatase	2.59E-04	Fatty acid metabolism

Table 3. Top 10 upstream regulators for BA19 and CB that drive differential protein expression between ASD cases and controls based on downstream target expression changes. P-values are derived from one-tailed Fisher's exact test.

types of analyses ($p < 0.05$) (Supplementary Table 7). These data support the validity of the quantitative analyses produced by the unbiased, label-free MS analysis.

Discussion

Intense efforts have been mounted in recent years to understand the biological basis of autism. These investigations have implicated dysfunction of numerous biochemical pathways and biological networks in autism biology; several investigative approaches, including cytogenomic and exome analyses of affected individuals and transcriptomic and metabolic analyses of ASD tissues, have been especially influential in shaping our current understanding of molecular mechanisms in ASD pathogenesis. This study describes the first global proteomic analysis of autistic brain. Our aim was to evaluate which proteins are differentially expressed in two functionally and geographically distinct regions of autism vs. control brain and, using these data, determine pathways and protein-protein interaction networks that are relevant to autism biology.

Current data indicate that there are a large number of known causes of autism, nearly all of which are uncommon or rare^{9,10}. It is interesting that in this set of idiopathic autism brain samples 20 proteins in BA19 (13.7%) and 30 proteins in CB (15.7%) are differentially expressed in the examined ASD brains and are reported to have associations with either syndromic or non-syndromic autism in the SFARI Gene database¹¹. Within the protein-protein interaction networks, 11 proteins in the BA19 network and 6 proteins in the CB network are represented in the SFARI database including, for example, SHANK3 and CTNBN1 in BA19 and GRID2 and ATP1A3 in CB (Fig. 2).

Based on the differential protein expression data, there is enrichment for pathways in both BA19 and CB relating to glutamate signaling and glutathione-mediated detoxification/oxidative stress. The protein-protein interaction network analyses for both brain regions indicate the importance of synapse formation/function, glutamatergic signaling, cytoskeletal support of axon/neuron structure and of intracellular vesicle trafficking, and the transduction of intracellular calcium signals. Involvement of these pathways and processes has been implicated in ASD pathogenesis^{17–21}. For example, there are substantial data relating to glutamate and its receptors in a wide variety of neurobiological and behavioral findings associated with ASD^{32,33}. Similarly, there are abundant data

indicating evidence of oxidative stress in non-CNS tissues and brain region specific glutathione redox imbalance in ASD^{34–36}.

The canonical pathway and protein network analyses also revealed brain region specific processes. Pathways that appear differentially expressed in the two regions include Sertoli cell signaling in BA19 and pathways of fatty acid, organic acid and branched chain amino acid metabolism in CB. The association with Sertoli cell signaling raises two issues: whether this finding may relate to earlier work suggesting an association with ASD and inhibin³⁷ and the larger issue of a possible relationship with maleness, a strong but incompletely understood risk factor for autism³⁸. Associations between both rare monogenic and common idiopathic forms of autism and branched chain amino acid metabolism are well established, as are associations with specific disturbances of organic acid metabolism^{19,39–41}. Taken together, the differential protein expression data from human brain tissue provide direct support for shared autism-related pathobiology in the different brain regions studied here, as well as region-specific associations with autism.

In terms of networks, while both brain regions highlight synapse formation/function and glutamatergic signaling, the BA19-specific network highlights the function of DLG4 (PSD-95), a postsynaptic density protein that forms a multimeric scaffolding for the clustering of receptors, ion channels and associated signaling proteins, mediating NMDA receptor function and modulating synaptic plasticity via AMPA receptors⁴². The CB-specific network, while similar to BA19 in emphasizing cytoskeletal proteins that support axonal guidance/structure and intracellular vesicle trafficking, features in distinction to BA19 the up-regulation of tubulin and proteins associated with tubulin assembly and stability and a cerebellum-specific glutamate receptor, GRID2.

