

UNIQUE GENETIC PATHWAYS FOR AUTISM

1

1

2

3

4 **Characterizing Genetic Pathways Unique to Autism Spectrum Disorder at Multiple Levels** 5 **of Biological Analysis**

6

7 Lukas S. Schaffer^{1,2}, Sophie Breunig^{1,2}, Jeremy M. Lawrence^{1,2}, Isabelle F. Foote¹, Andrew D.
8 Grotzinger^{1,2}

9

¹ Institute for Behavioral Genetics, University of Colorado Boulder, Boulder, CO

10 ² Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO

11

UNIQUE GENETIC PATHWAYS FOR AUTISM

2

12 **Abstract**

13 Autism spectrum disorder (ASD) is a neurodevelopmental condition characterized by
14 atypical patterns of social functioning and repetitive/restricted behaviors. ASD commonly co-
15 occurs with ADHD and, despite their clinical distinctiveness, the two share considerable genetic
16 overlap. Given their shared genetic liability, it is unclear which genetic pathways confer unique
17 risk for ASD independent of ADHD. We applied Genomic Structural Equation Modeling (SEM)
18 to GWAS summary statistics for ASD and ADHD, decomposing the genetic signal for ASD into
19 that which is unique to ASD (*uASD*) and that which is shared with ADHD. We computed genetic
20 correlations between *uASD* and 75 external traits to estimate genetic overlap between *uASD* and
21 other clinically relevant phenotypes. We went on to apply Stratified Genomic SEM to identify
22 classes of genes enriched for *uASD*. Finally, we implemented Transcriptome-Wide SEM (T-
23 SEM) to explore patterns of gene-expression associated with *uASD*. We observed positive
24 genetic correlations between *uASD* and several external traits, most notably those relating to
25 cognitive/educational outcomes and internalizing psychiatric traits. Stratified Genomic SEM
26 showed that heritability for *uASD* was significantly enriched in genes involved in evolutionarily
27 conserved processes, as well as for a histone mark in the germinal matrix. T-SEM revealed 83
28 unique genes with expression associated with *uASD*, many of which were novel. These findings
29 delineate the unique biological underpinnings of ASD which exist independent of ADHD and
30 demonstrate the utility of Genomic SEM and its extensions for disambiguating shared and
31 unique risk pathways for genetically overlapping traits.

32

33

34

35

36 **Introduction**

37

38 Autism spectrum disorder (ASD) is a heterogenous neurodevelopmental disorder which
39 occurs in >1% of the population [1, 2]. While the phenotypic presentation of ASD is highly
40 variable, it is characterized by two core symptom domains: (i) impairments in social
41 communication and interaction as well as (ii) repetitive, restricted patterns of behavior or
42 interests [3]. Although the etiology of ASD involves an array of risk factors, extant literature has
43 demonstrated a strong genetic component with heritability estimates from twin and family
44 studies ranging between 64 and 91% [4–6]. A substantial proportion of this heritability (~12%) is
45 attributable to common genetic variation as evidenced by recent genome-wide association
46 studies (GWASs) of ASD [2, 7].

47 ASD often presents alongside other neuropsychiatric conditions; the most frequent
48 comorbidity is attention-deficit/hyperactivity disorder (ADHD) [8], a childhood-onset disorder
49 characterized by symptoms of either inattention, hyperactivity, or both [3]. Indeed, conservative
50 estimates suggest that one in every three children with ASD will also meet diagnostic criteria for
51 ADHD [9]. Converging evidence indicates that a common genetic liability partially underlies
52 risk for both disorders [10]. For example, observations from family studies find that ASD and
53 ADHD tend to co-aggregate in families [11], and this co-aggregation is due, in part, to shared
54 additive genetic influences [12, 13]. These findings are corroborated by molecular and statistical
55 genetic studies, which have estimated moderate genetic correlations between ASD and ADHD
56 [2, 10, 14], indicating a shared genetic architecture.

