1 **Title: Genome-wide DNA methylation patterns in Autism Spectrum Disorder**
2 **and mitochondrial function.**

3

4 **Authors:**

5 Sofia Stathopoulos¹, PhD

6 Renaud Gaujoux², PhD

7 Colleen O’Ryan^{1*}, PhD

8 **Affiliations:**

9 ¹Department of Molecular and Cell Biology, University of Cape Town, Cape Town,
10 South Africa.

11 ²Cytoreason LTD.

12 *Correspondence to: colleen.oryan@uct.ac.za.

13

Abstract

15

16 Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterised by
 17 phenotypic heterogeneity and overlapping co-morbidities. The genetic architecture of
 18 ASD is complex, with 100's of risk genes cumulatively contributing to the aetiology of
 19 ASD. Epigenetic mechanisms, particularly DNA methylation, have been associated
 20 with ASD. The vast majority of ASD molecular research has focused on Northern
 21 European populations, with a paucity of data from Africa. This study examines
 22 genome-wide DNA methylation patterns in a novel cohort of South African children
 23 with ASD and matched, unrelated controls. We performed a whole-genome DNA
 24 methylation screen using the Illumina 450K Human Methylation Array. We identify
 25 differentially methylated loci associated with ASD across 898 genes ($p\text{-value} \leq 0.05$).
 26 Using a pathway analysis framework, we find nine enriched canonical pathways
 27 implicating 32 of the significant genes in our ASD cohort. These pathways converge
 28 on two crucial biological processes: mitochondrial metabolism and protein
 29 ubiquitination, both hallmarks of mitochondrial function. The involvement of
 30 mitochondrial function in ASD aetiology is in line with the recently reported
 31 transcriptomic dysregulation associated with the disorder. The differentially
 32 methylated genes in our cohort overlap with the gene co-expression modules
 33 identified in brain tissue from five major neurological disorders, including ASD. We
 34 find significant enrichment of three gene modules, two of which are classified as
 35 Mitochondrial and were significantly downregulated in ASD brains. Furthermore, we
 36 find significant overlap between differentially methylated and differentially expressed
 37 genes from our dataset with a RNA seq dataset from ASD brain tissue. This overlap
 38 is particularly significant across the Occipital brain region ($\text{padj} = 0.0002$) which has
 39 known association to ASD. Our differential methylation data recapitulate the
 40 expression differences of genes and co-expression module functions observed in
 41 ASD brain tissue which is consistent with a central role for DNA-methylation leading
 42 to mitochondrial dysfunction in the aetiology of ASD.

43

44

45

46 **Keywords**

47 DNA methylation, epigenome, ASD, autism, mitochondria, metabolism, protein
48 ubiquitination, gene co-expression modules.

49 **Background**

50

51 It is forecast that by 2050 (1) the African continent will be home to 40% of the world's
 52 children, of whom 10-20 million are predicted to have Autism Spectrum Disorder
 53 (ASD). This prediction is based on the reported ASD incidence estimated in
 54 northern-hemisphere countries (2). There is a paucity of ASD information with little or
 55 no information available pertaining to epidemiology, diagnosis, intervention and
 56 particularly molecular genetics research from African populations (3). Currently there
 57 are only two publications reporting the genetic component of ASD in Sub-Saharan
 58 Africa (4), which highlights the importance of characterising the molecular basis of
 59 ASD in an African population. This knowledge gap is hamstrung by the ability to
 60 complete ADS genetic research in Africa, including South Africa, which is limited by
 61 the absence of ASD biological databanks with associated ASD phenotypes for
 62 African populations. This is in stark contrast to the European and North American
 63 ASD consortiums available to researchers working on those populations (4). South
 64 Africa has a unique population demography, and molecular investigations of a well-
 65 characterized ASD cohort can make a meaningful contribution to understanding the
 66 aetiology of this complex disorder in an understudied population(3,4).

67

68 ASD is a neurodevelopmental disorder characterised by deficits in social
 69 communication, restricted interests and repetitive behaviours (5). The genetic
 70 architecture of ASD is heterogeneous, with only 10% of ASD attributable to known
 71 genetic causes and 49% of ASD liability being quantitatively associated with
 72 common genetic variation (6). Although ASD is highly heritable, molecular
 73 investigations have yet to identify a consistent association of ASD with a biological
 74 marker (7). This reflects the complexity of the ASD phenotype, which is varied and
 75 overlaps with other co-morbidities (8). Although ASD is traditionally considered a
 76 brain disorder, a recent paradigm shift has emerged which considers ASD as a
 77 systems-level disorder (7,9), because of its numerous co-morbidities and the
 78 involvement of pleiotropic pathways in its aetiology. In this systems-level view, ASD
 79 is underpinned by a combination of overlapping biological networks that are
 80 dysregulated and affect multiple systems. The idea of overlapping biological
 81 networks being affected in ASD is reflected in the transcriptome, where post-mortem
 82 brain tissue of ASD individuals reveal distinct patterns of gene expression across a

83 number of biological pathways (10–12). A recent study (9) using transcriptomic
84 profiling coupled with GWAS data from five major psychiatric disorders including
85 ASD, found shared pathways of transcriptional dysregulation, as well as disease-
86 specific signatures of gene expression. The authors evaluated the genetic
87 component of ASD using GWAS data in a subset of gene co-expression modules
88 and raised the question of the contribution of epigenetic factors regulating gene
89 expression, given that genetic variation accounted for a fraction of the transcriptional
90 dysregulation observed in psychiatric disorders. Of the disease-specific
91 transcriptional dysregulation signatures associated with the psychiatric disorders,
92 specific gene co-expression modules involved in neuronal and mitochondrial function
93 were downregulated in ASD, linking ATP energetics to synaptic transmission in
94 psychiatric disorders (9).

95

96 Mitochondrial homeostasis is a fundamental physiological process and its
97 dysfunction affects all tissues. Congenital mitochondrial diseases are often lethal or
98 have severe consequences (13,14). There are more than 1500 gene products are
99 involved in mitochondrial function (14) and expression changes of any of these
100 genes can result in a range of mitochondrial dysregulation phenotypes. A 50 fold
101 increase in mitochondrial disease in children with ASD (15), altered oxidative stress
102 response in ASD lymphoblastoid cell lines (Rose et al. (16) and higher DNA damage,
103 DNA hypomethylation and oxidised protein levels(17) have been reported. Given the
104 high energetic demand of the brain, mitochondrial dysfunction has also been linked
105 to other neurologic disorders (18). For example, multiple lines of evidence suggest
106 that mitochondrial dysfunction plays a pivotal role in Parkinson's Disease (PD) via
107 mitochondrial respiration defects, aberrant programmed cell-death and mitophagy
108 (mitochondrial-specific type of autophagy), abnormal mitochondrial fission, fusion
109 and biogenesis (19,20).

