

1 Chromatin, transcriptional and immune dysregulation in children with
2 neurodevelopmental regression

3

4 Hiroya Nishida^{1#}, Velda X Han^{1,2,3#}, Brooke A Keating¹, Katherine G Zyner⁴, Brian
5 Gloss⁵, Nader Aryamanesh^{6,7}, Pinki Dey⁸, Lee L Marshall⁶, Erica Tsang¹, Sarah
6 Alshammery¹, Xianzhong Lau⁹, Ruwani Dissanayake¹⁰, Shekeeb Mohammad^{1,11},
7 Mark Graham¹², Shrujna Patel^{1,11^}, Russell C Dale^{1,11,13^*}

8

9 Joint first author[#]

10 Joint senior author[^]

11

12 ¹Kids Neuroscience Centre, The Children's Hospital at Westmead, Faculty of
13 Medicine and Health, University of Sydney, NSW, Australia

14 ²Khoo Teck Puat-National University Children's Medical Institute, National University
15 Health System, Singapore, Singapore

16 ³Department of Paediatrics, Yong Loo Lin School of Medicine, National University of
17 Singapore, Singapore

18 ⁴Computational Systems Biology Unit, Children's Medical Research Institute, Faculty
19 of Medicine and Health, The University of Sydney, Westmead, NSW, Australia

20 ⁵Westmead Research Hub, Westmead Institute for Medical Research, Westmead,
21 NSW, Australia

22 ⁶Bioinformatics Group, Children's Medical Research Institute, University of Sydney,
23 Westmead, NSW, Australia

24 ⁷Faculty of Medicine and Health, Sydney Medical School, University of Sydney,
25 Camperdown, NSW, Australia

26 ⁸School of Biotechnology and Biomolecular Sciences, University of New South
27 Wales, Sydney, NSW, Australia

28 ⁹Australian Genome Research Facility Ltd, Melbourne, VIC, Australia

29 ¹⁰Australian Genome Research Facility Ltd, Westmead, NSW, Australia

30 ¹¹The Children's Hospital at Westmead Clinical School, Faculty of Medicine and
31 Health, University of Sydney, Sydney, NSW, Australia

32 ¹²Biomedical Proteomics, Children's Medical Research Institute, The University of
33 Sydney

34 ¹³The University of Sydney, School of Medical Sciences and Discipline of Child and
35 Adolescent Health, Faculty of Medicine and Health, Sydney, NSW, Australia.

36

37 **Corresponding author:**

38 *Prof Russell Dale

39 Kids Neuroscience Centre, The Children's Hospital at Westmead, Faculty of
40 Medicine and Health, University of Sydney, Sydney, New South Wales, Australia

41 The Children's Hospital at Westmead Clinical School, Faculty of Medicine and
42 Health, University of Sydney, Sydney, New South Wales, Australia

43 The Brain and Mind Centre, The University of Sydney, Sydney, New South Wales,
44 Australia

45 Fax: +61 2 9845 3389

46 Email: russell.dale@sydney.edu.au

47

48

49

50

51

52 **Summary**

53 Neurodevelopmental including autistic regression in children is often associated with
54 inconclusive diagnostic investigations, leaving the underlying pathophysiology largely
55 unexplored and poorly understood. To address this, we conducted peripheral blood
56 multi-omics profiling in 15 children with neurodevelopmental regression (median age
57 9 (4-14) years, 60% males), and 15 healthy controls (median age 12 (7-16) years,
58 60% males). Our comprehensive analysis included bulk RNA sequencing (8
59 regression vs. 8 controls), single-cell RNA sequencing (two cohorts of 4 regression
60 vs. 4 controls), bulk proteomics (4 regression vs. 4 controls), and
61 phosphoproteomics (4 regression vs. 4 controls). Despite clinical heterogeneity, we
62 identified convergent pathophysiological processes, including dysregulated pathways
63 in epigenetic pathways (chromatin remodelling, histone modifications, transcription
64 factors) and immune responses, as well as downregulated ribosomal and
65 translational pathways. These disruptions in epigenetic, immune and translational
66 processes may arise from complex genetic and environmental interactions, offering
67 opportunities for therapeutic interventions that specifically target these pathways.

68

69 Key words (10): neurodevelopmental disorders, autistic spectrum disorder,
70 developmental regression, epigenetics, omics

71

72

73

74

75

76

77 **Introduction**

78 Neurodevelopmental disorders (NDDs) affect 1 in 6 children, with rising prevalence¹.

79 These disorders arise from disruptions in brain development during critical prenatal
80 and postnatal developmental windows, influenced by both genetic and environmental
81 factors². Key developmental processes, including neurogenesis, synaptogenesis,
82 myelination, and synaptic pruning by microglia, actively shape brain architecture
83 during this period³. While genetic factors play a critical role in these processes, the
84 absence of a single pathogenic gene variant in many affected individuals highlights
85 the contribution of common gene variants⁴. Genes involved in chromatin
86 modification, transcriptional control, synaptic organisation, and immune regulation
87 have been implicated in the pathogenesis of NDDs⁵⁻⁷. Additionally, environmental
88 triggers also significantly impact these developmental processes, leading to changes
89 in brain structure, cognitive functions, and behaviour². The interplay between genetic
90 and environmental factors is complex and proposed to operate through epigenetic
91 modifications and immune pathways².

92

93 Developmental trajectories in children with NDDs are classically characterized by
94 gradual and steady progress. However, some children experience regression,
95 defined by the loss of previously acquired skills. This regression can be abrupt,
96 triggered by specific events, or gradual in nature. Autistic regression, observed in up
97 to one third of individuals with autistic spectrum disorders (ASD), typically manifests
98 in the second year of life⁸⁻¹², although sometimes occurs at an older age, referred to
99 as “childhood disintegrative disorder”. Autistic regression involves losing previously
100 acquired skills in social interaction, language, cognition, or behavioural changes⁸⁻¹².
101 In addition, some children lose developmental skills, such as cognitive decline,

102 memory loss or executive dysfunction, without fulfilling the criteria for autism or a
103 definable childhood dementia syndrome. Despite being widely recognised in the
104 literature, the pathophysiology of autistic and neurodevelopmental regression
105 remains unclear in a significant proportion of patients. For example, the current
106 diagnostic yield of exome sequencing in cases of autistic regression is only 9%¹³.
107 Likewise, neuroimaging and cerebrospinal fluid testing is typically normal or
108 unhelpful. It is therefore believed that most cases are not attributable to a rare de
109 novo or inherited DNA variation, but rather a combination of common vulnerability
110 genes and environmental factors¹⁴. Unfortunately, treatment in neurodevelopmental
111 regression remains limited, primarily focussed on managing comorbid behaviours
112 with anti-psychotics, stimulants and anti-depressants, and providing developmental
113 support with speech, occupational, physical and psychological therapies.

114

115 We hypothesize that epigenetic and immune dysregulation in both the blood and
116 brain, occurring during critical developmental periods, plays key roles in
117 neurodevelopmental regression. In this study, we investigated children with
118 neurodevelopmental regression with unexplained aetiology using bulk RNA
119 sequencing (RNAseq), single-cell RNA sequencing (scRNAseq), proteomics, and
120 phosphoproteomics of peripheral immune cells and found evidence of chromatin,
121 transcriptional and immune dysregulation. These insights provide opportunities for
122 biomarkers and future therapeutics targeting epigenetic and immune pathways.

123

124 **Methods**

125 *Participant selection*

126 Children (<18 years of age) were recruited into this study if they met criteria for a
127 prior history of neurodevelopmental (autistic or cognitive) regression.

128 1. Regression was defined as one or more episodes of losing previously
129 established skills; the loss of spoken language (such as a child who was
130 regularly speaking phrases who stops speaking) and/or loss of social and
131 communication abilities (such as a child previously using gestures and making
132 regular eye contact who loses such social responsiveness¹⁵), or cognitive
133 decline/memory loss. In all included children, there was a clear history of loss
134 of previously acquired neurodevelopmental or cognitive skills, associated with
135 functional decline- corroboration of the parental observations from teachers
136 and other family members was sought in all patients.

137 2. Normal investigation using routine application of neuroimaging, cerebrospinal
138 fluid and genomic testing (as defined by Australian Medicare supported
139 genomic testing). The investigation was determined by the treating clinicians.

140 3. All children had ongoing symptoms and were in the 'chronic phase' of their
141 illness. No patient had an infectious disease in the preceding two weeks prior
142 to blood sampling.

143 4. The patients who had scRNAseq and proteomics/phosphoproteomics (n=8)
144 were all screened for rare pathogenic variants that cause autistic spectrum
145 disorder (trio exome n=3, and autism gene panel n=5), and were negative.

146

147 *Control selection*

148 The inclusion criterion for healthy controls was absence of neurodevelopmental
149 disorders, autoimmune diseases, severe allergic conditions, as well as absence of

150 infection two weeks before blood draw. Controls were sex- matched to patients, and
151 there was no significant age difference between patients and controls.