57 These findings illustrate a broad challenge of parsing disorder-specific biological
58 pathways when two phenotypes are genetically and phenotypically correlated. These difficulties
59 necessitate the need for multivariate genomic analyses capable of isolating the genetic variance

UNIQUE GENETIC PATHWAYS FOR AUTISM

4

60 that is unique to a specific trait. Here, we approach these challenges by leveraging Genomic
61 Structural Equation Modeling (SEM) and its extensions to separate out the genetic signal unique
62 to ASD from that which is shared with ADHD [15–17]. We model GWAS summary statistics for
63 ASD and childhood-diagnosed ADHD using a Cholesky decomposition to derive a unique ASD
64 (henceforth, $uASD$) latent factor, reflecting the residual genetic signal for ASD after removing
65 genetic overlap with ADHD. We then apply downstream analyses to interrogate the genetic
66 architecture of $uASD$ at the genome-wide, functional, and gene-expression levels of analysis.
67 Collectively, these analyses delineate the biological mechanisms that contribute specifically to
68 the etiology of ASD and its associated symptoms, as opposed to those that may confer a broader
69 spectrum of shared neurodevelopmental risk.

70

71 **Method**

72

73 *Summary Statistics*

74

75 Summary statistics for ASD were used from the most recent GWAS meta-analysis [2].
76 Briefly, the original GWAS included 13,076 cases and 22,664 controls from the Danish
77 population-based cohort iPSYCH and 5,305 cases and 5,305 pseudo-controls (i.e., non-
78 transmitted parental alleles) from family-based trio samples from the Psychiatric Genomics
79 Consortium (PGC). Together, the meta-analysis totaled 18 381 ASD cases and 27 969
80 controls/pseudo-controls. GWAS summary statistics for childhood-diagnosed ADHD were
81 utilized from the GWAS conducted by Rajagopal et al. (2022), which stratified ADHD cases by
82 age of diagnosis. Here, we specifically utilize summary statistics for ADHD diagnosed in
83 childhood due to its higher genetic overlap with ASD compared to persistent or adulthood-
84 diagnosed ADHD [14, 18]. The childhood-diagnosed GWAS included 14 878 cases and 38 303
85 controls from the iPSYCH cohort.

UNIQUE GENETIC PATHWAYS FOR AUTISM

5

86 To evaluate genetic overlap with other relevant phenotypes, we leveraged publicly
87 available European-ancestry summary statistics for 77 external traits spanning domains of
88 cognition, psychopathology, health/lifestyle behaviors, interpersonal relations, and physical
89 activity. We used a SNP-based h^2 z -statistic cutoff of 4, as recommended by the original linkage
90 disequilibrium score regression (LDSC) developers [19], to limit our pool of external traits to
91 those with interpretable genetic covariance. Based on this cutoff, 75 of our original 77 traits were
92 carried forward for analysis. A comprehensive list of included external traits and relevant
93 characteristics is reported in **Supplementary Table 1**.

94 *Genomic Structural Equation Modeling*

95

96 Prior to analysis, all GWAS summary statistics underwent an identical set of QC filters
97 using the *munge* function in the *GenomicSEM* R package. These filters included restricting
98 analyses to HapMap3 SNPs and removing SNPs with a minor allele frequency (MAF) < 1% and
99 imputation score (INFO) < .9 (when available). Once processed, these GWAS summary statistics
100 were used as input for multivariable LDSC using the *ldsc* function within *GenomicSEM*, which
101 produces the genetic covariance and sampling covariance matrices across included traits. The
102 genetic covariance matrix includes the SNP-based heritability (h^2_{SNP}) on the diagonal and the
103 genetic covariances on the off-diagonal. The sampling covariance matrix contains squared
104 standard errors (sampling variances) on the diagonal and the sampling covariances (sampling
105 dependencies) on the off-diagonal that will arise in the presence of participant sample overlap.
106 The sampling covariance matrix is estimated directly from the data using a block jackknife
107 resampling procedure and allows for GWAS with varying degrees of power and sample overlap
108 to be included in the same statistical model.

UNIQUE GENETIC PATHWAYS FOR AUTISM

6

109 For binary traits, estimates were converted to the liability scale using the population
110 prevalence and the sum of effective sample size across contributing cohorts [20]. For ASD, the
111 effective sample size was estimated directly from the data. This is because the ASD GWAS used
112 pseudocontrol subjects, which reduces power to detect GWAS associations, such that using the
113 observed sum of effective sample size would produce downwardly biased estimates of
114 heritability [21]. LDSC requires that estimates are produced within a single ancestry group as the
115 LD weights used to estimate the regression model will vary across ancestral populations. Due to
116 limited availability of data in other ancestral groups, GWAS statistics were limited to
117 participants of European ancestry, and the LDSC model was estimated using the 1000 Genomes
118 Phase 3 European LD scores. These scores excluded the major histocompatibility complex
119 (MHC) due to complex LD structures in this region that can bias estimates.