110

111 The nuclear genome can epigenetically modulate mitochondrial function via histone
112 modification, DNA methylation or metabolic signals from the mitochondria (21). It is
113 accepted that the epigenome represents the interface between genes and the
114 environment (22) with epigenetic mechanisms driving temporal and spatial gene
115 regulation during cell differentiation and neurodevelopment (7,23,24). DNA
116 methylation has increasingly been associated with complex behavioural phenotypes

(25,26). For example, an epigenome-wide ASD monozygotic twin study revealed multiple differentially methylated genes associated with the disorder (27) and DNA methylation has been correlated with neuroimaging in ADHD-discordant monozygotic twins (28).

The enrichment of genetic ASD variants partly accounts for transcriptomic regulation in brain, with the role of epigenetic and environmental factors not yet fully characterised (9). This study aims to assess the contribution of DNA methylation in the differential gene expression observed in ASD. We identify differentially methylated genes associated with ASD a cohort of South African children using a whole-genome approach, and analyse these genes within a pathway analysis framework. The differentially methylated genes associated with ASD converge in pleiotropic pathways that result in mitochondrial dysregulation which implicate mitochondrial metabolic function and programmed cell-death via protein ubiquitination. Our data recapitulate the expression differences of genes and co-expression module functions observed in ASD brain tissue, supporting a central role of epigenetically-driven mitochondrial dysfunction in the aetiology of ASD.

Results and Discussion

ASD study cohort and phenotyping. There is a large body of knowledge regarding genetic variation and disease associations from Northern Hemisphere populations, with African populations being understudied (29). This bias is also observed in ASD genetic research (4) which is exacerbated by the lack of a centralised ASD biobank for African populations. Therefore, the first step of our study was to build an ASD cohort of South African children with age- and gender-matched controls. All participants were phenotypically characterised using the autism assessment tool ADOS-2. Since autism is a spectrum disorder, its defining traits are quantitative and are present along a continuum which overlaps with typically developing individuals (30), thus all the control participants were also phenotypically assessed. We administered ADOS-2 assessments to 127 participants (72 case and 55 control), and found that 13.6% of the children in the case group, who had a prior independent diagnosis of ASD, did not meet ADOS-2 criteria for ASD. Furthermore, 18.2% of the children in the control group, were excluded because of the presence of defining autistic traits, as defined by the ADOS-2 criteria. This dropout rate highlights the importance of phenotyping controls, a procedure which is not common-practice in ASD studies, and allowed us to narrow the heterogeneity in our control cohort before proceeding to the molecular analyses.

Differentially methylated genes in ASD. To examine the epigenetic component of ASD, we selected 48 study participants from our South African cohort (32 case and 16 control) to investigate differential DNA methylation using the Illumina 450K Human Methylation beadchip assay. We applied stringent quality control filtering criteria to our data, retaining 408604 CpG probes for subsequent analyses. To ensure that we were working with homogeneous epithelial cells samples, we tested methylation signals for the potential contamination from blood cells, which was confirmed to be absent. We tested for the covariates ethnicity and age, to confirm that these were not contributing to the observed differential methylation, and corrected for potential batch effects. We identified differentially methylated CpG sites across 898 genes between case and control, of which 39 were highly significant ($p \leq 0.005$, Figure 1, Table S1). Differentially methylated genes were distributed

across all chromosomes, including the sex chromosomes (Figure 1). The presence of numerous differentially methylated genes on the Y chromosome is consistent with the gender bias observed in ASD (2). Of the 898 genes differentially methylated in our cohort, 34 have previous associations with ASD (from the human SFARI gene database, <https://gene.sfari.org/database/human-gene/>, Table S2) including two genes that are among the top differentially methylated genes in our cohort ($p \leq 0.005$): GLRA2, a ligand-gated channel mediating a chloride-dependent inhibitory neurotransmission (31), and AHI1, a gene involved in both cerebellar and cortical development in humans (32). In support of a role for epigenetic mechanisms in ASD aetiology, we also found two differentially methylated genes that are directly involved in DNA methylation and that had been previously associated to ASD: SETD5 and MTR. SETD5 is a methyltransferase within the 3p25 microdeletion; mutations of SETD5 have been identified in several studies of ASD probands, leading to the classification of this gene as a strong ASD candidate gene (33) (Table S2). MTR encodes the enzyme 5-methyltetrahydrofolate-homocysteine methyltransferase which catalyses the final step in methionine biosynthesis. MTR mRNA levels in postmortem human cortex were found to be significantly lower in autistic individuals than in controls, especially at younger ages, comparable to the age-range of our South African ASD cohort (34).

Canonical pathways implicated in ASD: Mitochondrial energetics.

The limitations of single-gene analysis are apparent when examining the complicated genetic interplay typical of complex neurological disorders (35). Indeed, a number of transcriptomic studies have proposed that common pleiotropic pathways underpin not only ASD, but also a number of other neuropsychiatric disorders (9,36). Therefore, we analysed our data using a framework of pathway analysis, rather than a single-gene approach, to investigate the aetiology of ASD in our cohort.

We analysed the 898 genes with differentially methylated probes associated with ASD, in a canonical pathway framework. We identified nine enriched canonical pathways in our dataset (Table 1), which included 32 genes. Eight of the nine enriched canonical pathways identified in our dataset were metabolic pathways, with the majority of these occurring in the mitochondria (Table 1). Using IPA, we examined the overlap of the significant genes in our dataset with all the canonical

pathways and functions associated with mitochondrial function and biogenesis. This revealed additional differentially methylated ASD genes in both the Mitochondrial Dysfunction canonical pathway and the Mitochondrial Biogenesis toxicity list (Table S3), which support the observed enrichment of mitochondrial functions in our dataset and a central role for mitochondrial homeostasis in ASD (37).

Table 1_ Enriched Canonical Pathways and Associated Genes

Ingenuity Canonical Pathways	-log(p-value)	Cellular location	Genes
Methylmalonyl Pathway	2.06	Mitochondria	MUT,PCCB
2-oxobutanoate Degradation I	1.85	Mitochondria	MUT,PCCB
Glutaryl-CoA Degradation	1.72	Mitochondria	L3HYPDH,HACD2,EHHADH
Triacylglycerol Biosynthesis	1.56	Endoplasmatic Reticulum	LPGAT1,LPCAT3,PLPP5,MOGAT1,AGPAT1
*Protein Ubiquitination Pathway	1.51	Cytoplasm	HSPA2,DNAJC13,PSMB5,DNAJC8,PSMC2,USP51,DNAJ C28,USP31,UBB,UBE2C,USP49,MDM2,PSMD10,SMURF 1,RBX1,DNAJC5G,PSMA5
Acyl Carrier Protein Metabolism	1.41	Mitochondria	AASDHPPT
Acetyl-CoA Biosynthesis III (from Citrate)	1.41	Cytoplasm	ACLY
Lipoate Salvage and Modification	1.41	Mitochondria	LIPT1
Superpathway of Methionine Degradation	1.33	Mitochondria	MTR,MUT,PCCB,PRMT1