152

153 *Semi-structured clinical interview*

154 Clinicians recorded the child's demographics, family history, birth history,
155 comorbidities and medication history after a semi-structured clinical review for all
156 participants, using DSM-5 criteria¹⁶. Details around the regression episodes were
157 collated including development prior to regression, timings of regression, and
158 developmental domains affected, including loss of social communication, loss of
159 language/meaningful words, loss of imaginative play/restricted play, and
160 restricted/repetitive patterns of behavioural interest. Videos were used to corroborate
161 parental observations of loss of developmental skills.

162

163 *Sample collection*

164 After written consent, venous blood samples were collected for testing in PAXgene™
165 blood RNA tubes (Qiagen, Hilden, Germany) and Acid Citric Dextrose tubes
166 (BD). Peripheral blood mononuclear cells (PBMCs) were isolated from the whole
167 blood within 12 hours of sample collection using a Ficoll density medium (Cytiva) in
168 SepMate™ tubes (Stemcell Technologies). PBMCs were cryopreserved in 10%
169 dimethyl sulfoxide (Amresco, WN182-10ML) in foetal bovine serum (FBS, Thermo
170 Fisher Scientific) and stored at -80°C in controlled cooling chambers for 24 hours
171 before being transferred to liquid nitrogen for storage until single cell RNA-
172 sequencing.

173

174 *Bulk RNA sequencing*

175 Bulk RNA sequencing was performed in 8 children with neurodevelopmental
176 regression compared to 8 healthy controls (Figure 1, refer to Table S1). This
177 workflow includes RNA extraction from PAXgene™ blood RNA tubes, depletion of
178 ribosomal RNA via hybrid capture (Illumina Ribo-Zero), and Illumina TruSeq
179 Stranded Total RNA Library Preparation (input 200-1,000 ng of Total RNA). The
180 stranded RNA samples were sequenced on the Illumina NovaSeq 6000 next
181 generation sequencing platform (2 x 150 base pairs) for a depth of 50 million paired
182 end reads. The cleaned sequence reads were aligned against the *Homo sapiens*
183 genome (Build version hg38), and the STAR aligner (v2.5.3a) was used to map
184 unique reads to the genomic sequences¹⁷.

185

186 HIVE™ *Single-cell RNA sequencing*

187 Single-cell RNA sequencing was performed in 8 children with neurodevelopmental
188 regression compared to 8 healthy controls (2 cohorts of 4 patients vs 4 healthy
189 controls). Red blood cells (RBC) were depleted from whole blood by
190 immunomagnetic negative selection, using the EasySep™ RBC depletion reagent
191 protocol (catalog #18000), preserving all leukocyte populations, including
192 granulocytes. Two cohorts of 8 HIVE devices (4 patients and 4 controls for each
193 cohort) were each loaded with 30,000 cells in 1 mL of DPBS + 1% FBS, followed by
194 3mL of cell media (DPBS + 1% FBS), according to the manufacturer's protocol. The
195 leukocytes were loaded into the HIVE collector within an hour of blood sampling from
196 all individuals to minimise neutrophil activation. Single-cells settled into picowells of
197 the HIVE containing 3' transcript-capture beads. HIVEs were then placed onto a spin
198 plate and spun at 30 x G for 3 minutes at room temperature. Media was removed
199 from the HIVE, and 2 mL sample wash solution was added to each, following the

200 manufacturer's instructions for the HIVE sample capture kit (Honeycomb
201 Biotechnologies). The sample wash solution was removed, and the cell-loaded
202 HIVEs were frozen at -80°C after adding 1 mL of the cell preservation solution,
203 before being transferred to the Australian Genome Research Facility (AGRF Ltd) and
204 processed through to single-cell NGS libraries.

205

206 HIVE devices were processed according to the manufacturer's instructions. Briefly,
207 cell-loaded HIVE devices were sealed with a semi-permeable membrane, enabling
208 the use of the strong lysis solution and the addition of a hybridization solution. Beads
209 with captured transcripts were extracted from the HIVE device by centrifugation.
210 Subsequent HIVE library preparation steps were performed in a 96-well plate format.

211 The size distribution of the final amplified libraries was determined using the
212 TapeStation and the final concentrations of the pooled libraries were accurately
213 quantified using the qPCR assay. HIVE scRNAseq libraries were paired-end
214 sequenced (2 × 150 bp) with custom primers (provided in the HIVE single-cell RNA
215 seq processing kit) on an Illumina NovaSeq 6000 platform.

216

217 *Proteome analysis and preparation from PBMCs*

218 Proteomics and phosphoproteomics were performed in 4 children with
219 neurodevelopmental regression and 4 controls (second cohort who had scRNAseq).
220 Following isolation, PBMCs were lysed in 200 µL lysis buffer containing 0.8% Triton
221 X-100, PhosSTOP, 50 mM HEPES (pH 7.4, adjusted with NaOH), and 2 mM
222 phenylmethylsulfonyl fluoride (PMSF, in ethanol). The lysates were heated at 85°C
223 for 10 minutes in a heating block and stored at -80°C. The samples were thawed,
224 benzonase treated, reduced, alkylated, precipitated and digested as described

225 previously¹⁸. Approximately 90 µg of peptide from each sample was labelled with a
226 TMTpro reagent and combined. A portion of the combined sample was fractionated by
227 hydrophilic interaction liquid chromatography on a Vanquish Neo HPLC system
228 (Thermo Fisher Scientific) with a 250 mm long and 1 mm inside diameter TSKgel
229 Amide-80 column (Tosoh Biosciences). The HILIC gradient used 90% acetonitrile,
230 0.1% TFA (Buffer A) and a solution of 0.1% TFA (Buffer B). The flow rate was 50
231 µL/min. The gradient was from 100% Buffer A to 60% Buffer A in 35 min. Fractions
232 were collected into a 96-well plate using an FC204 fraction collector (Gilson) at 1 min
233 intervals, monitored by absorbance of UV at 214 nm. UV signal was used to combine
234 select fractions into similar amounts of peptide. Fractions were dried and
235 reconstituted in 0.1% formic acid. The LC-MS/MS was performed as described
236 previously¹⁸ using a Dionex UltiMate 3000 RSLC nano system and Q Exactive Plus
237 quadrupole-orbitrap mass spectrometer (Thermo Fisher Scientific). We obtained
238 5,961 proteins suitable for analysis.

239

240 *Phosphoproteomics analysis*

241 Phosphopeptides were enriched using a previously described method¹⁹. Hydrophilic
242 interaction liquid chromatography was used to fractionate the phosphopeptides as
243 for the proteomics. LC-MS/MS was performed as for the proteomics except that
244 gradient was from 1% buffer B to 5% buffer B in 1 min, to 25% buffer B in 74 min, to
245 35% buffer B in 8 min, to 99% buffer B in 1 min, held at 99% buffer B for 2 min, to
246 99% buffer A in 1 min and held for 8 min as the flow rate increased to 275 nL/min.
247 The MS/ settings were the same as for proteomics except that maximum ion time
248 was 115 ms and the normalized collision energy was 34. We obtained 7,550
249 phosphopeptides suitable for analysis.

250

251

252 *Proteomics and phosphoproteomics database searching*

253 The raw LC-MS/MS data was processed with MaxQuant v1.6.7.0 using the following
254 settings: variable modifications were oxidation (M), acetyl (protein N-terminus),
255 deamidation (NQ) and phospho (STY); carbamidomethyl (C) was a fixed
256 modification; digestion was set to trypsin/P with a maximum of 3 missed cleavages;
257 the TMTpro correction factors were entered for lot XA340093; minimum reporter
258 peptide ion fraction was 0.6; the *Homo sapiens* reference proteome with canonical
259 and isoform sequences downloaded Dec 6 2023 with 80,685 entries and 20,596
260 genes; the inbuilt contaminants fasta file was also used; minimum peptide length
261 was 6 and maximum peptide mass was 6,000 Da; second peptides search and
262 dependent peptides search were enabled; all modified peptides and counterpart non-
263 modified peptides were excluded from protein quantification. All other settings were
264 default.

265

266 *Bioinformatic analysis*

267 *Analysis of omics datasets*

268 Omics data were analyzed in the R statistical environment²⁰ with tidyverse. For bulk
269 RNA sequencing and proteomics, filtering and normalization steps were first
270 performed. Subsequently, normalization with removal of unwanted variation, via the
271 remove unwanted variation (*RUV*) R package was performed²¹. *RUV* is commonly
272 used in human samples to better discern biologically relevant variations associated
273 with the disease of interest, particularly in the presence of significant inter-sample
274 biological diversity²². This method relies on having a set of endogenous negative

275 control genes known *a priori* not to be differentially expressed with respect to the
276 biological factor of interest^{21,22}. For this study, a set of 500 empirical negative control
277 proteins with little or no change in RNA expression across samples was identified
278 from an initial ANOVA test. In the bulk RNA sequencing (k=10), proteomics (k=4),
279 and phosphoproteomics (k=4) (factors of unwanted variation) were used respectively
280 to remove genes that had minimal differential expression, compared to negative
281 control genes. For linear modelling, the limma R package was used and the p-values
282 were calculated using the empirical Bayes method 'eBayes' function²³. The false
283 discovery rate correction was applied to the p-values by calculating the adjusted p-
284 values. Significant differentially expressed genes/proteins were defined as those with
285 adjusted p-values/false discovery rate (FDR) < 0.05.