Striking data were noted regarding the upstream regulators, molecules that are predicted to drive the differential protein expression observed between ASD cases and controls through regulation of downstream targets. Upstream regulator analysis for both BA19 and CB brain regions showed highly significant enrichment for proteins that are central to many neurodegenerative disorders: MAPT (microtubule associated protein tau), APP (amyloid precursor protein), PSEN1 (presenilin-1) and HTT (huntington) (Table 3). Each of these proteins has multiple functions and it is possible that their actions as upstream regulators occur through different mechanisms^{43–46}. However, a more parsimonious explanation is that there are overlapping protein interaction networks or other shared properties of these proteins that when perturbed result in converging pathogenetically relevant pathways; there are extensive data in support of this concept for many neurodegenerative disorders^{47–49}. In terms of the upstream regulator data and a pathogenetic theory of autism, we hypothesize that perturbation of normal function of these proteins would impact part of a shared interactome, resulting in the broad CNS phenotype of autism. Two pathogenetic mechanisms are supported by the proteomic data of BA19 and CB and by the literature on shared interactomes and converging pathways in neurodegenerative disorders: impacts on synapse biology and on mitochondrial function^{48,49}. There also are abundant data in support of disturbances of synaptic function and of mitochondrial bioenergetics in the pathogenesis of autism^{15–19}. The data presented here suggest that the action(s) of several proteins that heretofore have had a thematic commonality of neurodegeneration also drive abnormal synaptic biology and/or mitochondrial dysfunction that is central to autism neurobiology.

There are both important similarities and differences of our findings and those of other studies using post-mortem ASD brains. Major themes of prior transcriptomic analyses of post-mortem brains, analyzing different brain regions, have included differential regulation of synaptic formation/function, immune response, mitochondrial activity and microglial functions^{29–31,50,51}. Dysregulation in pathways governing cell number, cortical patterning and differentiation in young autism cases and dysregulation of signaling and repair pathways was noted in adult autism cases in another study⁵². Like these investigators, our proteomic analysis indicates dysregulation of synaptic biology and mitochondrial bioenergetics and provides limited support for immune response differences; our data do not affirm other themes noted above.

In addition, this work provides insights into whether post-mortem autistic brains involve attenuation of regional protein expression differences. Some prior ASD vs. control brain transcriptomic analyses noted that regional brain differentiation between frontal cortex and temporal cortex appeared to be attenuated in autism^{30,31}, although other transcriptomic analyses evaluating possible attenuation of regional differentiation between ASD and control cerebellar and occipital cortex did not corroborate that finding⁵³. The current proteomic data of cerebellar and occipital cortex show similar levels of differential expression between these two regions (Fig. 2; Supplementary Table 3), suggesting that the previously reported attenuation of regional brain gene expression in ASD may be anatomically specific. Consequently, this issue warrants further investigation, especially in terms of evaluating additional brain regions and the incorporation of proteomic methods.

There are significant preanalytic and analytic challenges in carrying out a proteomic investigation using human post-mortem brain tissue⁵⁴. Additional challenges are faced when one studies post-mortem ASD brains as the number of available brain samples are limited. In addition, analyses of ASD cases are potentially confounded by etiological heterogeneity, medical co-morbidities, and various treatments. Issues regarding ASD/control brain sample selection and our use of inclusion/exclusion criteria were described previously, including the rationale for use of idiopathic ASD brain samples to minimize confounding of data that can occur though inclusion of samples with diverse identifiable etiologies⁵⁵.

Yet another challenge in neuroproteomics concerns the cellular complexity of brain tissue²⁴. Although care was taken to obtain matched areas of brain for each case and paired control, it is possible that there could be differences in cellular composition between paired samples, either because of differences due to the different biological process affecting the ASD cases relative to controls or because of artifact or chance difference. To our knowledge, there have been no attempts to address this issue in proteomic analyses of ASD brain. We addressed this by evaluating the expression of cell type-specific markers between cases and controls for each major CNS cell type. We noted no significant differences in a pairwise analysis, although modest differences were noted for several markers using a regression model. Overall, the data suggest that there are no major cell population differences between ASD vs. control brain samples (Supplementary Fig. 1, Supplementary Table 8).

Conclusions

This first proteomic evaluation of ASD human brains provides direct evidence of dysregulation of numerous proteins in two brain regions, BA19 and posterior inferior cerebellum, compared to controls. Bioinformatic analyses of the differentially expressed proteins of each of these brain regions suggests dysregulation of multiple pathways and protein interaction networks, especially implicating glutamatergic signaling. Overall, the proteomic data provide support for shared dysregulated pathways and upstream regulators for two brain regions in human ASD brain, suggesting a common ASD pathophysiology that has distinctive regional expression.