120 The output from LDSC was used as input to all subsequent analyses in Genomic SEM.
121 We began by fitting a Cholesky decomposition model to our observed variables based on GWAS
122 summary statistics for ASD and childhood-diagnosed ADHD (henceforth, simply referred to as
123 ADHD). Both ASD and ADHD were regressed onto a latent factor, $cADHD$, which represents
124 the genetic variance of ADHD as well as the proportion of genetic variance for ASD that is
125 shared with ADHD. ASD was additionally regressed onto $uASD$, representing the residual
126 genetic variance that is unique to ASD after accounting for that which is shared with ADHD. By
127 construction, $uASD$ and $cADHD$ were orthogonal ($r_g = 0$). The genetic residual variances of ASD
128 and ADHD were fixed to 0 so that all variance in the disorders was explained by the latent
129 factors. At the genome-wide level, this model was expanded to compute the genetic correlations
130 between $uASD$ and each of our pre-selected 75 external traits (see **Supplementary Table 2**). In

UNIQUE GENETIC PATHWAYS FOR AUTISM

7

131 interpreting the statistical significance of each genetic correlation, we apply a strict Bonferroni-
132 adjusted significance threshold ($p < 6.7E-4$).

133 *Stratified Genomic SEM*

134

135 We applied Stratified Genomic SEM to identify enrichment for functional annotations
136 (i.e., categories of genes) for *uASD*. We first ran multivariable Stratified LDSC (S-LDSC) to
137 obtain genetic covariance estimates within each annotation. We originally included a total of 168
138 annotations for analysis. However, 12 of these annotations were excluded due to the model
139 failing to converge ($n = 10$) or negative heritability estimates ($n = 2$). The remaining 156 were
140 examined to ensure that none required a smoothing of the covariance matrix resulting in a z -
141 statistic difference > 1.96 (as recommended by Grotzinger et al., 2022) before moving forward.
142 Our final analysis included 49 annotations from the 1000 Genomes Baseline LD Version 2.2
143 [22], as well as neuronal and brain tissue annotations from DEPICT [23], gnomAD [24], GTEx
144 v8 [25], and the Roadmap Epigenomics Project [26] (see **Supplementary Table 3** for full list of
145 annotations). Using the *enrich* function within *GenomicSEM*, we then estimated enrichment for
146 *uASD* within each annotation. Enrichment is calculated as the ratio-of-ratios. For the current
147 analyses, the numerator of this ratio reflects the proportion of *uASD* genetic variance explained
148 by an annotation (i.e., the within-annotation genetic variance for *uASD* divided by the total *uASD*
149 genetic variance). The denominator of the ratio reflects the proportional size of the annotation
150 (i.e., the number of SNPs in the annotation divided by the total number of SNPs analyzed across
151 all annotations). The null for this enrichment ratio-of-ratios is 1, where values above 1 index
152 functional annotations that account for a greater proportion of genetic variance in *uASD* than
153 would be expected based solely on the proportional size of that annotation. Given the non-

UNIQUE GENETIC PATHWAYS FOR AUTISM

8

154 independent nature of functional annotations, we applied the FDR correction for multiple
155 comparisons to the accompanying p -values using the *p.adjust* R package.

156 *Transcriptome-Wide SEM (T-SEM)*

157

158 Transcriptome-wide SEM (T-SEM) was applied to identify patterns of gene expression
159 associated with *uASD*. First, FUSION [27] was used to perform univariate transcriptome-wide
160 association studies (TWASs) on both ASD and ADHD. We utilized functional weights for 13
161 brain tissue types from the Genotype Tissue Expression Project (GTEx v8, [28]), two
162 dorsolateral prefrontal cortex weights from the CommonMind Consortium (CMC, [29]), and one
163 set of weights for the prefrontal cortex from PsychEncode [30]. This resulted in 16 total
164 functional weights from which we derived 73 412 genes with imputed expression data across
165 different brain regions and tissues. These univariate FUSION outputs were then input into the
166 *read_fusion* function in Genomic SEM.