286

287 For single cell transcriptomics, the data was analysed using the Seurat package²⁴.
288 Cells with a high mitochondrial transcript ratio (>0.15) were excluded. Experiments
289 were integrated using the *FindIntegrationAnchors* function in Seurat then immune
290 cell types were assigned using *scPred*²⁵. Merged data were then split by cell type
291 and separately normalized, scaled, integrated between patient using harmony²⁶, then
292 UMAP (uniform manifold approximation and *projection*) projections were made using
293 the first 30 dimensions. Differentially expressed genes with significant FDR values
294 <0.05 were identified using *FindMarkers*.

295

296 Pathway enrichment analysis

297 *Gene set enrichment analysis (GSEA) for bulk RNAseq, scRNAseq and proteomics:*
298 The genes and proteins were ranked based on their sign(logFC) x log10Pvalue
299 scores^{27,28}. Enriched gene sets were identified based on a running sum statistic

300 (normalized enrichment score (NES)) and statistical significance, based on the false
301 discovery rate (FDR). Significant GSEA GO pathways (FDR <0.05) were further
302 simplified using the *simplify* function in clusterProfiler²⁹.

303

304 *Over representation enrichment (ORA) analysis for phosphoproteomics:* We took an
305 approach that deliberately reduced the significant pathways identified by (a) limiting
306 the background to only proteins we could detect by mass spectrometry and (b) using
307 the ranked list method within gProfiler, which weights enrichment toward highly
308 ranked genes. We separated up- and down-regulated phosphorylation, assigned the
309 maximum positive and negative quantitative values for phospho-regulation to each
310 protein, and ranked the proteins using both the quantitative value and significance of
311 maximal change.

312

313 Bar and dot plots of GSEA results were plotted using *ggplot2* package, and
314 heatmaps of GSEA results were made using the *pheatmap* package. Connectivity
315 network (CNET) plots were created using *enrichplot* package where the enriched
316 pathways are represented by their respective colors, and corresponding genes'
317 adjusted p value. To improve understanding in some of the differential pathways, we
318 performed sub clustering of the pathways. Genes associated with the pathway were
319 extracted from the GO pathway results and subjected to further GO Enrichment
320 Analysis using Reactome/GO molecular function. To refine the results, pathways with
321 a reference gene count exceeding 500 were excluded to avoid overly broad terms,
322 and the remaining pathways were ranked by FDR significance. Starting from the top-
323 ranked pathways, we sequentially included pathways in the CNET plot, excluding
324 those where more than half of the genes overlapped with higher-ranked pathways.

325

326 **Ethics approval**

327 Ethical approval was granted by the Sydney Children's Hospitals Network Human
328 Research Ethics Committee (HREC/18/SCHN/227, 2021/ETH00356). Informed
329 written consent was obtained from all participants.

330

331 **Results**

332 *Clinical data*

333 A total of 15 children with a clear history of neurodevelopmental regression (median
334 age 9 (4-14) years at time of sampling, 60% males, with an average time from first
335 regression to sampling of 4.7 years (median 5 (0.5-11) years) and a total of 15
336 controls (median age 12 (7-16) years, 60% males) were recruited for this study (refer
337 to Table S1). Family history of neurodevelopmental, neuropsychiatric or autoimmune
338 disorders are summarized in Table S2. Parental history of neurodevelopmental or
339 neuropsychiatric disorders was present in 12 of 15 patients (maternal n=10, paternal
340 n=4, including both parental n=2), and an autoimmune condition was present in 8
341 (maternal n=8, paternal n=2, including both parental n=2). Pregnancy related
342 complications (n=3), infections (n=3) and significant stressors during pregnancy
343 (n=3) were also present (Table S2).

344

345 For the 15 patients in this study, 7 had some developmental issues prior to the onset
346 of regression, this included ADHD (n=4), ASD (n=2), and language delay (n=1). The
347 first regression occurred at a median age of 4 (1.5 - 12) years, and 10 children were
348 reported to have triggers associated with the regression event, including infections
349 (n=9, impetigo, Staphylococcal scalded skin syndrome, sore throat, and urinary tract

350 infection), and vaccination (n=1). During the autistic and cognitive regression,
351 symptoms included new onset emotional lability/depression (including anxiety, rage,
352 agitation, irritability) (n=15), new onset or exacerbation of ritualistic and compulsive
353 behaviour (n=15), academic and/or cognitive decline (n=13), loss of language or
354 meaningful words (n=12), behavioural developmental regression or loss in self-help
355 skills (n=12), loss of social skills (n=11), sensory or motor abnormalities (n=11), tics
356 (n=10), loss of imaginative play (n=9), sleep disturbance (n=8) and restricted food
357 intake (n=7). The first regression lasted an average of 11.7 weeks (median 7 weeks,
358 range 2-52 weeks), excluding the two individuals whose symptoms persisted after
359 the initial regression. (Table S2).

360

361 The patients had severe neuropsychiatric and behavioural symptoms requiring visits
362 to a hospital emergency department (n=8), and long school absence (>3 months)
363 (n=4). The majority (n=13) reported recurrent regressions (occurring 1-6 times a
364 year), typically triggered by infections, and these deteriorations lasted between 2 and
365 52 weeks, until subsequent restoration or partial restoration of neurodevelopmental
366 skills (Table S2).

367

368 *Blood bulk RNA sequencing*

369 Exploratory bulk RNAseq was performed in 8 children with neurodevelopmental
370 regression compared to 8 healthy controls (Figure 1, Table S1). There was no
371 significant difference in blood cell type composition between patient and control
372 samples (Table S1). Unsupervised principal component analysis (Figure 2A) showed
373 clear discrimination between groups. Top 10 upregulated GSEA GO (FDR <0.05)
374 pathways (Figure 2B, in red) involved immune function (immunoglobulin complex,

375 myeloid cell differentiation/activation, leukocyte migration), epigenetic mechanisms
376 (demethylation), and signaling transduction (secretory granule lumen, endocytosis,
377 tyrosine kinase signaling).

378

379 Genes which enriched the upregulated myeloid cell differentiation pathway (Figure
380 2C) were further subclustered via Reactome and included four themes including TLR
381 2 cascade, signaling by interleukins, RNA polymerase II transcription, and signaling
382 by Rho GTPases. Genes which enriched the upregulated demethylation pathway
383 (Figure 2D) were also further subclustered via Reactome and included three themes
384 including HDMs (histone demethylases including *KDM1B*, *KDM2A*, *KDM3A*, *KDM3B*,
385 *KDM4A*, *KDM4B*, *KDM5A*, *KDM5B*, *KDM5C*, *KDM6A*, *KDM6B*, *KDM7A*), DNA
386 demethylation enzymes (TET (ten-eleven translocation), TDG (thymine DNA
387 glycosylase)), and DNA damage reversal.

388

389 Top 10 GSEA GO (FDR <0.05) downregulated pathways (Figure 2B, in blue) include
390 ribosomal biogenesis, translation, and mitochondrial expression pathways. Genes
391 which enrich the top 2 downregulated ribosomal pathways include multiple ribosomal
392 genes associated with small subunit (40S) or large subunit (60S) of the eukaryotic
393 and mitochondrial ribosome: *RPS*, *RPL*, and eukaryotic translation initiation factor
394 genes: *EIF* (Figure S1).

395

396 *Blood single-cell RNA sequencing, cohort 1 and cohort 2*

397 We performed two independent cohorts of scRNAseq (Figure 1, each 4 children with
398 neurodevelopmental regression versus 4 controls). These 8 patients were screened
399 for rare pathogenic gene variants that cause ASD and were negative. Cells were

400 clustered into 9 different cell types based on cell markers (Figure 3A), including
401 neutrophils, classical monocytes, natural killer (NK) cells, dendritic cells, basophils,
402 eosinophils, B cells, CD8 T cells, and CD4 T cells.

403

404 scRNAseq cohort 1 was performed in 4 children (mean age 9.5 (7-14) years, all
405 males) with neurodevelopmental regression versus 4 healthy controls (11.3 (7-14)
406 years, all males). A total of 7,240 cells from patients and 6,761 cells from controls
407 were sequenced. Differentially expressed genes (DEGs) per cell type ranged from 0
408 to 868 (Figure 3B, left).

409

410 Upregulated GSEA GO pathways (FDR <0.05) (Figure 3C, in red) clustered in
411 functions involving innate immune (immune response, immune receptor activity),
412 transcription factor, and signaling/matrix pathways (transmembrane signaling
413 receptor activity, and actin filament binding). The upregulated innate immune
414 pathways occurred predominantly in neutrophils, monocytes, NK cells and CD4 T
415 cells. The upregulated transcription factor pathways occurred in neutrophils.