Methods

Subjects and samples. Brain samples were procured from the Autism Tissue Program (ATP, www.atp-portal.org) through the Harvard Brain Tissue Resource Center (www.brainbank.mclean.org) and the National Institute for Child Health and Human Development Brain and Tissue Bank (www.btbank.org). The samples were used in accordance with the principles of the Declaration of Helsinki. Tissues were procured from a total of 9 ASD and 9 control subjects. Brain samples from both cases and controls are from the left posterior inferior cerebellar hemispheric cortex (CB) and BA19 (occipital) cortex. The samples from autism cases were obtained from validated cases of ASD who met detailed inclusion/exclusion criteria; controls were matched for gender, age, brain region and post-mortem interval⁵⁵ (Supplementary Table 1). All 9 ASD brains were evaluated for *FMR1* nucleotide repeat expansion mutation, by two chromosomal microarray analyses, and with RNA-Sequencing based transcriptomic analyses; 8 of the 9 ASD brains were subject to deep sequencing of 78 known ASD candidate genes^{29,55–57}. The result of this detailed characterization of the ASD cases was that 8 of 9 ASD cases did not have an identified pathogenic chromosomal lesion or monogenic disorder that might account for ASD; one case (ANO9714) has an etiologically relevant DNA sequence variant, *SCN2A* c.4543C > T, p.R1515*. Eight of the 9 ASD cases, therefore, meet criteria for idiopathic autism. Tissue for proteomic analyses was available from 9 ASD and 9 control subjects for CB and from 7 ASD and 6 control subjects for BA19.

Global label-free expression proteomic analysis. Samples were rinsed in 100 mM Tris pH 8.0, transferred to a Covaris cryobag and frozen in liquid nitrogen, pulverized, transferred to glass vials, and then lysed with 2% SDS and 0.1% protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO). Samples were then sonicated, followed by filter-aided sample preparation detergent clean-up⁵⁸. 10 mcg protein of each sample were digested with LysC for 1 hour and trypsin overnight at 37 °C. Reverse phase LC-MS/MS was performed as described⁵⁹ except that 600 ng of peptide digests were loaded on the HPLC column and the gradient of solvent B ranged from 2% to 40% over 210 min.

Raw LC-MS/MS data files for each sample were processed using Rosetta Elucidator (Rosetta Biosoftware, Seattle, WA) (Version 3.3.01 SP4 25). Automated differential quantification of peptides was performed as previously described^{60,61}. Briefly, LC-MS/MS raw data were imported, and for each MS spectrum profile of each LC-MS/MS run, chromatographic peaks and monoisotopic masses were extracted and aligned. Chromatographic peaks were first aligned by retention time and monoisotopic mass. Peak lists with the monoisotopic mass and corresponding MS/MS data were then generated for each sample and searched using Mascot (Matrix Science, London, UK; version 2.4.1). Resultant peptide identifications were imported into Elucidator and monoisotopic masses annotated with peptide identifications. The false discovery rate for protein identifications was calculated to be 2%. The MS/MS peak lists were searched by Mascot using the human UniProt database. Search settings were as follows: trypsin enzyme specificity; mass accuracy window for precursor ion, 25 ppm; mass accuracy window for fragment ions, 0.8 Da; variable modifications including carbamidomethylation of cysteines, one missed cleavage and oxidation of methionine.

Statistical methods. MS data were obtained for each region; these files contained intensity values, with rows corresponding to peptides and columns corresponding to the sample. Missing values were imputed using a weighted k-nearest neighbors algorithm⁶². Next, data were log₂-transformed to improve symmetry and normal distribution approximation. Data were visualized with principal component analysis. One outlier was noted (BA19_6994) and removed from subsequent analyses. These preprocessing steps were performed using InfernoRDN⁶³. As the principal component analysis showed clear separation of samples based on brain region, data from each brain region were separately used for subsequent statistical and bioinformatic analyses.