*Non-metabolic pathways

When synthesizing these seemingly disparate enriched metabolic canonical pathways within a framework of integrated metabolism, we observed that six of these pathways were directly linked to ATP production in the mitochondria (Figure 2). The Lipoate Salvage and Modification pathway transfers a lipoate moiety via the action of Lipoyltransferase 1 gene, LIPT1. This results in a lipoylated protein, the lipoyl-

domain of two crucial enzymes of the Tricarboxylic acid cycle (TCA), pyruvate dehydrogenase (PDH) and alpha ketoglutarate dehydrogenase (α-KGDH). Similarly, the Glutaryl-CoA Degradation, 2-Oxobutanoate Degradation, Methylmalonyl pathway, Superpathway of Methionine Degradation Acyl Carrier Protein Metabolism, and Acetyl-CoA Biosynthesis III (from Citrate) all directly affect the substrates of the TCA, which culminates in the production of ATP.

In addition to the genes in the enriched canonical pathways, we found six differentially methylated genes associated with ASD that are part of the Mitochondrial Dysfunction canonical pathway (Figure 3: pathway, Table S3) that also affect ATP production and Reactive Oxygen Species (ROS). These include four genes from the NDUF family (NDUFA4, NDUFB2, NDUFB4, NDUFB6) which are part of complex I, one gene in complex III (UQCRC2) and one gene part of complex IV (COX7B) of the electron transport chain. When inspecting our top differentially methylated genes ($p < 0.005$), we find that the most differentially methylated gene in our dataset, Stomatin Like 2 (STOML2), codes for a mitochondrial protein that regulates mitochondrial fusion in response to ROS damage. Importantly, the differential expressions of STOML2 and UQCRC2 have been validated in ASD brain tissue and contribute to two gene co-expression modules enriched in mitochondrial Gene Ontology pathways, which are downregulated in ASD (9).

Differentially methylated genes are enriched with co-expression modules linked to mitochondrial activity and ASD traits. To verify whether the results of our epigenomic screen were supported by the gene expression differences observed in ASD brain tissue, we computed the enrichment of our results in conjunction with available transcriptomic data (9). This recent study identified 13 gene co-expression modules that were robustly shared across ASD and four other psychiatric disorders in a cohort of 700 subjects of which there were 50 ASD subjects (9). The modules were defined from a co-expression network meta-analysis and were subsequently characterized according to their respective enrichment in GO categories and biological pathways, as well as their direct association with each psychiatric trait in the original cohort. Furthermore, the core expression pattern of each module was represented by a set of 20 top hub genes, as prioritized by their correlation with their respective module's eigengene (9).

We assessed the concordance between the 13 modules and our results in three ways. First, we computed the overlap between the selected list of differentially methylated genes from our data (898 genes) and the top 20 hub genes of each module (Fisher exact test). Secondly, we computed the overlap between the differentially methylated genes from our data and all the genes in each module. Thirdly, we computed the enrichment of each module among the genes associated with the most differential methylation, using Gene Set Enrichment Analysis (GSEA). Amongst the 13 modules, modules CD1, CD5, CD7 and CD10 showed a significant overlap with our ASD differentially methylated genes in at least one of the analyses ($p\text{-value} \leq 0.05$, Figure 4). Of particular interest are modules CD5 and CD10, which were strongly associated with both mitochondrial activity and ASD traits, while module CD1 was associated with both ASD and neuron and synapse functions. All three modules (CD5, CD10 and CD1) were enriched in neuron-specific genes and were significantly down-regulated in ASD (9). The significant concordance between our epigenomic ASD signature and mitochondrial/neuronal enriched gene co-expression modules in ASD brains, is also supported by the association of mitochondrial gene-enriched modules with neuronal firing rate (37), linking ATP production to synaptic function in neuropsychiatric disorders.

Down-regulated gene expression in ASD occipital cortex is associated with differential methylation. Furthermore, we tested whether our differentially methylated ASD genes were associated with changes in gene expression in different brain regions. To do so, we used a published RNA-seq gene expression dataset (Synapse:syn4587609) that contains samples of post-mortem cortex brain of 24 ASD and 17 non-psychiatric subjects to obtain ASD gene expression signatures within four cortical regions (frontal, temporal, parietal and occipital lobes) (9). We used GSEA (38) to compute the enrichment of our list of differentially methylated genes (898 genes) in each brain region signature (ordered by moderated t-statistic). We found a significant, negative enrichment in the occipital brain region (Figure 5, Table S4) which suggests that the observed down-regulated gene expression in this brain region of ASD subjects may be linked to corresponding methylation differences. Interestingly, a recent study of brain imaging data from ASD and typically developing boys revealed a decrease of structural connectivity and resting-state brain activity in

the occipital cortex of ASD brains, which is thought to be associated with the impaired social communication traits typical of ASD (39).

Therefore, we propose that differential methylation of a combination of genes associated with ASD in our cohort (Table S1) results in the dysregulation of canonical pathways which culminate in mitochondrial dysfunction, leading to the many varied traits seen in ASD. Consistent with this hypothesis, the observed differential methylation of the genes in the enriched canonical pathways regulates the expression of their respective encoded enzymes (Figure 2), as is the significant overlap of our data with the transcriptomic data on ASD brain tissue. Furthermore, we propose that the flux of metabolites which ultimately affect the TCA cycle and ATP production, via the electron transport chain in the mitochondria, is perturbed in ASD. For example, the Lipoate Salvage and modification pathway enriched in our cohort, if dysregulated, would impact lipoic acid production. Lipoic acid is crucial for the TCA, with deletions of the gene impacting its biosynthesis or salvage resulting in early embryonic lethal phenotypes (40) or early onset fatal disease (41). Furthermore, among the differentially methylated genes in our ASD cohort that have a mitochondrial function is Glycine amidinotransferase (GATM), a gene encoding a mitochondrial enzyme involved in creatinine biosynthesis. Mutations of this gene lead to GATM deficiency, and were found to be associated with intellectual disability and behavioural disorders, including ASD, in a study of 27 patients (42).

Several studies have identified abnormal metabolic biomarkers in individuals with ASD (18). For example, a study by Filipek et al. (43) examined levels of serum biomarkers of oxidative phosphorylation in 100 children with ASD. The study revealed a combination of relatively low serum carnitine levels, slightly elevated lactate and significantly elevated alanine and ammonia levels, all suggestive of mild mitochondrial dysfunction in children with ASD (43). Furthermore, markers of abnormal ETC function have been identified in post-mortem brains of individuals with ASD, with decreased levels of complex I, complex III, complex IV and complex V proteins and impaired complex I and complex IV activity in temporal lobe of these brains (44). Noteworthy, most mitochondrial abnormalities were found in the brains

of children under 10 years old, a similar age range as our cohort, suggesting a vulnerability of the developing brain (44).