416 Downregulated GSEA GO pathways (FDR <0.05) (Figure 3C, in blue) included
417 immune pathways (MHC class II protein complex) and ribosome/translational
418 pathways. Downregulated ribosome/translation pathways were present across most
419 cell types.

420

421 scRNAseq of cohort 2 was performed in a second cohort of 4 children with
422 neurodevelopmental regression (mean age 8.3 (7-10) years, 50% males) versus a
423 second cohort of 4 healthy controls (mean age 10.5 (8-15) years, 50% males). A total

424 of 10,523 cells from patients and 17,237 cells from controls were sequenced.
425 Differentially expressed genes per cell type ranged from 0 to 429 (Figure 3B, right).
426
427 Upregulated GSEA GO pathways (FDR <0.05) (Figure 3D, in red) related to immune
428 functions (response to molecule of bacterial origin), and epigenetic pathways
429 (nuclear body, ncRNA gene silencing activity, chromatin organization, transcription
430 regulator activity, protein-DNA complex organization). Downregulated GSEA GO
431 pathways (FDR <0.05) (Figure 3D, in blue) included immune (endocytic vesicle, cell
432 killing, antigen binding/processing, T cell mediated cytotoxicity/immunity), and
433 ribosome/translational pathways. Downregulated ribosome/translation pathways
434 were present across most cell types (except CD4 T cells).

435

436 *PBMC proteomics in cohort 2*

437 The 4 children with neurodevelopmental regression, versus 4 healthy controls who
438 were in the second cohort of scRNAseq, also had proteomics and
439 phosphoproteomics performed (Figure 1). Unsupervised principal component
440 analysis of blood proteomics (Figure S4) showed clear discrimination between
441 groups. A total of 432 upregulated and 720 downregulated differentially abundant
442 proteins were found. The top 5 upregulated GSEA GO pathways (FDR <0.05)
443 (Figure 4A) include receptor-mediated endocytosis and histone acetyltransferase
444 complex. Key proteins that enrich the upregulated receptor-mediated endocytosis
445 pathway (Figure 4B) have functions in signal transduction (*RAB21*, *EGF*),
446 cytoskeletal dynamics (*ANXA2*, *TBC1D5*), synaptic function (*FMR1*, *DTNBP1*, *HIP1*),
447 cellular stress response (*CLU*, *SERPINE1*, *PPT1*), chromatin function (*H1-1*) and
448 mRNA processing (*HNRNPK*). Key proteins that enrich the upregulated histone

449 acetyltransferase complex pathways are involved in chromatin remodelling and
450 transcription regulation (*EP400, ACTL6A, RUVBL1, YEATS*) (Figure S4).
451
452 Downregulated GSEA GO pathways (FDR <0.05) (Figure 4A) include cytosolic
453 ribosome and translation pathways. Key proteins that enrich the downregulated
454 cytoplasmic translation pathway (Figure 4C) include ribosomal genes associated
455 with small subunit (40S) or large subunit (60S) of the eukaryotic ribosome (*RPS*,
456 *RPL*), eukaryotic translation initiation factor genes (*EIF*), and RNA methyltransferase
457 (*METTL*).
458
459 *PBMC phosphoproteomics in cohort 2*
460 Unsupervised principal component analysis of blood phosphoproteomics (Figure S5)
461 showed clear discrimination between groups. A total of 2,109 upregulated and 2,396
462 downregulated differentially abundant phosphopeptides were identified. The top 5
463 upregulated ORA GO pathways (FDR <0.05) (Figure 4D) include guanyl-nucleotide
464 exchange factor activity, GTPase activator activity, actin filament binding and
465 cadherin binding. Top 5 downregulated ORA GO pathways (FDR <0.05) (Figure 4E)
466 include regulation of DNA-templated transcription, transcription corepressor activity,
467 and chromatin binding/organization. Key phosphoproteins that enrich the
468 downregulated 'chromatin binding' pathway (Figure 4F, Figure S5) were further
469 subclustered into GO molecular function pathways including histone binding (*CHDs*,
470 *BRDs, KDM1B*), DNA-binding transcription factor binding (*DNMT3A, MEF2A*), and
471 chromatin DNA binding (*HMGNs, HNRNPU*).

472

473 **Discussion**

474 We conducted in-depth peripheral blood multi-omics profiling in a group of children
475 with neurodevelopmental regression, conditions that significantly impair school
476 functioning and quality of life. Most of the children experienced regression
477 compatible with "autistic regression", although some did not exhibit symptoms that
478 fully met the criteria for an ASD diagnosis. Despite the heterogenous clinical
479 presentations, we identified convergent pathophysiological processes underlying
480 these cases. Through peripheral blood bulk RNA sequencing, single-cell RNA
481 sequencing, and bulk proteomics and phosphoproteomics, we identified key findings.
482 These include dysregulated pathways related to epigenetics (chromatin remodelling,
483 histone modifications, transcription factors) and immune responses, as well as
484 downregulated ribosomal and translational pathways in patients with
485 neurodevelopmental regression. Epigenetic and immune processes are increasingly
486 recognized to play critical roles in the pathogenesis of ASD. Highly penetrant
487 chromatin, histone-related and immune genes are well known contributors to ASD⁵⁻⁷.
488 However, genetic testing in our scRNA and proteomic cohorts, including trio whole
489 exome sequencing and autism panel testing, yielded negative results for pathogenic
490 variants associated with autism. However, it is possible that highly penetrant variants
491 in coding or non-coding regions were not captured using this methodology.
492 Additionally, patients in our study may harbor common vulnerability genes that are
493 missed by clinical genomic screening designed to detect rare pathogenic DNA
494 variations. Thus, the abnormal epigenetic and immune landscape found in our study
495 could arise from susceptibility genes, environmental factors, or an interplay of both
496 factors.
497

498 Firstly, we identified a prominent signal of epigenetic dysregulation, involving
499 chromatin organization, histone modifications, and transcription factors. Epigenetic
500 modifications can disrupt critical brain developmental processes, such as neuronal
501 differentiation, migration, synaptic function, and brain connectivity³⁰. Modifications
502 such as histone acetylation, methylation, phosphorylation, and ubiquitination play a
503 pivotal role in shaping chromatin structure, directly impacting DNA accessibility to
504 transcriptional machinery. These modifications influence the tightness of nucleosome
505 packaging within chromatin; for instance, increased histone acetylation results in a
506 more relaxed chromatin configuration, which enhances DNA accessibility and
507 promotes gene transcription. In contrast, histone methylation can have dual effects,
508 either loosening or tightening chromatin, depending on the specific methylation
509 pattern, thereby either activating or repressing gene expression. Potential
510 mechanisms that regulate chromatin remodelling include GTPase and actin
511 binding³¹. GTPase is proposed to regulate actin filaments within chromatin
512 remodellers, influencing their interaction with transcription factors, and playing roles
513 in modulating gene regulation³¹. Our phosphoproteomic analysis identified GTPase
514 and actin pathways, however their specific relevance to chromatin modulation in
515 neurodevelopmental regression require further examination. Collectively, these
516 findings point towards the role of epigenetic modifications in neurodevelopment,
517 through pathways which contribute to chromatin remodelling and gene regulation,
518 driving disruptions in brain development in these children.

519

520 Secondly, dysregulated ribosomal and translational pathways were consistently
521 identified through the multi-omics profiling in our cohort of children with
522 neurodevelopmental regression. The translation of mRNA into functional proteins

523 occurs in ribosomes, which play critical roles in brain development. Disruptions in the
524 translational machinery have been linked to NDDs^{32,33}. Monogenic mutations
525 affecting components of the translational process, such as ribosomal proteins
526 (ribosomopathies), are associated with NDDs³². Additionally, environmental
527 exposures, such as infections and stressors, can impair ribosomal biogenesis,
528 through epigenetic processes³⁴⁻³⁶. During stress conditions, such as proteostasis
529 imbalance, nutrient deprivation, and oxidative stress, cells have been shown to
530 assemble heterochromatin in rDNA, leading to its transcriptional repression^{37,38}.
531 Perturbations in ribosomal biogenesis can reduce protein synthesis crucial for
532 dendritic and synaptic function, contributing to changes in neuronal connectivity
533 associated with NDDs. In our study, the consistent downregulation of ribosomal and
534 translational pathways observed across bulk RNA sequencing, single-cell RNA
535 sequencing, and proteomic analyses highlights the importance of these pathways.
536 These findings align with translational-immune signatures reported in peripheral
537 blood and brain transcriptomic studies of more classical NDDs³⁹. Leukocyte RNA
538 expression from two independent cohorts of ASD boys aged 1 to 4 years (142
539 discovery participants and 73 replication participants) using Illumina microarrays,
540 showed differentially co-expressed genes enriched in translation and
541 immune/inflammation functions³⁹. The gene signature observed in our cohort with
542 neurodevelopmental regression suggests shared underlying pathophysiological
543 mechanisms between regressive and non-regressive NDD phenotypes. Together, we
544 propose that ribosomal protein dysregulation is a central feature in NDDs and a
545 potential marker of abnormal gene (epigenetic) regulation¹⁸.
546