167 The gene expression estimates were then added to the LDSC covariance matrix for ASD
168 and ADHD, and the *userGWAS* function was used to estimate the effect of gene expression on
169 both *uASD* and *cADHD*. Given the scope of the analyses, we specifically focus on the
170 relationship between gene expression and *uASD*. Finally, to identify gene sets with significant
171 enrichment in *uASD*, we used the *WebGestalt* package to conduct an overrepresentation analysis
172 (ORA, [31]) on significant T-SEM hits. An FDR correction was used in interpreting significance
173 of T-SEM and ORA results. We also carried forth a drug repurposing analysis on these hits using
174 methods outlined by Grotzinger et al. 2023. As these analyses did not produce any findings, no
175 results are reported below.

176

177 **Results**

178

179 *Genome-Wide Results Reveal Genetic Correlates of uASD*

UNIQUE GENETIC PATHWAYS FOR AUTISM

9

180
181 ASD and ADHD were moderately genetically correlated ($r_g = .45$, $SE = .05$). The
182 Cholesky decomposition revealed that *cADHD* explained ~20% of the genetic variance in ASD
183 and *uASD* explained the remaining (~80%) ASD genetic variance (see **Figure 1** for visual
184 representation of the structural equation model and partitioning of genetic variance). The
185 remaining analyses sought to clarify what genetically differentiates ASD from ADHD by
186 examining associations with the *uASD* factor at multiple levels of analysis. We began by
187 genetically correlating *uASD* with 75 external traits in order to evaluate the extent of genetic
188 overlap with other phenotypes relevant to mental and physical health. By virtue of our statistical
189 definition of *uASD*, these correlations represented associations extending above and beyond
190 those with *cADHD*. A full list of external traits and relevant outputs is available in
191 **Supplementary Table 2**. Our analyses revealed that *uASD* was significantly correlated with 20
192 external traits. These 20 correlations tended to span four primary phenotypic dimensions –
193 cognition, psychopathology, physical movement, and interpersonal relations – that we review
194 below (and in **Figure 2**, panels **A**, **B**, **C**, and **D**).

195 *Cognition*. Cognitive-related phenotypes demonstrated the most robust genetic
196 correlations with the latent factor *uASD*, both in statistical significance and magnitude. We
197 observed positive correlations between *uASD* and educational attainment ($r_g = .48$, $SE = .05$),
198 childhood intelligence ($r_g = .45$, $SE = .1$), general intelligence ($r_g = .42$, $SE = .05$), noncognitive
199 skills of educational attainment ($r_g = .27$, $SE = .05$), word reading ($r_g = .29$, $SE = .08$), and verbal
200 numerical reasoning ($r_g = .39$, $SE = .05$). Interestingly, a number of these traits (childhood
201 intelligence, noncognitive skills of educational attainment, and word reading) were not found to
202 be significantly associated with ASD, indicating that *uASD* may be capturing additional genetic
203 variance uniquely related to cognitive- and education-related traits. While we observed

UNIQUE GENETIC PATHWAYS FOR AUTISM

10

204 consistent associations with traits indexing intellectual abilities, we did not observe a statistically
205 significant correlation between *uASD* and the trail-making task B, a well-established measure of
206 executive functioning [32, 33]. Collectively these cognitive results indicate that the genetic
207 component unique to ASD is specifically associated with more general cognitive processes
208 independent of self-regulatory processes.

209 *Psychopathology.* Several psychiatric phenotypes, especially those falling within the
210 internalizing spectrum, were positively and significantly associated with *uASD*. These included
211 anxiety ($r_g = .22$, SE = .07), major depressive disorder (MDD; $r_g = .20$, SE = .05), self-harm ($r_g =$
212 $.36$, SE = .10), consideration of self-harm ($r_g = .44$, SE = .08), and sensitivity to environmental
213 stress and adversity ($r_g = .16$, SE = .05). A positive genetic correlation was also observed
214 between *uASD* and schizophrenia ($r_g = .22$, SE = .04). These genetic correlations with
215 psychopathology were similar in directionality and magnitude to those observed for ASD and
216 *cADHD*. This suggests that, despite their high level of genetic overlap, ASD and ADHD each
217 have unique genetic pathways that link them to other forms of psychopathology.