To increase statistical power, we treated individual peptides as observations of a given protein. Data were imported into the R statistical programming environment for subsequent analyses. Using the lme4 package⁶⁴, a linear mixed effects model was applied to each protein with the following form:

$$intensity = \beta_1 \cdot CONDITION + \beta_2 \cdot PMI + \beta_3 \cdot AGE + \beta_4 \cdot PEPTIDE + \epsilon$$

where intensity refers to the log₂-transformed intensity values for each peptide observed for that protein, condition a binomial categorical variable (ASD or Control), PMI and age are continuous variables, peptide is a multi-level random effects categorical variable with the number of levels dependent on the protein, and ϵ is independent normally-distributed residual error with mean 0 and standard deviation σ . PMI data were unavailable for 3 samples, so these were assumed to be missing at random and imputed using predictive mean modeling, to allow use of all available intensity data in our model. For determination of proteins that varied significantly with changing PMI, we excluded these samples when running the model. A likelihood ratio was constructed for each coefficient in the model to assess its contribution to the variance in the data, and p-values were obtained. An additional analysis comparing regional contributions to differential expression within region, a binomial categorical variable, was run on the ASD and control datasets. P-values were adjusted using the false discovery rate (FDR) method described by Benjamini and Hochberg (BH)⁶⁵. Multiple proteins were identified as cell type-specific markers. To assess potential contributions of cellular heterogeneity, these cell type-specific markers were assessed

for differential expression through both the aforementioned regression model and separately with paired t-tests comparing expression changes between the paired autism and control samples. The pairs were compared using averages of the peptide intensities to reflect the specific protein concentration in each sample.

Bioinformatics analyses. Differentially expressed proteins identified by at least two unique peptides underwent bioinformatic analysis. Network, pathway, and upstream regulator analyses were performed for each brain region with Ingenuity Pathway Analysis® (IPA, www.qiagen.com/ingenuity); proteins with unadjusted p-value < 0.05 were used for this purpose. Enrichment scores and p-values for canonical pathway and upstream regulator analyses were determined by a one-tailed Fisher's exact test using the Ingenuity Knowledge Base as a reference. The latter is a database that is curated by expert reviewers that includes sets of member molecules of canonical pathways and downstream targets of upstream regulators, as well as all known interactions, direct and indirect, between proteins (http://pages.ingenuity.com/rs/ingenuity/images/IPA_data_sheet.pdf). Enrichment scores of above 1.3 correspond to a p-value < 0.05. Canonical pathway p-values were adjusted for multiple comparisons using the Benjamini-Hochberg (BH) procedure. Networks of protein-protein interactions were curated based on data from CNS and immune cell lines and tissues. The IPA bioinformatics tool generated interaction networks of the differentially expressed proteins through an iterative algorithm focused on interconnectivity between network members based on the Ingenuity Knowledge Base.

Results from IPA pathway analyses were validated using the open-source bioinformatics tool Reactome (<https://reactome.org>). P-values for Reactome-generated pathways from over-representation analysis were adjusted for multiple comparisons using the Benjamini Hochberg procedure.

Targeted MS analysis with selective reaction monitoring. Validation of the global proteomic data was done through selected reaction monitoring MS to quantify the levels of several differentially expressed proteins noted in the global, label-free proteomic analysis. The targeted MS analysis was performed as described previously⁶⁶. One or two peptides from each of 11 proteins of biological interest was synthesized (Sigma-Aldrich, St. Louis, MO) and used as a labeled internal standard that was spiked into each brain homogenate sample and used to quantify its unlabeled peptide analogue and thereby compute the concentration of the protein containing that peptide. Analysis of the selected peptides between the targeted and untargeted analyses was done using one-tailed Wilcoxon signed-rank tests. A p-value of 0.1 was considered significant given the nature of using Wilcoxon signed-rank tests on a small number of samples. Metrics assessed included consistency in directionality of expression between the label-free global vs. targeted analyses as well as whether statistical significance of differential expression was maintained. The peptides that were quantified, their amino acid sequences, and the proteins that they belong to are indicated in Supplementary Table 7.

Data Availability

The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE⁶⁷ repository with the dataset identifier PXD012755 and 10.6019/PXD012755. The code used for analysis is available upon request to the corresponding author.