547 Thirdly, we identified dysregulated immune pathways in RNA and proteome
548 sequencing in our cohorts. Single-cell RNA sequencing revealed heterogenous
549 patterns of immune regulation across cell types, with a mix of up- and
550 downregulation instead of a uniform trend. Immune dysregulation has also been
551 observed in various NDDs, through studies of peripheral blood and brain samples in
552 humans⁴⁰⁻⁴². In peripheral blood clinical studies, abnormal cytokine profiles,
553 immunoglobulin levels, immune cell subpopulations, and gene expression profiling of
554 lymphocytes have been elucidated in ASD, ADHD and Tourette syndrome^{40,42}.
555 Several molecular signaling pathways, including those involving cytokines, major
556 histocompatibility complex molecules, complement factors, and microglia, have been
557 identified as links between peripheral and central immune activation, and the
558 development of NDD phenotypes⁴⁰. Several ASD risk genes are involved in immune
559 functions and the transcriptional regulation of immune networks⁴³. Analysis of
560 transcription factor networks in differentially expressed genes from brain tissue and
561 peripheral mononuclear cells in ASD has identified *RELA* and *IRF* as key targets,
562 both of which play crucial roles in regulating the immune system⁴⁴.
563
564 Epigenetic modifications and immune dysregulation in NDDs can also arise from
565 environmental influences. Environmental influences on child brain development
566 begins in the pre-natal period and continues postnatally. Animal models and human
567 epidemiological data support an association between maternal inflammation during
568 pregnancy and offspring NDDs^{2,45}. In our study, there was high prevalence of
569 maternal autoimmunity, psychiatric disorders (e.g. anxiety, depression) and
570 pregnancy related complications (e.g. gestational diabetes, infections). These
571 diverse maternal conditions during pregnancy lead to inflammatory perturbations to

572 the *in utero* environment, and have been shown to adversely affect fetal
573 neurodevelopment and increase the risk of NDDs in human offspring^{2,46}. A meta-
574 analysis of two high-risk prospective human cohort studies involving umbilical cord
575 blood revealed transcriptional changes to chromatin, autoimmune, and
576 environmental response genes in children with NDDs⁴⁷. This suggests that
577 epigenetic factors and immune processes play important roles at the gene-
578 environment interface early in the development of NDDs. Chromatin genes have
579 high expression during gestation and are downregulated after birth. This observation
580 supports the window of vulnerability hypothesis, in which the early stages of brain
581 development are the most sensitive to perturbations, and disruptions during the
582 critical periods can affect neurodevelopment. Additionally, animal models show that
583 immune insults during pregnancy can prime microglia via epigenetic reprogramming,
584 heightening susceptibility to postnatal “second hit” events in offspring, such as
585 infections or stress, which can trigger neurobehavioral abnormalities^{48,49}. In our
586 cohort, infection provoked regressions were highly prevalent, often recurrent, and
587 had significant lasting effects on child development. While some children were able
588 to recover between regressions, others experienced progressive symptoms without
589 return to developmental baseline. Postnatal immune challenges can induce
590 peripheral immune activation, which communicates with brain immune cells through
591 central-peripheral cross-talk, triggering microglia to adopt an activated state and
592 heightened pro-inflammatory response^{49,50}. Post-mortem ASD brain studies reveal
593 upregulation of immune gene expression in microglial markers, corresponding to
594 hypomethylation of these genes, postulated to be driven by environmental
595 factors^{51,52}. Collectively, studies indicate that epigenetic and immune pathways

596 represent a key convergence point for the diverse genetic and environmental risk
597 factors associated with NDDs throughout life.

598

599 The strengths of our study include an in-depth multi-omics analysis in children with
600 autistic and cognitive regression, revealing consistent epigenetic, transcriptional,
601 immune, and ribosomal abnormalities. We hypothesize that epigenetic modifications
602 in brain and blood immune cells contribute to downstream transcriptional, immune,
603 and ribosomal dysregulation, all of which are critical for normal brain development.
604 Limitations of this study include small sample size and heterogeneity in clinical
605 symptomatology. Clinical histories of regression were sometimes obtained
606 retrospectively, which may introduce recall bias. However, extensive efforts were
607 made to corroborate these accounts through video evidence and observations from
608 multiple independent sources (e.g. teachers). Secondly, there is absence of
609 standardised definitions of neurodevelopmental regression, despite ongoing efforts
610 to standardize and validate these criterias¹⁵. Thirdly, omics analyses were performed
611 many years after initial regression event, and the majority of the children were taking
612 psychiatric medications (including anti-depressants) during this study, which can
613 theoretically influence gene and protein expression. Fourthly, this study focused on
614 children with loss of skills or developmental regression, and did not include more
615 classical presentations of NDDs for comparison. As a result, direct comparisons of
616 gene expression profile between our cohort and more classical ASD, OCD or
617 tics/Tourette syndrome was not possible. Fifthly, it is important to highlight that bulk
618 and HIVE single-cell RNA sequencing of whole blood included all leukocyte
619 populations, whereas proteomics and phosphoproteomics were performed on
620 PBMCs (excluding neutrophils). Also, we performed investigations only in blood and

621 not in brain. In clinical cohorts like this, omics analysis of brain tissues is not feasible,
622 making investigations using peripheral blood a more practical and accessible
623 approach. We have shown that multi-omic analysis of peripheral blood has potential
624 clinical utility in distinguishing between patients and controls. Our results compliment
625 previous findings of dysregulated epigenetic, chromatin, and immune pathways in
626 brain tissue of individuals with neurodevelopmental disorders. We propose that
627 peripheral blood RNA and proteomic sequencing can effectively capture the
628 interactions between genetic and environmental factors that are critical for the
629 expression of NDDs, offering valuable insights into their underlying pathogenesis.
630
631 Future studies involving larger, treatment-naïve cohorts of NDDs, together with
632 careful evaluation of genetic backgrounds and environmental exposures, will enable
633 further genotype-phenotype correlations, including those across different
634 developmental trajectories. Previous investigators have proposed that the relative
635 genetic and environmental contributions to ASD include rare pathogenic de novo
636 variants (9.5%), rare pathogenic inherited variants (2.6%), common inherited
637 variants (49.8%) and environmental factors (38.1%)¹⁴. Exploring common gene
638 variants and utilizing polygenic scores could provide valuable insights into whether
639 our omics signature is associated with common genetic variants, environmental
640 factors, or a combination of both. This approach will help identify convergent or
641 divergent pathways associated with these different developmental phenotypes,
642 which will help to develop targeted therapeutics for these different subgroups of
643 patients. Omics analyses should also be conducted at different time points in the
644 disease, to examine the stability of peripheral blood genes and protein expression.
645 Additionally, future studies should incorporate concurrent epigenetic profiling,

646 including ATAC sequencing, histone modifications, and methylation to further
647 elucidate underlying epigenetic modifications for these gene expression
648 perturbations. While our study highlights epigenetic mechanisms, the timing and
649 triggers of epigenetic modifications in these children remain unclear. These changes
650 could stem from heritable epigenetic marks from their parents or may have occurred
651 during pregnancy or early life. To date, treatments for NDDs primarily focus on
652 managing symptoms and providing developmental support. However, therapies that
653 target immune dysregulation or epigenetic modifications, such as butyrate, may offer
654 effective treatment options for these patients^{53,54}.

655

656 **Resource availability**

657 Anonymized data not published within this article and R code will be made available
658 by request from any qualified investigator.

659

660 **Acknowledgements**

661 We would like to thank the patients and their families for participating in this study.

662 We would like to thank the doctors and nurses in the Department of Paediatric

663 Endocrinology at The Children's Hospital at Westmead, Sydney for their help and

664 support to recruit controls. We would like to thank Children's Medical Research

665 Institute for access to the Biomedical Proteomics Facility and bioinformatic support

666 and thank the Australian Genome Research Facility for providing expertise in

667 sequencing.

668

669 **Author contributions**

670 Conceptualization: HN, VXH, SP, RCD; Methodology: HN, VH, SP, RCD;

671 Investigation: HN, VXH, BK, KZ, SA, SP; Formal analysis: HN, VXH, KZ, BG, NA,

672 LLM, XL, RD, MG, SP, RCD, Writing – Original Draft: HN, VH, RCD; Writing –
673 Review & Editing: MG, SP, RCD; Funding Acquisition, RCD; Supervision: SP and
674 RCD.

675

676 **Conflicts of interest**

677 The authors declare no competing interests.

678

679 **Funding**

680 Financial support for the study was granted by the Dale NHMRC Investigator Grant
681 AP1193648, Petre Foundation, Cerebral Palsy Alliance, Bioplatforms Australia,
682 PANS Australia, PANDAS Network and International OCD Foundation. VXH was
683 supported by ExxonMobil NUS Research Scholarship and National University
684 Hospital Singapore (NUHS) Clinician-Scientist Program during the course of this
685 study.