218 Given the shared signal between *uASD* and other psychiatric disorders, we went on to
219 implement a follow-up model to examine the magnitude of unique genetic signal in ASD after
220 partialling out genetic overlap with multiple psychiatric disorders. This involved running a
221 multiple regression model with ADHD, MDD, anxiety, and schizophrenia as correlated
222 predictors of ASD. The multiple regression model path diagram is visualized in **Supplementary**
223 **Figure 1** and its output is provided in **Supplementary Table 3**. This model revealed that 74%
224 of the residual genetic variance in ASD was unexplained by the other psychiatric disorders. This
225 demonstrates that ASD is not merely a genetic amalgamation of ADHD and other psychiatric

UNIQUE GENETIC PATHWAYS FOR AUTISM

11

226 disorders, but has a sizable proportion of unique genetic variance distinguishing it from other
227 forms of psychopathology.

228 *Physical movement.* We examined the genetic correlations with accelerometer data (i.e.,
229 physical movement) in 1-hour increments across a 24-hour period. This accelerometer data can
230 be considered a useful indicator of atypical patterns of movement that may reflect disturbances in
231 sleep and circadian rhythm, wherein these processes have been discussed as transdiagnostic risk
232 factors for psychiatric and neurodevelopmental disorders [16, 34, 35]. We observed positive
233 associations between *uASD* and movement at hours 0-1 (i.e., 12:00AM-1:00AM; $r_g = .29$, SE =
234 .08), 21-22 (i.e., 9:00PM-10:00PM; $r_g = .28$, SE = .07), 22-23 ($r_g = .29$, SE = .08), and 23-24 (r_g
235 = .44, SE = .08), indicating that this genetic overlap was restricted to movement during periods
236 of early morning and late night. We also observed overlap between *uASD* and physical inactivity
237 (r_g for hours of moderate exercise = -.25, SE = .07).

238 *Interpersonal relations.* A select few traits relating to increased social behavior displayed
239 genetic correlations with *uASD*. Notably, these positively associated traits tended to revolve
240 around social relations with family members, such as family satisfaction ($r_g = .41$, SE = .07) and
241 the frequency of friend and family visits ($r_g = .48$, SE = .07). Age of first sexual encounter was
242 also positively genetically correlated with *uASD* ($r_g = .32$, SE = .06), and this association was not
243 observed with ASD.

244 *Functional Results Identify uASD Enrichment*

245
246 Stratified Genomic SEM revealed four significantly enriched functional annotations for
247 *uASD*. Three of these annotations reflected classes of genes implicated in evolutionarily
248 conserved processes, including genes conserved in primates ($p = 1.00E-7$), and two annotations
249 indexing genes conserved in mammals ($p = 5.68E-4$ and $9.45E-4$). We also found significant

UNIQUE GENETIC PATHWAYS FOR AUTISM

12

250 enrichment for the H3K4me1 histone mark in the germinal matrix ($p = 6.20E-4$), a transient
251 brain region present only during gestational brain development (see **Figure 2** and
252 **Supplementary Table 3** for the magnitude of these enrichments). These annotations represent a
253 subset of eight significantly enriched annotations in ASD, which also included genes conserved
254 in vertebrates, as well as enrichment in histone marks in the anterior caudate and fetal male and
255 female brains. The *cADHD* factor captured 26 significantly enriched annotations which
256 encompassed all of the annotations observed in *uASD* and ASD, albeit with smaller point
257 estimates for the evolutionary annotations. Additional significant annotations for *cADHD* were
258 characterized predominantly by markers for genetic modifications in several brain regions and
259 hormonal centers. A full list of annotations and their relative enrichments is provided in
260 **Supplementary Table 3**.

261 *T-SEM Uncovers 83 Genes Associated with uASD*

262
263 We obtained 73 412 gene expression estimates for *uASD* (many of which reflect
264 expression levels for the same gene in different tissues). T-SEM revealed 278 significant hits
265 across 83 unique gene IDs; many of these hits were clustered on chromosomes 8 and 17. These
266 results are visualized as a Miami plot in **Figure 4**. The most highly significant hit corresponded
267 with the downregulation of *PINXI* ($z = -5.79$, $p = 6.89E-9$), a potent inhibitor of telomerase [36].
268 The univariate TWAS of ASD revealed 231 significant hits across 69 genes at the same
269 significance threshold used for the *uASD* T-SEM analysis ($p < 9.37E-5$). The *uASD* T-SEM
270 revealed 34 novel gene hits relative to the ASD univariate TWAS. Despite subtracting out shared
271 signal with ADHD, novel genes can arise in this model for genes with particularly discordant
272 effects across ASD and ADHD.