References

- Lai, M.-C., Lombardo, M. V. & Baron-Cohen, S. Autism. *Lancet* **383**, 896–910 (2014).
- Lord, C., Elsabbagh, M., Baird, G. & Veenstra-Vanderweede, J. Autism spectrum disorder. *Lancet* **392**, 508–520 (2018).
- Bauman, M. L. Medical comorbidities in autism: challenges to diagnosis and treatment. *Neurotherapeutics* **7**, 320–327 (2010).
- Muskens, J. B., Velders, F. P. & Staal, W. G. Medical comorbidities in children and adolescents with autism spectrum disorders and attention deficit hyperactivity disorders: a systematic review. *Eur Child Adolesc Psychiatry* **26**, 1093–1103 (2017).
- GBD 2017 disease and injury incidence and prevalence collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1789–1858 (2018).
- Buescher, A. V. S., Cidav, Z., Knapp, M. & Mandell, D. S. Costs of Autism Spectrum Disorders in the United Kingdom and the United States. *JAMA Pediatr* **168**, 721–728 (2014).
- Elsabbagh, M. *et al.* Global prevalence of autism and other pervasive developmental disorders. *Autism Res* **5**, 160–172 (2012).
- Lyall, K. *et al.* The changing epidemiology of autism spectrum disorders. *Annu Rev Public Health* **38**, 81–102 (2017).
- Betancur, C. Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res* **1380**, 42–77 (2011).
- Kleijer, K. T. E., Huguot, G., Tastet, J., Bourgeron, T. & Burbach, J. P. H. Anatomy and cell biology of Autism Spectrum Disorder: lessons from human genetics. *Translational Anatomy and Cell Biology of Autism Spectrum Disorder, Advances in Anatomy, Embryology and Cell Biology*, Springer, pp 1–25 (2017).
- Larsen, E. *et al.* A systematic variant annotation approach for ranking genes associated with autism spectrum disorders. *Molec Autism* **7**, 44 (2016).
- Bölte, S., Girdler, S. & Marschik, P. B. The contribution of environmental exposure to the etiology of autism spectrum disorder. *Cell Mol Life Sci* **76**, 1275–1297 (2019).
- Gialloreti, L. E. & Curatolo, P. Autism Spectrum Disorder: why do we know so little? *Front Neurol* **9**, 1–4 (2018).
- Geschwind, D. H. & Levitt, P. Autism spectrum disorders: developmental disconnection syndromes. *Curr Opin Neurobiol* **17**, 103–111 (2007).
- Bourgeron, T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nat Rev Neurosci* **16**, 551–563 (2015).
- Mullins, C., Fishell, G. & Tsien, R. W. Unifying views of Autism Spectrum Disorders: a consideration of autoregulatory feedback loops. *Neuron* **89**, 1131–1156 (2016).
- Sztainberg, Y. & Zoghbi, H. Y. Lessons learned from studying syndromic autism spectrum disorders. *Nat Neurosci* **19**, 1408–1417 (2016).
- de la Torre-Ubieta, L., Won, H., Stein, J. L. & Geschwind, D. H. Advancing the understanding of autism disease mechanisms through genetics. *Nat Med* **22**, 345–361 (2016).
- Hollis, F., Kanellopoulos, A. & Bagni, C. Mitochondrial dysfunction in Autism Spectrum Disorder: clinical features and perspectives. *Curr Opin Neurobiol* **45**, 178–187 (2017).
- Oron, O. & Elliott, E. Delineating the common biological pathways perturbed by ASD's genetic etiology: lessons from network-based studies. *Int J Molec Sci* **18**, 828 (2017).