686

687

688

689

690

691

692

693 **Figures**

694 **Figure 1 Study workflow**

695 15 children with neurodevelopmental regression were recruited for this study and
696 underwent comprehensive clinical assessments. Exploratory bulk RNA sequencing
697 was performed, comparing samples from 8 affected children to 8 healthy controls.

698 Genetic screening of 8 of the children with neurodevelopmental regression who were
699 used in scRNA sequencing and proteomic/phosphoproteomic studies (1 overlap with

700 the initial group) revealed no pathogenic mutations. These 8 children were further
701 analyzed using single-cell RNA sequencing in two separate cohorts, each
702 comprising four affected children and four matched controls. Additionally, the second
703 cohort (n=4) underwent proteomic and phosphoproteomic analyses to further
704 investigate the underlying molecular mechanisms.

705

706 **Figure 2 Blood bulk RNA sequencing of children with neurodevelopmental**
707 **regression versus healthy controls**

- 708 **(A)** Principal component analysis (PCA) performed on peripheral blood bulk RNA
709 sequencing performed in children with neurodevelopmental regression and
710 controls. The x-axis represents Principal Component 1 (PC1), while the y-axis
711 represents Principal Component 2 (PC2). Unbiased hierarchical clustering of
712 gene expression between children with neurodevelopmental regression and
713 controls showed separation of data indicating strong group discrimination post
714 RUV.
- 715 **(B)** Bar plot of Top 10 up- and down-regulated Gene Set Enrichment Analysis
716 (GSEA) Gene Ontology (GO) bulk RNA sequencing pathways in children with
717 developmental regression compared to the controls. The top 10 up-regulated
718 GSEA GO pathways (red) were mainly involved in immune function
719 (immunoglobulin complex, myeloid cell differentiation/activation, leukocyte
720 migration), epigenetic mechanisms (demethylation), and signaling
721 transduction (secretory granule lumen, endocytosis, tyrosine kinase
722 signaling). The top 10 down-regulated GSEA GO pathways (blue) included
723 ribosomal biogenesis, translation, and mitochondrial expression pathways.
- 724 **(C)** Genes which enriched the up-regulated myeloid cell differentiation pathway
725 were further clustered via Reactome and included 4 themes including TLR 2
726 cascade, signaling by interleukins, RNA polymerase II transcription, and
727 signaling by Rho GTPases.
- 728 **(D)** Genes which enriched the up-regulated demethylation pathway were also
729 further clustered via Reactome and included 3 themes including HDM (histone
730 demethylases), DNA demethylation enzymes (TET enzymes and TDG
731 (thymine DNA glycosylase)), and DNA damage reversal.

732

733

734

735

736

737

738

739

740 **Figure 3 Single-cell blood RNA sequencing (scRNAseq) cohort 1 and 2 of**
741 **children with neurodevelopmental regression versus controls**

- 742 **(A)** Uniform manifold approximation and projection (UMAP) projection of cohort 1
743 including 4 children with neurodevelopmental regression and 4 control
744 samples in single-cell transcriptomics identified 9 unique cell types based on
745 cell markers, including neutrophils, classical monocytes, natural killer (NK)
746 cells, dendritic cells, basophils, eosinophils, B cells, CD8 T cells and CD4 T
747 cells.

748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797

- (B)** Bar chart of differentially expressed genes (DEGs) in individual cell types in cohort 1 (left) and cohort 2 (right), showing genes that are up-regulated (red) or down-regulated (blue) in neurodevelopmental regressive patients compared to controls.
- (C)** Dot plot visualizing the top 5 GSEA pathways across all cell types within cohort 1. The dot colour represents the normalized enrichment score (NES), with red indicating upregulation and blue indicating downregulation in neurodevelopmental regressive patients compared to controls, while white represents zero. The size of each dot corresponds to the $-\log_{10}(p_{adj})$ of the pathway. Upregulated GSEA GO pathways (FDR <0.05) were clustered in functions involving innate immune (immune response, immune receptor activity), transcription factor, signaling/matrix pathway (transmembrane signaling receptor activity, and actin filament binding). The upregulated innate immune pathways occurred predominantly in neutrophils, monocytes, NK cells and CD4 T cells. The upregulated transcription factor pathways occurred in neutrophils. Downregulated GSEA GO pathways (FDR <0.05) included immune pathways (MHC class II protein complex) and ribosome/translational pathways. Downregulated ribosome/translation pathways were present across most cell types.
- (D)** Dot plot visualizing the top 5 GSEA pathways across all cell types within cohort 2. In neurodevelopmental regression patients compared to controls, upregulated GSEA GO pathways (FDR <0.05) had immune functions (response to molecule of bacterial origin), and epigenetic pathways (nuclear body, ncRNA gene silencing activity, chromatin organization, transcription regulator activity, protein-DNA complex organization). Downregulated GSEA GO pathways (FDR <0.05) included immune (endocytic vesicle, cell killing, antigen binding/processing, T cell mediated cytotoxicity/immunity), and ribosome/translational pathways. Downregulated ribosome/translation pathways were present across most cell types (except CD4 T cells).

Figure 4 Blood proteomics and phosphoproteomics in cohort 2 of children with neurodevelopmental regression versus healthy controls

- (A)** Bar plot of the top 5 up- and down-regulated Gene Set Enrichment Analysis (GSEA) Gene Ontology (GO) proteome pathways in children with

798 developmental regression compared to the controls. The top 2 up-regulated
799 pathways (red) include regulation of receptor-mediated endocytosis, and
800 histone acetyltransferase complex. The top 5 down-regulated pathways (blue)
801 were predominantly ribosomal and translational pathways.

802

803 **(B)** Connectivity network (CNET) plot of the top up-regulated GSEA GO
804 proteome pathway in patients: positive regulation of receptor-mediated
805 endocytosis. The enriched pathways are represented by their respective
806 colors, and corresponding genes' adjusted p value. Key proteins that enrich
807 the pathway have functions in signal transduction (*RAB21*, *EGF*), cytoskeletal
808 dynamics (*ANXA2*, *TBC1D5*), synaptic function (*FMR1*, *DTNBP1*, *HIP1*),
809 cellular stress response (*CLU*, *SERPINE1*, *PPT1*), chromatin function (*H1-1*),
810 and mRNA processing (*HNRNPK*).

811

812 **(C)** CNET plot of down-regulated GSEA GO proteome pathway in patients:
813 cytoplasmic translation. Key proteins that enrich the pathway include
814 ribosomal genes associated with small subunit (40S) or large subunit (60S) of
815 the eukaryotic ribosome (*RPS*, *RPL*), eukaryotic translation initiation factor
816 genes (*EIF*), RNA methyltransferase (*METTL*).

817

818 **(D)** Bar plot of the top 5 up-regulated over representation analysis (ORA) GO
819 phosphoproteomic pathways in patients compared to controls. The top 5
820 upregulated ORA GO pathways include guanyl-nucleotide exchange factor
821 activity, GTPase activator activity, actin filament binding, and cadherin
822 binding.

823

824 **(E)** Bar plot of the top 5 downregulated ORA GO phosphoproteomic pathways in
825 patients compared to controls, include regulation of DNA-templated
826 transcription, transcription corepressor activity, and chromatin
827 binding/organization.

828

829 **(F)** Key phosphoproteins that enrich the downregulated phosphoproteomic
830 'chromatin binding' pathway was further clustered by GO molecular function
831 pathways and included histone binding (*CHDs*, *BRDs*, *KDM1B*), DNA-binding
832 transcription factor binding (*DNMT3A*, *MEF2A*) and chromatin DNA binding
833 (*HMGNs*, *HNRNPU*).

834

835

836

837

838

839

840

841

842

843

844

845

846

847

Supplemental information

848 Table S1: Children with neurodevelopmental regression and healthy controls
849 Table S2: Clinical characteristics of children with neurodevelopmental regression
850 Table S3: Bulk RNA sequencing pathway analysis
851 Table S4: Single-cell RNA seq pathway analysis – Cohort 1
852 Table S5: Single-cell RNA seq pathway analysis – Cohort 2
853 Table S6: Proteomics pathway analysis
854 Table S7: Phosphoproteomics pathway analysis
855
856 Figure S1: Exploratory bulk RNA seq
857 Figure S2: Single-cell RNA seq analysis – Cohort 1
858 Figure S3: Single-cell RNA seq analysis – Cohort 2
859 Figure S4: Proteomics analysis – Cohort 2
860 Figure S5: Phosphoproteomics analysis – Cohort 2