UNIQUE GENETIC PATHWAYS FOR AUTISM

13

273 To identify potential biological pathways implicated in *uASD*, we applied ORA to the
274 significant genes identified by T-SEM. The analysis revealed two gene sets associated with the
275 *uASD* genes; both of which were implicated in skin-related pathologies. The first set relates to
276 bacterial skin diseases (enrichment ratio = 74.53, $p = 6.55E-9$). The second set corresponded to
277 erythema (redness of the skin, often manifesting as a rash) and included all of the genes in the
278 gene list for bacterial skin diseases, with the exception of *FAM167A* (enrichment ratio = 43.27, p
279 = 2.20E-6).

280

281 **Discussion**

282 The present study leveraged Genomic SEM to dissect the genetic architecture specific to
283 ASD after accounting for shared genetic variance with another neurodevelopmental disorder,
284 ADHD diagnosed in childhood. These ADHD summary statistics evince the highest level of
285 genetic overlap with ASD across the psychiatric space, thereby providing a stringent benchmark
286 for indexing the genetic variance unique to ASD. We find that the majority of genetic variance in
287 ASD, as well as genetic overlap with other clinically relevant traits, is unique from ADHD
288 despite their high levels of genetic overlap [16]. Implementing genome-wide, functional, and
289 gene-expression analyses, we investigated the unique genetic variance at increasing levels of
290 biological granularity. At each level, we interrogated this specific variance and identified distinct
291 biological pathways with specific relevance to ASD.

292 *Genome-Wide Level*

293 Partitioning the genetic variance unique to ASD, we find a plethora of genetic
294 correlations with cognitive, psychiatric, and other behavioral traits. Notably, we find the
295 strongest correlations between *uASD* and traits related to cognition and education. While the

UNIQUE GENETIC PATHWAYS FOR AUTISM

14

296 phenotypic literature surrounding ASD and cognitive abilities is mixed [37–39], our findings that
297 *uASD* has a positive genetic association with intelligence and educational attainment corroborate
298 prior genetic studies of ASD [2, 40–42]. Interestingly, we did not see genetic associations with
299 executive function despite meta-analytic research demonstrating ubiquitous executive
300 functioning deficits in ASD [43, 44]. Cognitive traits showed some of the more divergent
301 patterns of association between *uASD* and *cADHD*, with the latter having large and negative
302 genetic associations with the aforementioned cognitive traits. For many cognitive traits, the
303 magnitude of the correlations seen with *uASD* surpassed those seen within ASD broadly. Thus, it
304 appears the genetic variance unique to ASD may have opposing effects to the variance shared
305 between ASD and ADHD, with the unique component driving correlations in the positive
306 direction.

307 Contrary to cognitive traits, we observed general convergence in the directionality of
308 genetic associations with psychiatric traits, especially in those relating to mood or anxiety
309 disturbances. We therefore conclude that the genetic relationships between ASD and many
310 psychiatric (especially internalizing) phenotypes are not driven solely by the genetic similarity
311 between ASD and ADHD. We also show that ASD is not a simple conglomerate of the genetic
312 components of ADHD and other psychiatric disorders, but rather a genetically distinct construct
313 within the psychiatric space.

314 *Functional Genomic Level*

315 Functional analyses revealed that the genetic signal unique to ASD was concentrated in
316 evolutionarily conserved genes and the H3K4me1 histone mark in the germinal matrix, a
317 transitory brain region present during the prenatal period which serves as a hub for neural
318 progenitor cells [45]. The enrichment of the H3K4me1 histone mark, indicative of active

UNIQUE GENETIC PATHWAYS FOR AUTISM

15

319 enhancer elements, reinforces the importance of early epigenetic modifications underlying
320 neurodevelopmental processes in the pathogenesis of ASD [46]. Moreover, our observation of
321 this unique enrichment underscores the value of employing multivariate genomic analyses to
322 dissect disorder-specific biological pathways, facilitating a deeper understanding of the
323 molecular mechanisms underlying ASD etiology.