21. Quesnel-Vallieres, M., Weatheritt, R. J., Cordes, S. P. & Bienkows, B. J. Autism spectrum disorder: insights into convergent mechanisms from transcriptomics. *Nat Rev Genet* **20**, 51–63 (2019).
22. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347–355 (2016).
23. Aebersold, R., Bensimon, A., Collins, B. C., Ludwig, C. & Sabido, E. Applications and developments in targeted proteomics: from SRM to DIA/SWATH. *Proteomics* **16**, 2065–2067 (2016).
24. Hosp, F. & Mann, M. A primer on concepts and applications of proteomics in neuroscience. *Neuron* **96**, 558–571 (2017).
25. Abraham, J., Szoko, N. & Natowicz, M. R. Proteomic investigations of Autism Spectrum Disorder: past findings, current challenges, and future prospects. *Adv Exp Med Biol* **1118**, 235–252 (2019).
26. Bruchhage, M. M. K., Bucci, M. P. & Becker, E. B. E. Cerebellar involvement in autism and ADHD. *Handb Clin Neurol* **155**, 61–72 (2018).
27. Hazlett, E. A. *et al.* Regional glucose metabolism within cortical Brodmann areas in healthy individuals and autistic patients. *Neuropsychobiology* **49**, 115–125 (2004).
28. Poulin-Lord, M.-R. *et al.* Increased topographical variability of task-related activation in perceptive and motor associative regions in adult autistics. *NeuroImage* **4**, 444–453 (2014).
29. Gupta, S. *et al.* Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. *Nat Commun* **5**, 5748 (2014).
30. Voineagu, I. *et al.* Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **474**, 380–384 (2011).
31. Parikshak, N. N. *et al.* Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature* **540**, 423–427 (2016).
32. Choudhury, P. R., Lahiri, S. & Fajama, U. Glutamate mediated signaling in the pathophysiology of autism spectrum disorders. *Pharm Biochem Behav* **100**, 841–849 (2012).
33. Moretto, E., Murru, L., Martano, G., Sassone, J. & Passafaro, M. Glutamatergic synapses in neurodevelopmental disorders. *Prog Neurophysiol Biol Psychiatr* **84**, 328–342 (2018).
34. Chauhan, A., Audhya, T. & Chauha, V. Brain region-specific glutathione redox imbalance in autism. *Neurochemical Res* **37**, 1681–1689 (2012).
35. Rose, S. *et al.* Evidence of oxidative damage and inflammation associated with low glutathione redox status in the autism brain. *Transl Psychiatry* **2**, e134 (2012).
36. Frye, R. E. & James, S. J. Metabolic pathology of autism in relation to redox metabolism. *Biomark Med* **8**, 321–330 (2014).
37. Pankhurst, M. W. & McLennan, I. S. Inhibin B and anti-Müllerian hormone/Müllerian inhibiting substance may contribute to the male bias in autism. *Transl Psychiatry* **2**, e148 (2012).
38. Baron-Cohen, S. *et al.* Why are autism spectrum conditions more prevalent in males? *PLoS Biol* **9**, e1001081 (2011).
39. Novarino, G. *et al.* Mutations in BCKD-kinase lead to a potentially treatable form of autism with epilepsy. *Science* **338**, 394–397 (2012).
40. Tarlungeanu, D. C. *et al.* Impaired amino acid transport at the blood brain barrier is a cause of autism spectrum disorder. *Cell* **167**, 1481–1494 (2016).
41. Smith, A. M. *et al.* Amino acid dysregulation metabolites: potential biomarkers for diagnosis and individualized treatment for subtypes of autism spectrum disorder. *Biol Psychiatry* **85**, 345–354 (2019).
42. Coley, A. A. & Gao, W.-J. PSD95: a synaptic protein implicated in schizophrenia or autism? *Prog Neuropsychopharmacol. Biol Psychiatry* **82**, 187–194 (2018).
43. Zhang, C.-C., Xing, A., Tan, M.-S., Tan, L. & Yu, J.-T. The role of MAPT in neurodegenerative diseases: genetics, mechanisms and therapy. *Mol Neurobiol* **53**, 4893–4904 (2016).
44. Vetrivel, K. S., Zhang, Y. W., Xu, H. & Thinakaran, G. Pathological and physiological functions of presenilins. *Mol Neurodegen* **1**, 4 (2006).
45. Müller, U. C., Deller, T. & Korte, M. Not just amyloid: physiological functions of the amyloid precursor protein family. *Nat Rev Neurosci* **18**, 281–298 (2017).
46. Saudou, F. & Humbert, S. The biology of huntingtin. *Neuron* **89**, 910–926 (2016).
47. Hosp, F. *et al.* Quantitative interaction proteomics of neurodegenerative disease proteins. *Cell Rep* **11**, 1134–1146 (2015).