861
862
863

Reference

- 864 1. Zablotzky, B., Black, L.I., Maenner, M.J., Schieve, L.A., Danielson, M.L., Bitsko, R.H.,
865 Blumberg, S.J., Kogan, M.D., and Boyle, C.A. (2019). Prevalence and Trends of
866 Developmental Disabilities among Children in the United States: 2009-2017.
867 *Pediatrics* *144*. [10.1542/peds.2019-0811](https://doi.org/10.1542/peds.2019-0811).
- 868 2. Han, V.X., Patel, S., Jones, H.F., and Dale, R.C. (2021). Maternal immune activation
869 and neuroinflammation in human neurodevelopmental disorders. *Nature Reviews*
870 *Neurology* *17*, 564-579. [10.1038/s41582-021-00530-8](https://doi.org/10.1038/s41582-021-00530-8).
- 871 3. Schafer, D.P., Lehrman, E.K., Kautzman, A.G., Koyama, R., Mardinly, A.R., Yamasaki, R.,
872 Ransohoff, R.M., Greenberg, M.E., Barres, B.A., and Stevens, B. (2012). Microglia
873 sculpt postnatal neural circuits in an activity and complement-dependent manner.
874 *Neuron* *74*, 691-705. [10.1016/j.neuron.2012.03.026](https://doi.org/10.1016/j.neuron.2012.03.026).
- 875 4. Anttila, V., Bulik-Sullivan, B., Finucane, H.K., Walters, R.K., Bras, J., Duncan, L., Escott-
876 Price, V., Falcone, G.J., Gormley, P., Malik, R., et al. (2018). Analysis of shared
877 heritability in common disorders of the brain. *Science* *360*. [10.1126/science.aap8757](https://doi.org/10.1126/science.aap8757).
- 878 5. De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y.,
879 Liu, L., Fromer, M., Walker, S., et al. (2014). Synaptic, transcriptional and chromatin
880 genes disrupted in autism. *Nature* *515*, 209-215. [10.1038/nature13772](https://doi.org/10.1038/nature13772).
- 881 6. Gabriele, M., Lopez Tobon, A., D'Agostino, G., and Testa, G. (2018). The chromatin
882 basis of neurodevelopmental disorders: Rethinking dysfunction along the molecular
883 and temporal axes. *Prog. Neuropsychopharmacol. Biol. Psychiatry* *84*, 306-327.
884 [10.1016/j.pnpbp.2017.12.013](https://doi.org/10.1016/j.pnpbp.2017.12.013).
- 885 7. Lasalle, J. (2013). Autism genes keep turning up chromatin. *OA Autism* *1*.
886 [10.13172/2052-7810-1-2-610](https://doi.org/10.13172/2052-7810-1-2-610).
- 887 8. Al Backer, N.B. (2015). Developmental regression in autism spectrum disorder. *Sudan*
888 *J Paediatr* *15*, 21-26.
- 889 9. Boterberg, S., Charman, T., Marschik, P.B., Bölte, S., and Roeyers, H. (2019).
890 Regression in autism spectrum disorder: A critical overview of retrospective findings
891 and recommendations for future research. *Neuroscience & Biobehavioral Reviews*
892 *102*, 24-55. <https://doi.org/10.1016/j.neubiorev.2019.03.013>.
- 893 10. Richler, J., Luyster, R., Risi, S., Hsu, W.L., Dawson, G., Bernier, R., Dunn, M., Hepburn,
894 S., Hyman, S.L., McMahon, W.M., et al. (2006). Is there a 'regressive phenotype' of

- 895 Autism Spectrum Disorder associated with the measles-mumps-rubella vaccine? A
896 CPEA Study. *J Autism Dev Disord* 36, 299-316. 10.1007/s10803-005-0070-1.
- 897 11. Mouridsen, S.E. (2003). Childhood disintegrative disorder. *Brain and Development* 25,
898 225-228.
- 899 12. Rosman, N.P., and Bergia, B.M. (2013). Childhood disintegrative disorder: distinction
900 from autistic disorder and predictors of outcome. *Journal of Child Neurology* 28,
901 1587-1598.
- 902 13. Goin-Kochel, R.P., Trinh, S., Barber, S., and Bernier, R. (2017). Gene Disrupting
903 Mutations Associated with Regression in Autism Spectrum Disorder. *J Autism Dev*
904 *Disord* 47, 3600-3607. 10.1007/s10803-017-3256-4.
- 905 14. Huguet, G., Benabou, M., and Bourgeron, T. (2016). The Genetics of Autism Spectrum
906 Disorders. In *A Time for Metabolism and Hormones*, P. Sassone-Corsi, and Y. Christen,
907 eds. (Springer Copyright 2016, The Author(s).), pp. 101-129. 10.1007/978-3-319-
908 27069-2_11.
- 909 15. Furley, K., Mehra, C., Goin-Kochel, R.P., Fahey, M.C., Hunter, M.F., Williams, K., and
910 Absoud, M. (2023). Developmental regression in children: Current and future
911 directions. *Cortex* 169, 5-17. <https://doi.org/10.1016/j.cortex.2023.09.001>.
- 912 16. Association, A.P. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth*
913 *Edition, 5th Edition* (American Psychiatric Publishing).
- 914 17. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P.,
915 Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner.
916 *Bioinformatics* 29, 15-21. 10.1093/bioinformatics/bts635.
- 917 18. Tsang, E., Han, V.X., Flutter, C., Alshammery, S., Keating, B.A., Williams, T., Gloss, B.S.,
918 Graham, M.E., Aryamanesh, N., Pang, I., et al. (2024). Ketogenic diet modifies
919 ribosomal protein dysregulation in KMT2D Kabuki syndrome. *EBioMedicine* 104,
920 105156. 10.1016/j.ebiom.2024.105156.
- 921 19. Engholm-Keller, K., and Larsen, M.R. (2016). Improving the Phosphoproteome
922 Coverage for Limited Sample Amounts Using TiO₂-SIMAC-HILIC (TiSH)
923 Phosphopeptide Enrichment and Fractionation. *Methods Mol Biol* 1355, 161-177.
924 10.1007/978-1-4939-3049-4_11.
- 925 20. R Core Team (2013). R: A language and environment for statistical computing. R
926 Foundation for Statistical Computing, Vienna, Austria. URL [http://www.R-](http://www.R-project.org/)
927 [project.org/](http://www.R-project.org/). (2013).
- 928 21. Risso, D., Ngai, J., Speed, T.P., and Dudoit, S. (2014). Normalization of RNA-seq data
929 using factor analysis of control genes or samples. *Nat Biotechnol* 32, 896-902.
930 10.1038/nbt.2931.
- 931 22. Gagnon-Bartsch, J.A., and Speed, T.P. (2012). Using control genes to correct for
932 unwanted variation in microarray data. *Biostatistics* 13, 539-552.
933 10.1093/biostatistics/kxr034.
- 934 23. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015).
935 limma powers differential expression analyses for RNA-sequencing and microarray
936 studies. *Nucleic Acids Res* 43, e47. 10.1093/nar/gkv007.
- 937 24. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee,
938 M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal
939 single-cell data. *Cell* 184, 3573-3587.e3529. 10.1016/j.cell.2021.04.048.

- 940 25. Alquicira-Hernandez, J., Sathe, A., Ji, H.P., Nguyen, Q., and Powell, J.E. (2019). scPred:
941 accurate supervised method for cell-type classification from single-cell RNA-seq data.
942 *Genome Biology* 20, 264. 10.1186/s13059-019-1862-5.
- 943 26. Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y.,
944 Brenner, M., Loh, P.R., and Raychaudhuri, S. (2019). Fast, sensitive and accurate
945 integration of single-cell data with Harmony. *Nat Methods* 16, 1289-1296.
946 10.1038/s41592-019-0619-0.
- 947 27. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A.,
948 Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene
949 set enrichment analysis: a knowledge-based approach for interpreting genome-wide
950 expression profiles. *Proc Natl Acad Sci U S A* 102, 15545-15550.
951 10.1073/pnas.0506580102.
- 952 28. Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A.,
953 Wadi, L., Meyer, M., Wong, J., Xu, C., et al. (2019). Pathway enrichment analysis and
954 visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap.
955 *Nature Protocols* 14, 482-517. 10.1038/s41596-018-0103-9.
- 956 29. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et
957 al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data.
958 *Innovation (Camb)* 2, 100141. 10.1016/j.xinn.2021.100141.
- 959 30. Zahir, F.R., and Brown, C.J. (2011). Epigenetic Impacts on Neurodevelopment:
960 Pathophysiological Mechanisms and Genetic Modes of Action. *Pediatric Research* 69,
961 92-100. 10.1203/PDR.0b013e318213565e.
- 962 31. Klages-Mundt, N.L., Kumar, A., Zhang, Y., Kapoor, P., and Shen, X. (2018). The Nature
963 of Actin-Family Proteins in Chromatin-Modifying Complexes. *Front Genet* 9, 398.
964 10.3389/fgene.2018.00398.
- 965 32. Narla, A., and Ebert, B.L. (2010). Ribosomopathies: human disorders of ribosome
966 dysfunction. *Blood* 115, 3196-3205. 10.1182/blood-2009-10-178129.
- 967 33. Paul, M.S., Duncan, A.R., Genetti, C.A., Pan, H., Jackson, A., Grant, P.E., Shi, J., Pinelli,
968 M., Brunetti-Pierri, N., Garza-Flores, A., et al. (2023). Rare EIF4A2 variants are
969 associated with a neurodevelopmental disorder characterized by intellectual
970 disability, hypotonia, and epilepsy. *Am J Hum Genet* 110, 120-145.
971 10.1016/j.ajhg.2022.11.011.
- 972 34. Vind, A.C., Genzor, A.V., and Bekker-Jensen, S. (2020). Ribosomal stress-surveillance:
973 three pathways is a magic number. *Nucleic Acids Res* 48, 10648-10661.
974 10.1093/nar/gkaa757.
- 975 35. Terhorst, A., Sandikci, A., Whittaker, C.A., Szórádi, T., Holt, L.J., Neurohr, G.E., and
976 Amon, A. (2023). The environmental stress response regulates ribosome content in
977 cell cycle-arrested *S. cerevisiae*. *Front Cell Dev Biol* 11, 1118766.
978 10.3389/fcell.2023.1118766.
- 979 36. Szaflarski, W., Leńniczak-Staszak, M., Sowiński, M., Ojha, S., Aulas, A., Dave, D., Malla,
980 S., Anderson, P., Ivanov, P., and Lyons, S.M. (2022). Early rRNA processing is a stress-
981 dependent regulatory event whose inhibition maintains nucleolar integrity. *Nucleic
982 Acids Res* 50, 1033-1051. 10.1093/nar/gkab1231.
- 983 37. Hirai, H., Takemata, N., Tamura, M., and Ohta, K. (2022). Facultative heterochromatin
984 formation in rDNA is essential for cell survival during nutritional starvation. *Nucleic
985 Acids Research* 50, 3727-3744. 10.1093/nar/gkac175.