324 *Transcriptome-Wide Level*

325 Transcriptome-wide analyses revealed 83 unique genes with differential expression
326 linked with ASD independent of ADHD. Notably, we find several novel gene hits unique to
327 ASD, reflecting genes with highly discordant effects across ASD and ADHD. Focusing on genes
328 with expression associated with *uASD*, we find overlap with gene sets implicated in two classes
329 of skin-related pathologies: bacterial skin disease and erythema (non-specific reddening of the
330 skin). There are well-documented links between ASD and various immune-mediated conditions,
331 with medical research focusing on the association with atopic dermatitis (i.e., eczema) [47–50].
332 Dysfunction of the immune system and inflammatory processes have been hypothesized to
333 contribute jointly to ASD and skin disorders such as atopic dermatitis [51]; however, there is a
334 dearth of literature apart from the current study which has explicitly identified gene clusters
335 implicated in both syndromes.

336 *Limitations*

337 The current analyses were restricted to GWAS summary statistics derived exclusively
338 from individuals of European ancestry [52], which hinders the generalizability of our findings
339 due to differences in allele frequency and LD structure across ancestrally diverse populations
340 [53]. Efforts to broaden representation in GWAS analyses are crucial for extending genetic
341 insights to other ancestral groups [54]. It is also critical to recognize that ASD is a highly

UNIQUE GENETIC PATHWAYS FOR AUTISM

16

342 heterogeneous condition characterized by a broad range of clinical profiles and severity levels
343 [55, 56]. While these diverse phenotypes of ASD are likely attributable to different genetic
344 backgrounds [2], the use of a general ASD GWAS restricts our ability to parse any phenotypic
345 heterogeneity in the sample. Furthermore, many behavioral and psychiatric traits, including
346 ASD, show discrepancies between family- and SNP-based estimates of heritability [57], with a
347 component of ASD's missing heritability thought to be driven by rare, high-impact variants
348 undetected by GWASs [58–60]. It is therefore important to recognize that we are only examining
349 a subset of the genetic factors thought to independently contribute to ASD etiology.

350 *Conclusions*

351 Taken together, we provide insights across multiple levels of biology that characterize the
352 genetic signal unique to ASD. Relative to ADHD, we find evidence for divergent patterns of
353 relationships with a range of clinically relevant correlates (e.g., cognition) along with unique
354 patterns of functional enrichment and gene expression that implicate neurobiological processes
355 and disease states linked to ASD in the extant literature. While ASD has often been discussed as
356 unique within the psychiatric space, the current findings clarify and characterize the biological
357 substrata that differentiate this complex neuropsychiatric disorder.

358

359

360

361

362

363

364

UNIQUE GENETIC PATHWAYS FOR AUTISM

17

365

366 **Acknowledgments**

367 We are extremely grateful for the efforts put forth by the investigators and participants
368 from data resources such as the Psychiatric Genetics Consortium, iPSYCH, and UK Biobank for
369 making these analyses possible.