Our mitochondrial hypothesis of epigenetic ASD aetiology is also consistent with the recent paradigm shift of considering ASD a systems-level disorder, rather than exclusively a brain disorder (7,45). Indeed, mitochondrial dysfunction has been proposed to explain the diverse symptoms associated with ASD (15,18,46). A recent systematic review and meta-analysis (15) based on a total of 18 publications examining children diagnosed with both ASD and mitochondrial disease, identified a 50-fold increase of mitochondrial disease in children with ASD, compared to the general population (15). Importantly, only 22% of these ASD individuals had a genetic cause underpinning mitochondrial disease, suggesting secondary mitochondrial dysfunction in the majority of cases. Mitochondrial dysfunction affects all tissues, particularly those with high energy demands like the central nervous system which depends on oxidative phosphorylation for successful neuronal development. It is therefore not surprising that defects in mitochondrial functioning have been associated with a range neurological disorders (47,48) including bipolar disorder (49), Alzheimer's disease (50,51) and Parkinson's disease (19,20).

Canonical pathways implicated in ASD: Protein Ubiquitination and Mitophagy.

Only one pathway of the nine enriched canonical pathways identified in our dataset was non-metabolic: the Protein Ubiquitination Pathway, with 17 differentially methylated genes (Table 1). These included SMURF1 (SMAD-specific E3 ubiquitin protein ligase 1) as well as four subunits of the proteasome (PSMA5, PSMB5, PSMC2 and PSMD10), all of which are crucial in the protein ubiquitination pathway, which leads to the degradation of damaged mitochondria (52). Mitophagy, the selective degradation of damaged mitochondria by autophagy, is a response to mitochondrial damage. Mitochondrial damage is "sensed" via complex protein interactions that include Pink1, Parkin and numerous translocase proteins of the mitochondrial outer membrane (TOM) complex, culminating in mitophagy via the ubiquitin – proteasome system (52–54). In our ASD cohort, we found three differentially methylated genes that are part of the Pink interaction network and 32 differential methylated genes in the Parkin (PRKN) interaction network, including

STOML2, the top differentially methylated gene in our dataset (Table S1). Changes in TCA flux and oxidative phosphorylation, leading to the altered ATP and ROS production proposed to occur in ASD based on our results, could trigger molecular pathways that culminate in mitophagy which are similar to those observed in well-studied neurodegenerative diseases such as Parkinson's Disease (54) and Alzheimer's Disease (55).

Fifty additional differentially methylated genes in our dataset were associated with canonical pathways and networks of autophagy, apoptosis and other cellular functions central to cell death and survival (Table S5). It has been reported that a decrease in ATP leads directly to an increase in proteasome activity (56) with even mild oxidative stress increasing the activity of the ubiquitin – proteasome system (57). Supporting a role for ubiquitin – proteasome system in ASD, irrespective of the trigger for this system, is that five genes involved in ubiquitination (AHI1, CUL3, FBXO15, PSMD10 and UBA6) are also present in the SFARI database, and these genes are all differentially methylated in our dataset (Table S2). For example, CUL3 plays a role in polyubiquitination and the degradation of proteins, with rare mutations of this gene seen in individuals with ASD (58,59). Degradation of damaged mitochondria also occurs via caspase-dependent cell death (20), with caspase-3 proposed as a target to prevent the progression of Parkinson's Disease (60). Caspases are a family of protease enzymes playing essential roles in programmed cell death and Caspase-2 is differentially methylated in our dataset (Table S5) and is involved in stress-induced cell death pathways and is implicated in neurodegenerative disorders like Alzheimer's Disease, Huntington's disease and temporal lobe epilepsy (61). Furthermore, Caspase-2 is inactivated by CSNK2A1, our third top differentially methylated gene associated with ASD. In addition, we found differentially methylated genes in the Apoptosis signalling and autophagy canonical pathways, as well as AHI1, which is linked to cell death (62). AHI1 is one of our top differentially methylated genes (Tables S1 and S2) and it is associated with syndromic autism and Joubert syndrome (32); AHI1 is required for cerebellar and cortical development in humans. In addition, four more of our top differentially methylated genes (ASB15, EIF2B1, BLZF1 and EMP3) have functions that are part of the ubiquitin-proteasome system or are involved in autophagy (based on their

respective IPA gene cards, www.ingenuity.com). Thus, the multiple differentially methylated genes identified in our data, the literature on neurodegenerative disorders and that reported on the SFARI data base, support the implication of the ubiquitin – proteasome system in ASD aetiology, in line with the well-studied role of this system mammalian brain formation (63).

Conclusions

Our study is the first genetic study from Sub-Saharan Africa to examine a well-characterised ASD cohort using a whole-genome approach, and we identify novel differentially methylated genes associated with ASD. Our results highlight the need for ASD genetic research in under-studied populations (29). Our data supports a central role for DNA methylation in ASD aetiology which culminates in disrupted ATP homeostasis via mitochondrial dysfunction, suggesting a central role for epigenetic regulation in the recently reported transcriptomic ASD gene-expression signatures. Furthermore, our data suggests a role for aberrant protein degradation and mitophagy in ASD aetiology, consistent with what is observed in other well-described neurological conditions (20)(55). Given that there are a number of pre-existing therapeutic options already in use for mitochondrial dysfunction (18), our data identifies exciting new avenues for researching the treatment of ASD symptomology.

Materials and Methods:

Cohort design. There is no ASD national database containing biological material and phenotypic profiles of individuals in South Africa. We recruited a cohort of children with ASD which were age- and gender-matched to typically developing controls. Given that methylation levels at some genes are associated with age (64), our age range was restricted to pre-pubertal children (6-12 years old) (65). Due to the higher genetic load and co-morbidities present in girls with ASD (66,67), we focused on boys only, in order to reduce the genetic heterogeneity in our cohort. Furthermore, only children with a prior independent diagnosis of ASD and without known co-morbidities were recruited into our case group; these children all attended special-needs Autism schools in the Western Cape, South Africa. We used the

Autism Diagnostic Observation Schedule, ADOS-2, to characterise the ASD phenotype of each child. The level of ASD severity in our cohort ranges from Moderate to Severe according to the ADOS-2 scale. Our ASD cohort is therefore characterised by a range of phenotypes which are representative of the ASD spectrum. The control cohort consisted of genetically unrelated children with no diagnosis of ASD, ADHD, or any other developmental delays, attending 'mainstream' schools. In addition, the control cohort was screened for the complete absence of any of the traits that define ASD using ADOS-2. All ADOS assessments were done in English by a team of Research Reliable ADOS-2 administrators; recruited participants were instructed in English at their respective schools.