48. Wang, J. K. T., Langfelder, P., Horvath, S. & Palazzolo, M. J. Exosomes and homeostatic synaptic plasticity are linked to each other and to Huntington's, Parkinson's, and other neurodegenerative diseases by database-enabled analyses of comprehensively curated datasets. *Front Neurosci* **11**, 149 (2017).
49. Gan, L., Cookson, M. R., Petrucelli, L. & LaSpada, A. R. Converging pathways in neurodegeneration, from genetics to mechanisms. *Nat Neurosci* **21**, 1300–1309 (2018).
50. Garbett, K. A. *et al.* Immune transcriptome alterations in the temporal cortex of subjects with autism. *Neurobiol Dis* **30**, 303–311 (2008).
51. Anitha, A. *et al.* Brain region-specific altered expression and association of mitochondria-related genes in autism. *Mol Autism* **3**, 12 (2012).
52. Chow, M. L. *et al.* Age-dependent brain gene expression and copy number anomalies in autism suggest distinct pathological processes at young versus mature ages. *PLoS Genet* **8**, e1002592 (2012).
53. Ginsberg, M. R., Rubin, R. A. & Natowicz, M. R. Patterning of regional gene expression in autism: new complexity. *Sci Rep* **3**, 1831 (2013).
54. Szoko, N., McShane, A. J. & Natowicz, M. R. Proteomic explorations of autism spectrum disorder. *Autism Res* **10**, 1460–1469 (2017).
55. Ginsberg, M. R., Rubin, R. A., Falcone, T., Ting, A. H. & Natowicz, M. R. Brain transcriptional and epigenetic associations with autism. *PLoS One* **7**, e44736 (2012).
56. Wintle, R. F. *et al.* A genotype resource for postmortem brain samples from the Autism Tissue Program. *Autism Res* **4**, 89–97 (2011).
57. D'Gama, A. M. *et al.* Targeted DNA sequencing from autism spectrum disorder brains implicates multiple genetic mechanisms. *Neuron* **88**, 910–917 (2015).
58. Wisniewski, J. R. Filter-aided sample preparation for proteome analysis. *Methods Mol Biol* **1841**, 3–10 (2018).
59. Tomechko, S. E. *et al.* Proteomic and bioinformatics profile of paired human alveolar macrophages and peripheral blood monocytes. *Proteomics* **15**, 3797–3805 (2015).
60. Neubert, H. *et al.* Label-free detection of differential protein expression by LC/MALDI mass spectrometry. *J Proteome Res* **7**, 2270–2279 (2008).
61. Schlatzer, D. M., Sugalski, J., Dazard, J. E., Chance, M. R. & Anthony D. D. A quantitative proteomic approach for detecting protein profiles of activated human myeloid dendritic cells. *J Immunol Meth* **375**, 39–45 (2012).
62. Troyanskaya, O. *et al.* Missing value estimation methods for DNA microarrays. *Bioinformatics* **17**, 520–525 (2001).
63. Polpitiya, A. D. *et al.* DANTE: a statistical tool for quantitative analysis of omics data. *Bioinformatics* **24**, 1556–1558 (2008).
64. Bates, D., Maechler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J Statistical Software* **67**, 1–48 (2015).
65. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Society Series B (Methodological)* **57**, 289–300 (1995).
66. Schlatzer, D. M. *et al.* Urinary protein profiles in a rat model for diabetic complications. *Mol Cell Proteomics* **8**, 2145–2158 (2009).
67. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucl Acids Res* **47**, D442–D450 (2019).

Acknowledgements

We are grateful to the families who provided specimens from their loved ones for research and for the efforts of the Autism Tissue Program, Harvard Brain Tissue Resource Center and the National Institute for Child Health and Human Development Brain and Tissue Bank. We thank Drs Matthew Anderson, Jill James and Danica Wiredja for helpful discussions of this work. We also thank the anonymous reviewers for their suggestions.

Author Contributions

J.R.A. and M.R.N. designed the study. D.S., K.L. and X.L. performed proteomic analyses. J.R.A., J.B., M.R.N., R.A.R. and N.S. analyzed the data. J.R.A. and M.R.N. wrote the manuscript. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-49533-y>.

Competing Interests: M.R.N. is on the Scientific Advisory Board and has equity in Stemina Biomarker Discovery. All other authors declare no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019