- 986 38. Lorenz, D.R., Meyer, L.F., Grady, P.J., Meyer, M.M., and Cam, H.P. (2014).
987 Heterochromatin assembly and transcriptome repression by Set1 in coordination
988 with a class II histone deacetylase. *Elife* 3, e04506. 10.7554/eLife.04506.
- 989 39. Pramparo, T., Pierce, K., Lombardo, M.V., Carter Barnes, C., Marinero, S., Ahrens-
990 Barbeau, C., Murray, S.S., Lopez, L., Xu, R., and Courchesne, E. (2015). Prediction of
991 Autism by Translation and Immune/Inflammation Coexpressed Genes in Toddlers
992 From Pediatric Community Practices. *JAMA Psychiatry* 72, 386-394.
993 10.1001/jamapsychiatry.2014.3008.
- 994 40. Estes, M.L., and McAllister, A.K. (2015). Immune mediators in the brain and
995 peripheral tissues in autism spectrum disorder. *Nat. Rev. Neurosci.* 16, 469-486.
996 10.1038/nrn3978.
- 997 41. Hughes, H.K., Mills Ko, E., Rose, D., and Ashwood, P. (2018). Immune dysfunction and
998 autoimmunity as pathological mechanisms in autism spectrum disorders. *Front. Cell.*
999 *Neurosci.* 12, 405. 10.3389/fncel.2018.00405.
- 1000 42. Martino, D., Zis, P., and Buttiglione, M. (2015). The role of immune mechanisms in
1001 Tourette syndrome. *Brain Res* 1617, 126-143. 10.1016/j.brainres.2014.04.027.
- 1002 43. Arenella, M., Matuleviciute, R., Tamouza, R., Leboyer, M., McAlonan, G., Bralten, J.,
1003 and Murphy, D. (2023). Immunogenetics of autism spectrum disorder: A systematic
1004 literature review. *Brain Behav Immun* 114, 488-499. 10.1016/j.bbi.2023.09.010.
- 1005 44. Gevezova, M., Sbirkov, Y., Sarafian, V., Plaimas, K., Suratane, A., and Maes, M.
1006 (2023). Autistic spectrum disorder (ASD) - Gene, molecular and pathway signatures
1007 linking systemic inflammation, mitochondrial dysfunction, transsynaptic signalling,
1008 and neurodevelopment. *Brain Behav Immun Health* 30, 100646.
1009 10.1016/j.bbih.2023.100646.
- 1010 45. Osman, H.C., Moreno, R., Rose, D., Rowland, M.E., Ciernia, A.V., and Ashwood, P.
1011 (2024). Impact of maternal immune activation and sex on placental and fetal brain
1012 cytokine and gene expression profiles in a preclinical model of neurodevelopmental
1013 disorders. *J Neuroinflammation* 21, 118. 10.1186/s12974-024-03106-7.
- 1014 46. Han, V.X., Patel, S., Jones, H.F., Nielsen, T.C., Mohammad, S.S., Hofer, M.J., Gold, W.,
1015 Brilot, F., Lain, S.J., Nassar, N., and Dale, R.C. (2021). Maternal acute and chronic
1016 inflammation in pregnancy is associated with common neurodevelopmental
1017 disorders: a systematic review. *Translational Psychiatry* 11, 71. 10.1038/s41398-021-
1018 01198-w.
- 1019 47. Mordaunt, C.E., Park, B.Y., Bakulski, K.M., Feinberg, J.I., Croen, L.A., Ladd-Acosta, C.,
1020 Newschaffer, C.J., Volk, H.E., Ozonoff, S., Hertz-Picciotto, I., et al. (2019). A meta-
1021 analysis of two high-risk prospective cohort studies reveals autism-specific
1022 transcriptional changes to chromatin, autoimmune, and environmental response
1023 genes in umbilical cord blood. *Mol Autism* 10, 36. 10.1186/s13229-019-0287-z.
- 1024 48. Martins-Ferreira, R., Leal, B., Costa, P.P., and Ballestar, E. (2021). Microglial innate
1025 memory and epigenetic reprogramming in neurological disorders. *Progress in*
1026 *Neurobiology* 200, 101971. <https://doi.org/10.1016/j.pneurobio.2020.101971>.
- 1027 49. Cheray, M., and Joseph, B. (2018). Epigenetics Control Microglia Plasticity. *Front. Cell.*
1028 *Neurosci.* 12, 243. 10.3389/fncel.2018.00243.
- 1029 50. Tian, L., Ma, L., Kaarela, T., and Li, Z. (2012). Neuroimmune crosstalk in the central
1030 nervous system and its significance for neurological diseases. *Journal of*
1031 *Neuroinflammation* 9, 155. 10.1186/1742-2094-9-155.

- 1032 51. Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., Mill, J., Cantor,
1033 R.M., Blencowe, B.J., and Geschwind, D.H. (2011). Transcriptomic analysis of autistic
1034 brain reveals convergent molecular pathology. *Nature* 474, 380-384.
1035 [10.1038/nature10110](https://doi.org/10.1038/nature10110).
- 1036 52. Ramaswami, G., Won, H., Gandal, M.J., Haney, J., Wang, J.C., Wong, C.C.Y., Sun, W.,
1037 Prabhakar, S., Mill, J., and Geschwind, D.H. (2020). Integrative genomics identifies a
1038 convergent molecular subtype that links epigenomic with transcriptomic differences
1039 in autism. *Nature Communications* 11, 4873. [10.1038/s41467-020-18526-1](https://doi.org/10.1038/s41467-020-18526-1).
- 1040 53. Berni Canani, R., Di Costanzo, M., and Leone, L. (2012). The epigenetic effects of
1041 butyrate: potential therapeutic implications for clinical practice. *Clinical Epigenetics*
1042 4, 4. [10.1186/1868-7083-4-4](https://doi.org/10.1186/1868-7083-4-4).
- 1043 54. Xie, A., Ensink, E., Li, P., Gordevičius, J., Marshall, L.L., George, S., Pospisilik, J.A., Aho,
1044 V.T.E., Houser, M.C., Pereira, P.A.B., et al. (2022). Bacterial Butyrate in Parkinson's
1045 Disease Is Linked to Epigenetic Changes and Depressive Symptoms. *Movement*
1046 Disorders 37, 1644-1653. <https://doi.org/10.1002/mds.29128>.
- 1047
1048
1049
1050

1051 **Reviewer Access Details**

1052

- 1053 • **Bulk RNA sequencing**

1054

1055 To review GEO accession GSE290991:

1056 Go to <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE290991>

1057 Enter token szavcmwaldydjsp into the box

1058

- 1059 • **Single cell RNA sequencing**

1060

1061 Website: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE288764>

1062 Reviewer access token: inojkmikvdujbm

1063

1064

- 1065 • **Proteome/phosphoproteome**

1066

1067 Website: <http://www.ebi.ac.uk/pride>

1068

1069 Submission details:

1070 **Project Name:** Proteomics and phosphoproteomics of children with neurodevelopmental regression

1071 **Project accession:** PXD060771

1072 **Project DOI:** 10.6019/PXD060771

1073

1074 **Reviewer access details**

1075 Log in to the PRIDE website using the following details:

1076 **Project accession:** PXD060771

1077 **Token:** 5zymFt4LbW8r

1078 Alternatively, reviewer can access the dataset by logging in to the PRIDE website using the following

1079 account details:

1080 **Username:** reviewer_pxd060771@ebi.ac.uk

1081 **Password:** oONiEgdWjlpO

1082