Study participants and tissue collection. We screened 127 children, 72 cases and 55 controls, using ADOS-2 over a four-year period (2014-2017). Prior to recruiting, University of Cape Town ethical approval (FSREC076-2014) and Western Cape Government approval (20141002-37506) were obtained. Informed parental consent was obtained for each participant after we explained our study via oral and written presentations. Study participants did not receive a compensation and could withdraw from the study at any stage. Given the complex demographic history of South Africa, our study focused on the two demographic groups most represented at the recruiting sites: children of European and Mixed-Race ancestry. To limit the presence of the confounding factor of cellular heterogeneity, we used DNA extracted from buccal cells because these reflect DNA methylation of brain more accurately than blood cells (68,69). Buccal cells were collected using Epicentre Catch-All™ collection swabs. DNA was extracted using Proteinase K/ Salting out, followed by EtOH precipitation (70). DNA was re-suspended in sterile distilled water and quantified using Qubit 2.0 Fluorometer.

Whole-epigenome methylation beadchip assay. DNA (500ng) from 48 children, 32 cases and 16 controls, was bisulphite converted using Zymo EZ DNA Methylation-Gold™ kit, as per manufacturer's instructions. This sample size is similar to other methylome studies that were recently published (27,71), and reflects the semi-functional nature of the assay. Bisulphite-converted DNA samples were sent to the Agricultural Research Council core facility in Onderstepoort, South Africa for the beadchip assay. Genome-wide methylation was assessed using the Illumina Infinium

HumanMethylation450 beadchip assay, which screens > 485 000 CpG sites across the entire genome at single-nucleotide resolution. The CpG target sites cover UTRs and the gene body of 99% of RefSeq genes, which include 96% of CpG islands and flanking regions. Arrays were scanned using Illumina HiScanSQ.

Methylation beadchip analysis pipeline. All analyses were performed in R/Bioconductor (R Development Core Team 2011, (72)). Raw beadchip data were pre-processed and normalized using functions from the minfi, missMethyl and waterRmelon packages (73–75). Quality control of raw data was completed using the function waterRmelon::pfilter to flag low quality samples (over 1% sites with detection p-values > 0.05). The raw data for unflagged samples were then used to assess blood cell composition differences (minfi::estimateCellCounts) and normalized using SWAN (missMethyl::SWAN) (76). We then applied a sequence of filters to remove probes that had a detection p-value > 0.01 in at least one sample, probes that contained known SNPs (minfi::dropLociWithSnps) and probes that have been reported to be cross-reactive (77). After these filtering steps, the normalized data comprised 35 samples and 408 604 probes. All subsequent analyses were performed on M-values computed from these normalized data (with offset 100). Differences in methylation probe levels were obtained using the RUV-2 method (78), using Illumina 450k array negative controls and 0.5 FDR threshold for the first RUV fit. Differentially methylated probes (FDR ≤ 0.05) were annotated using Bioconductor annotation package FDb.InfiniumMethylation.hg19 which assigned methylation probes to the gene with the nearest transcription start site. A gene was called differentially methylated if it was assigned with at least one differentially methylated probe. In order to control for gene probe coverage differences, gene-level p-values were obtained using 2000 permutations of the sample labels, which generated null distributions that retain gene-specific coverage. P-values were computed by counting the proportion of times a gene obtained a statistic (maximum across assigned probes) in the permuted data, as extreme as the one observed on the original data.

Pathway enrichment. In our case vs control analysis, we identified differentially methylated CpG sites across 898 genes. We used this gene list to identify the

enriched pathways that differ between case and control using Qiagen's Ingenuity Pathway Analysis, IPA (www.qiagen.com/ingenuity). All canonical pathways with a $-\log(p\text{-value}) > 1.3$ are considered significantly enriched and we identified nine enriched pathways in our dataset, comprising a total of 32 differentially methylated genes. Of the 898 differentially methylated genes in our dataset, 39 genes have highly significant associations to ASD ($p\text{-value} < 0.005$). We use IPA to identify the enriched molecular networks associated with the differentially methylated genes in our cohort and known diseases and functions. We overlay our dataset with candidate canonical pathways which are interrelated with the core enriched canonical pathways in ASD. We find a significant enrichment of pathways consistent with mitochondrial dysregulation, implicating mitochondrial metabolic function and programmed cell-death via protein ubiquitination

Enrichment analysis of gene co-expression modules and ASD cortical signatures. The co-expression modules and 20 top hub gene sets definitions were downloaded and extracted from Supplementary Table S2 from the recent paper reporting transcriptome profiling in brain tissue from individuals with ASD ($n=50$) and four other psychiatric disorders ($n=700$ individuals in total (9)). Gene symbols were standardized using the latest version of the Bioconductor Human annotation package (org.Hs.eg.db version 3.4.1). Overlap between the differentially methylated genes in our cohort and each module and module hub gene sets was assessed using a single-sided Fisher exact-test (alternative "greater"). Enrichment of each module in the differential methylation signature was computed using pre-ranked GSEA (Subramanian et al. 2005) via the fgsea package with 10000 permutations. Enrichment analysis in pan-cortical RNA-seq dataset was performed in a similar way, using the signatures defined in Supplementary Table S1 from Gandal et al. (2018) with standardized gene symbols and genes ordered by decreasing moderated t-test statistic in each region separately.

List of abbreviations

Alzheimer Disease, AD; Attention deficit hyperactivity disorder, ADHD; Autism Spectrum Disorder, ASD; Autism Diagnostic Observation Schedule-2, ADOS-2; Adenosine triphosphate, ATP; Endoplasmic Reticulum, ER; Electron transport chain, ETC; Gene Set Enrichment Analysis, GSEA; Genome Wide Association Study, GWAS; Ingenuity pathway analysis, IPA; Parkinson Disease, PD; Reactive Oxygen Species, ROS; Tricarboxylic acid cycle, TCA.

Figure legends

Figure 1 Differentially Methylated genes in a South African ASD cohort

Manhattan plot depicting genome distribution of differentially methylated genes in ASD. The dotted line denotes p -value = 0.05; highly significant genes ($p \leq 0.005$) are indicated in red.

Figure 2 Mitochondrial pathways implicated in ASD aetiology

Differentially methylated genes in our dataset interact in multiple enriched canonical pathways (Table 2). When examining the enriched pathways in our ASD dataset, we find metabolic pathways which converge on the TCA cycle resulting in aberrant ATP production via the Electron Transport Chain, ETC, and the accumulation of Reactive Oxygen Species, ROS. The differentially methylated genes have a putative association with mitochondrial dysfunction in ASD. Differentially methylated genes are coded in colour corresponding to their respective canonical pathways which are indicated in colour-coded boxes. In these pathways, only relevant starting, intermediate and end-molecules are shown, where arrows can depict multiple reaction steps. Genes shown in colour are differentially methylated in our dataset resulting in changes in enzyme flux in their respective pathways.

Figure 3 Differentially methylated genes in the Mitochondrial Dysfunction canonical pathway

Mitochondrial dysfunction canonical pathway (IPA; www.qiagen.com/ingenuity). We find eight genes that are differentially methylated within this canonical pathway in our South African ASD cohort: COX7B, GRX2, MAP2K4, NDUFA4, NDUFB2, NDUFB4, NDUFB6, UQCRC2. These genes are central to different steps of the Mitochondrial dysfunction pathway, which are coloured in light pink in this figure.

Figure 4 Differentially methylated genes are enriched with co-expression modules linked to mitochondrial activity and ASD traits

Heatmap of $-\log_{10}(\text{p-value})$ of the comparisons of co-expression module overlap ('Top 20' and 'Module') and GSEA analysis ('Methylation') with our differentially methylated ASD genes. Only modules with $\text{p-value} \leq 0.05$ are shown. The color indicates the strength of association, from weak (light blue) to strong (dark blue).

Figure 5 Down-regulated genes in the occipital region of ASD brain are enriched with differentially methylated genes

Normalized enrichment scores from GSEA in the ASD signature within each cortical region tested. Color indicates statistical significance from weak (light blue) to strong (dark blue).

Declarations

Ethics approval and consent to participate. University of Cape Town ethical approval (FSREC076-2014) and Western Cape Government approval (20141002-37506) were obtained to conduct all aspects of this study. Informed parental consent was obtained for all study participants; study participants did not receive a compensation and could withdraw from the study at any stage.

Availability of data and material. The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests. The authors declare that they have no competing interests.

Funding. This work was funded by the South African National Research Foundation and the South African Medical Research Council.

Authors' contributions. COR and SS conceptualised and designed the study, performed all experiments, analysed and interpreted the results and wrote the manuscript. RG implemented the bioinformatic pipeline for data analysis. All authors read and approved the final manuscript.

Acknowledgements. We thank the families and the staff of all schools that participated in our research.

References

1. UNICEF. Generation 2030 Africa. UNICEF Div Data Res Policy. 2014;e_173.
2. Christensen DL, EdS JB, Van Naarden Braun K, Bilder D, Charles J, Constantino JN, et al. Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. *MMWR Surveill Summ.* 2016;65(3):1–23.
3. Abubakar A, Ssewanyana D, Newton CR. A Systematic Review of Research on Autism Spectrum Disorders in Sub-Saharan Africa. *Behav Neurol.* 2016;2016:1–14.
4. Franz L, Chambers N, von Isenburg M, de Vries PJ. Autism spectrum disorder in sub-saharan africa: A comprehensive scoping review. *Autism Res [Internet].* 2017 Mar; Available from: <https://www.ncbi.nlm.nih.gov/pubmed/28266791>
5. Lord C, Bishop SL. Recent advances in autism research as reflected in DSM-5 criteria for autism spectrum disorder. *Annu Rev Clin Psychol.* 2015;11:53–70.
6. Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB, et al. Most genetic risk for autism resides with common variation. *Nat Genet.* 2014 Aug;46(8):881–5.
7. Hu VW. From genes to environment: using integrative genomics to build a “systems-level” understanding of autism spectrum disorders. *Child Dev.* 2013 Jan;84(1):89–103.
8. Mannion A, Leader G. An investigation of comorbid psychological disorders, sleep problems, gastrointestinal symptoms and epilepsy in children and adolescents with autism spectrum disorder: A two year follow-up. *Res Autism Spectr Disord.* 2016;22:20–33.
9. Gandal MJ, Haney JR, Parikshak NN, Leppa V, Ramaswami G, Hartl C, et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science.* 2018 Feb 9;359(6376):693.

- 606 10. Voineagu I. Gene expression studies in autism: moving from the genome to the
607 transcriptome and beyond. *Neurobiol Dis.* 2012 Jan;45(1):69–75.
- 608 11. Garbett K, Ebert PJ, Mitchell A, Lintas C, Manzi B, Mirnics K, et al. Immune transcriptome
609 alterations in the temporal cortex of subjects with autism. *Neurobiol Dis.* 2008 Jun;30(3):303–11.
- 610 12. Chow ML, Pramparo T, Winn ME, Barnes CC, Li H-R, Weiss L, et al. Age-Dependent Brain
611 Gene Expression and Copy Number Anomalies in Autism Suggest Distinct Pathological Processes at
612 Young Versus Mature Ages. *PLOS Genet.* 2012 Mar 22;8(3):e1002592.
- 613 13. Alston CL, Rocha MC, Lax NZ, Turnbull DM, Taylor RW. The genetics and pathology of
614 mitochondrial disease. *J Pathol.* 2017;241(2):236–50.
- 615 14. Area-Gomez E, Schon EA. Mitochondrial Genetics and Disease. *J Child Neurol.* 2014 Sep
616 1;29(9):1208–15.
- 617 15. Rossignol DA, Frye RE. Mitochondrial dysfunction in autism spectrum disorders: a systematic
618 review and meta-analysis. *Mol Psychiatry.* 2012 Mar;17(3):290–314.
- 619 16. Rose S, Bennuri SC, Wynne R, Melnyk S, James SJ, Frye RE. Mitochondrial and redox
620 abnormalities in autism lymphoblastoid cells: a sibling control study. *FASEB J.* 2017 Mar;31(3):904–9.
- 621 17. Melnyk S, Fuchs GJ, Schulz E, Lopez M, Kahler SG, Fussell JJ, et al. Metabolic Imbalance
622 Associated with Methylation Dysregulation and Oxidative Damage in Children with Autism. *J Autism
623 Dev Disord.* 2012 Mar;42(3):367–77.
- 624 18. Griffiths KK, Levy RJ. Evidence of Mitochondrial Dysfunction in Autism: Biochemical Links,
625 Genetic-Based Associations, and Non-Energy-Related Mechanisms. *Oxid Med Cell Longev.*
626 2017;2017:12.
- 627 19. Bose A, Beal MF. Mitochondrial dysfunction in Parkinson’s disease. *J Neurochem.*
628 2016;139:216–31.
- 629 20. Franco-Iborra S, Vila M, Perier C. The Parkinson Disease Mitochondrial Hypothesis: Where
630 Are We at? *The Neuroscientist.* 2016;22(3):266–77.
- 631 21. Matilainen O, Quirós PM, Auwerx J. Mitochondria and Epigenetics – Crosstalk in
632 Homeostasis and Stress. *Trends Cell Biol.* 2017 Jun 1;27(6):453–63.
- 633 22. Rakan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common
634 human diseases. *Nat Rev Genet.* 2011 Jul 12;12(8):529–41.
- 635 23. Vogel Ciernia A, LaSalle J. The landscape of DNA methylation amid a perfect storm of autism
636 aetiologies. *Nat Rev Neurosci.* 2016 Jul;17(7):411–23.
- 637 24. Zahir FR, Brown CJ. Epigenetic Impacts on Neurodevelopment: Pathophysiological
638 Mechanisms and Genetic Modes of Action. *Pediatr Res.* 2011;69(5):92R-100R.
- 639 25. Hass J, Walton E, Wright C, Beyer A, Scholz M, Turner J, et al. Associations between DNA
640 methylation and schizophrenia-related intermediate phenotypes - a gene set enrichment analysis.
641 *Prog Neuropsychopharmacol Biol Psychiatry.* 2015 Jun;59:31–9.
- 642 26. Feinberg JL, Bakulski KM, Jaffe AE, Tryggvadottir R, Brown SC, Goldman LR, et al. Paternal
643 sperm DNA methylation associated with early signs of autism risk in an autism-enriched cohort. *Int J
644 Epidemiol.* 2015 Aug;44(4):1199–210.

- 645 27. Wong CC, Meaburn EL, Ronald A, Price TS, Jeffries AR, Schalkwyk LC, et al. Methyloomic
646 analysis of monozygotic twins discordant for autism spectrum disorder and related behavioural
647 traits. *Mol Psychiatry*. 2014 Apr;19(4):495–503.
- 648 28. Chen YC, Sudre G, Sharp W, Donovan F, Chandrasekharappa SC, Hansen N, et al.
649 Neuroanatomic, epigenetic and genetic differences in monozygotic twins discordant for attention
650 deficit hyperactivity disorder. *Mol Psychiatry* [Internet]. 2017 Mar 21; Available from:
651 <https://www.ncbi.nlm.nih.gov/pubmed/28322272>
- 652 29. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature*. 2016 Oct
653 13;538(7624):161–4.
- 654 30. Lai MC, Lombardo MV, Chakrabarti B, Baron-Cohen S. Subgrouping the autism “spectrum”:
655 reflections on DSM-5. *PLoS Biol*. 2013;11(4):e1001544.
- 656 31. Racca C, Gardiol A, Antoine T. Dendritic and Postsynaptic Localizations of Glycine Receptor a
657 Subunit mRNAs. Vol. 17. 1997. 1691 p.
- 658 32. Ferland RJ, Eyaid W, Collura RV, Tully LD, Hill RS, Al-Nouri D, et al. Abnormal cerebellar
659 development and axonal decussation due to mutations in *AHI1* in Joubert syndrome. *Nat Genet*.
660 2004 Sep;36(9):1008–13.
- 661 33. Grozeva D, Carss K, Spasic-Boskovic O, Parker MJ, Archer H, Firth HV, et al. De novo loss-of-
662 function mutations in *SETD5*, encoding a methyltransferase in a 3p25 microdeletion syndrome
663 critical region, cause intellectual disability. *Am J Hum Genet*. 2014 Apr;94(4):618–24.
- 664 34. Muratore CR, Hodgson NW, Trivedi MS, Abdolmaleky HM, Persico AM, Lintas C, et al. Age-
665 dependent decrease and alternative splicing of methionine synthase mRNA in human cerebral cortex
666 and an accelerated decrease in autism. *PLoS One*. 2013;8(2):e56927.
- 667 35. Jin L, Zuo XY, Su WY, Zhao XL, Yuan MQ, Han LZ, et al. Pathway-based analysis tools for
668 complex diseases: a review. *Genomics Proteomics Bioinformatics*. 2014 Oct;12(5):210–20.
- 669 36. Gupta S, Ellis SE, Ashar FN, Moes A, Bader JS, Zhan J, et al. Transcriptome analysis reveals
670 dysregulation of innate immune response genes and neuronal activity-dependent genes in autism.
671 *Nat Commun*. 2014 Dec 10;5:5748.
- 672 37. Winden KD, Oldham MC, Mirnics K, Ebert PJ, Swan CH, Levitt P, et al. The organization of the
673 transcriptional network in specific neuronal classes. *Mol Syst Biol* [Internet]. 2009 Jan 1;5(1).
674 Available from: <http://msb.embopress.org/content/5/1/291.abstract>
- 675 38. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set
676 enrichment analysis: A knowledge-based approach for interpreting genome-wide expression
677 profiles. *Proc Natl Acad Sci*. 2005 Oct 25;102(43):15545.
- 678 39. Jung M, Tu Y, Lang CA, Ortiz A, Park J, Jorgenson K, et al. Decreased structural connectivity
679 and resting-state brain activity in the lateral occipital cortex is associated with social communication
680 deficits in boys with autism spectrum disorder. *NeuroImage* [Internet]. 2017 Sep 18; Available from:
681 <http://www.sciencedirect.com/science/article/pii/S1053811917307796>
- 682 40. Yi X, Maeda N. Endogenous production of lipoic acid is essential for mouse development.
683 *Mol Cell Biol*. 2005 Sep;25(18):8387–92.

- 684 41. Tort F, Ferrer-Cortes X, Ribes A. Differential diagnosis of lipoic acid synthesis defects. *J*
685 *Inherit Metab Dis*. 2016 Nov 1;39(6):781–93.
- 686 42. Mercimek-Mahmutoglu S, Stoeckler-Ipsiroglu S, Adami A, Appleton R, Araújo H, Duran M, et
687 al. GAMT deficiency: Features, treatment, and outcome in an inborn error of creatine synthesis. *Vol.*
688 *67*. 2006. 480 p.
- 689 43. Filipek PA, Juranek J, Nguyen MT, Cummings C, Gargus JJ. Relative Carnitine Deficiency in
690 Autism. *J Autism Dev Disord*. 2004 Dec 1;34(6):615–23.
- 691 44. Tang G, Gutierrez Rios P, Kuo S-H, Akman HO, Rosoklija G, Tanji K, et al. Mitochondrial
692 abnormalities in temporal lobe of autistic brain. *Neurobiol Dis*. 2013 Jun 1;54(Supplement C):349–
693 61.
- 694 45. Gandal MJ, Haney JR, Parikshak NN, Leppa V, Horvath S, Geschwind DH. Shared molecular
695 neuropathology across major psychiatric disorders parallels polygenic overlap. 2016;
- 696 46. Frye RE, Rossignol DA. Mitochondrial Dysfunction Can Connect the Diverse Medical
697 Symptoms Associated With Autism Spectrum Disorders. *Pediatr Res*. 2011;69(5):41R-47R.
- 698 47. Trushina E, McMurray CT. Oxidative stress and mitochondrial dysfunction in
699 neurodegenerative diseases. *Neuroscience*. 2007 Apr 14;145(4):1233–48.
- 700 48. Lin MT, Beal MF. Mitochondrial dysfunction and oxidative stress in neurodegenerative
701 diseases. *Nature*. 2006 online;443:787.
- 702 49. Scaini G, Rezin GT, Carvalho AF, Streck EL, Berk M, Quevedo J. Mitochondrial dysfunction in
703 bipolar disorder: Evidence, pathophysiology and translational implications. *Neurosci Biobehav Rev*.
704 2016 Sep;68:694–713.
- 705 50. Salminen A, Haapasalo A, Kauppinen A, Kaarniranta K, Soininen H, Hiltunen M. Impaired
706 mitochondrial energy metabolism in Alzheimer’s disease: Impact on pathogenesis via disturbed
707 epigenetic regulation of chromatin landscape. *Prog Neurobiol*. 2015 Aug 1;131(Supplement C):1–20.
- 708 51. Hroudová JS Namrata Fišar, Zdenjk. Mitochondrial Dysfunctions in Neurodegenerative
709 Diseases: Relevance to Alzheimer’s Disease. *BioMed Res Int*. 2014;2014:9.
- 710 52. Bragoszewski P, Turek M, Chacinska A. Control of mitochondrial biogenesis and function by
711 the ubiquitin-proteasome system. *Open Biol [Internet]*. 2017 Apr;7(4). Available from:
712 <https://www.ncbi.nlm.nih.gov/pubmed/28446709>
- 713 53. Bender A, Desplats P, Spencer B, Rockenstein E, Adame A, Elstner M, et al. TOM40 Mediates
714 Mitochondrial Dysfunction Induced by α -Synuclein Accumulation in Parkinson’s Disease. *PLOS ONE*.
715 2013;8(4):e62277.
- 716 54. Bingol B, Sheng M. Mechanisms of mitophagy: PINK1, Parkin, USP30 and beyond. *Free Radic*
717 *Biol Med*. 2016 Nov 1;100(Supplement C):210–22.
- 718 55. Kerr JS, Adriaanse BA, Greig NH, Mattson MP, Cader MZ, Bohr VA, et al. Mitophagy and
719 Alzheimer’s Disease: Cellular and Molecular Mechanisms. *Trends Neurosci*. 2017 Mar 1;40(3):151–
720 66.
- 721 56. Huang H, Zhang X, Li S, Liu N, Lian W, McDowell E, et al. Physiological levels of ATP
722 negatively regulate proteasome function. *Cell Res*. 2010 Dec;20(12):1372–85.

- 723 57. Aiken CT, Kaake RM, Wang X, Huang L. Oxidative stress-mediated regulation of proteasome
724 complexes. *Mol Cell Proteomics*. 2011 May;10(5):R110 006924.
- 725 58. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, et al. Synaptic,
726 transcriptional and chromatin genes disrupted in autism. *Nature*. 2014 Nov 13;515(7526):209–15.
- 727 59. Wang T, Guo H, Xiong B, Stessman HA, Wu H, Coe BP, et al. De novo genic mutations among
728 a Chinese autism spectrum disorder cohort. *Nat Commun*. 2016 Nov;7:13316.
- 729 60. Kim J-S, Ha J-Y, Yang S, Son JH. A Novel Non-Apoptotic Role of Procaspase-3 in the
730 Regulation of Mitochondrial Biogenesis Activators. *J Cell Biochem*. 2017;n/a-n/a.
- 731 61. Martin I. Caspase-2: Divide and conquer. *Sci Transl Med*. 2016 Nov 2;8(363):363ec176.
- 732 62. Liu X, Chen M, Lobo P, An J, Grace Cheng S, Moradian A, et al. Molecular and structural
733 characterization of the SH3 domain of AHI-1 in regulation of cellular resistance of BCR-ABL(+) chronic
734 myeloid leukemia cells to tyrosine kinase inhibitors. *Proteomics*. 2012 Jul;12(13):2094–106.
- 735 63. Valnegri P, Huang J, Yamada T, Yang Y, Mejia LA, Cho HY, et al. RNF8/UBC13 ubiquitin
736 signaling suppresses synapse formation in the mammalian brain. *Nat Commun*. 2017 Nov;8(1):1271.
- 737 64. Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, et al. Epigenome-wide scans identify
738 differentially methylated regions for age and age-related phenotypes in a healthy ageing population.
739 *PLoS Genet*. 2012;8(4):e1002629.
- 740 65. Lintas C, Sacco R, Persico AM. Differential methylation at the RELN gene promoter in
741 temporal cortex from autistic and typically developing post-puberal subjects. *J Neurodev Disord*.
742 2016;8:18.
- 743 66. Robinson EB, Lichtenstein P, Anckarsater H, Hapke F, Ronald A. Examining and interpreting
744 the female protective effect against autistic behavior. *Proc Natl Acad Sci U S A*. 2013 Mar
745 26;110(13):5258–62.
- 746 67. Levy D, Ronemus M, Yamrom B, Lee YH, Leotta A, Kendall J, et al. Rare de novo and
747 transmitted copy-number variation in autistic spectrum disorders. *Neuron*. 2011 Jun;70(5):886–97.
- 748 68. Berko ER, Suzuki M, Beren F, Lemetre C, Alaimo CM, Calder RB, et al. Mosaic epigenetic
749 dysregulation of ectodermal cells in autism spectrum disorder. *PLoS Genet*. 2014;10(5):e1004402.
- 750 69. Lowe R, Gemma C, Beyan H, Hawa MI, Bazeos A, Leslie RD, et al. Buccals are likely to be a
751 more informative surrogate tissue than blood for epigenome-wide association studies. *Epigenetics*.
752 2013 Apr;8(4):445–54.
- 753 70. Aljanabi SM, Martinez I. Universal and rapid salt-extraction of high quality genomic DNA for
754 PCR-based techniques. *Nucleic Acids Res*. 1997;25(22):4692–3.
- 755 71. Schroeder DI, Schmidt RJ, Crary-Dooley FK, Walker CK, Ozonoff S, Tancredi DJ, et al. Placental
756 methylome analysis from a prospective autism study. *Mol Autism*. 2016;7:51.
- 757 72. Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho BS, et al. Orchestrating
758 high-throughput genomic analysis with Bioconductor. *Nat Meth*. 2015 print;12(2):115–21.
- 759 73. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a
760 flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation
761 microarrays. *Bioinformatics*. 2014 May 15;30(10):1363–9.

762 74. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analysing methylation
763 data from Illumina's Human Methylation450 platform. *Bioinformatics*. 2015;

764 75. Pidsley R, Wong CC., Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to
765 preprocessing Illumina 450K methylation array data. *BMC Genomics*. 2013;14(293):1–10.

766 76. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset quantile within-array normalization for
767 Illumina Infinium HumanMethylation450 BeadChips. *Genome Biol* [Internet]. 2012;13. Available
768 from: <http://dx.doi.org/10.1186/gb-2012-13-6-r44>

769 77. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of
770 cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450
771 microarray. *Epigenetics*. 2013 Feb;8(2):203–9.

772 78. Gagnon-Bartsch JA, Speed TP. Using control genes to correct for unwanted variation in
773 microarray data. *Biostatistics*. 2012 Jul;13(3):539–52.

774

$-\log_{10}(\text{FDR})$