370

371 **Conflict of Interests**

372 All authors declare no competing financial interests or conflicts of interest.

373

374 **Financial Support**

375 LSS, JML, and ADG are supported by NIH Grant R01MH120219. ADG and IFF are
376 supported by NIA Grant RF1AG073593.

377

378

References

- 379 1. Newschaffer CJ, Croen LA, Daniels J, Giarelli E, Grether JK, Levy SE, et al. The
380 epidemiology of autism spectrum disorders. *Annu Rev Public Health*. 2007;28:235–258.
- 381 2. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, et al. Identification of
382 common genetic risk variants for autism spectrum disorder. *Nat Genet*. 2019;51:431–444.
- 383 3. American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental*
384 *disorders* (5th ed.). <https://doi.org/10.1176/appi.books.9780890425596>
- 385 4. Tick B, Bolton P, Happé F, Rutter M, Rijsdijk F. Heritability of autism spectrum disorders:
386 a meta-analysis of twin studies. *J Child Psychol Psychiatry*. 2016;57:585–595.
- 387 5. Sandin S, Lichtenstein P, Kuja-Halkola R, Hultman C, Larsson H, Reichenberg A. The
388 Heritability of Autism Spectrum Disorder. *JAMA*. 2017;318:1182–1184.
- 389 6. Xie S, Karlsson H, Dalman C, Widman L, Rai D, Gardner RM, et al. The Familial Risk of
390 Autism Spectrum Disorder with and without Intellectual Disability. *Autism Res*.
391 2020;13:2242–2250.
- 392 7. Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB, et al. Most genetic risk
393 for autism resides with common variation. *Nat Genet*. 2014;46:881–885.
- 394 8. Khachadourian V, Mahjani B, Sandin S, Klevzon A, Buxbaum JD, Reichenberg A, et al.
395 Comorbidities in autism spectrum disorder and their etiologies. *Transl Psychiatry*.
396 2023;13:71.
- 397 9. Stevens T, Peng L, Barnard-Brak L. The comorbidity of ADHD in children diagnosed with
398 autism spectrum disorder. *Res Autism Spectr Disord*. 2016;31:11–18.
- 399 10. Peyre H, Schoeler T, Liu C, Williams CM, Hoertel N, Havdahl A, et al. Combining
400 multivariate genomic approaches to elucidate the comorbidity between autism spectrum
401 disorder and attention deficit hyperactivity disorder. *J Child Psychol Psychiatry*.
402 2021;62:1285–1296.
- 403 11. Ghirardi L, Brikell I, Kuja-Halkola R, Freitag CM, Franke B, Asherson P, et al. The familial
404 co-aggregation of ASD and ADHD: a register-based cohort study. *Mol Psychiatry*.
405 2018;23:257–262.
- 406 12. Reiersen AM, Constantino JN, Grimmer M, Martin NG, Todd RD. Evidence for shared
407 genetic influences on self-reported ADHD and autistic symptoms in young adult Australian
408 twins. *Twin Res Hum Genet*. 2008;11:579–585.
- 409 13. Ronald A, Simonoff E, Kuntsi J, Asherson P, Plomin R. Evidence for overlapping genetic
410 influences on autistic and ADHD behaviours in a community twin sample. *J Child Psychol*
411 *Psychiatry*. 2008;49:535–542.
- 412 14. Rajagopal VM, Duan J, Vilar-Ribó L, Grove J, Zayats T, Ramos-Quiroga JA, et al.
413 Differences in the genetic architecture of common and rare variants in childhood, persistent
414 and late-diagnosed attention-deficit hyperactivity disorder. *Nat Genet*. 2022;54:1117–1124.
- 415 15. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al.
416 Genomic structural equation modelling provides insights into the multivariate genetic
417 architecture of complex traits. *Nat Hum Behav*. 2019;3:513–525.
- 418 16. Grotzinger AD, Mallard TT, Akingbuwa WA, Ip HF, Adams MJ, Lewis CM, et al. Genetic
419 architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and
420 molecular genetic levels of analysis. *Nat Genet*. 2022;54:548–559.

UNIQUE GENETIC PATHWAYS FOR AUTISM

19

- 421 17. Grotzinger AD, Singh K, Miller-Fleming TW, Lam M, Mallard TT, Chen Y, et al.
422 Transcriptome-Wide Structural Equation Modeling of 13 Major Psychiatric Disorders for
423 Cross-Disorder Risk and Drug Repurposing. *JAMA Psychiatry*. 2023;80:811–821.
- 424 18. Breunig S, Lawrence JM, Foote IF, Gebhardt HJ, Willcutt EG, Grotzinger AD. Examining
425 Differences in the Genetic and Functional Architecture of Attention-Deficit/Hyperactivity
426 Disorder Diagnosed in Childhood and Adulthood. *Biol Psychiatry Glob Open Sci*.
427 2024;4:100307.
- 428 19. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, et al. An atlas of
429 genetic correlations across human diseases and traits. *Nat Genet*. 2015;47:1236–1241.
- 430 20. Grotzinger AD, Fuente J de la, Privé F, Nivard MG, Tucker-Drob EM. Pervasive
431 Downward Bias in Estimates of Liability-Scale Heritability in Genome-wide Association
432 Study Meta-analysis: A Simple Solution. *Biol Psychiatry*. 2023;93:29–36.
- 433 21. Peyrot WJ, Robinson MR, Penninx BWJH, Wray NR. Exploring Boundaries for the Genetic
434 Consequences of Assortative Mating for Psychiatric Traits. *JAMA Psychiatry*.
435 2016;73:1189–1195.
- 436 22. Belsare S, Levy-Sakin M, Mostovoy Y, Durinck S, Chaudhuri S, Xiao M, et al. Evaluating
437 the quality of the 1000 genomes project data. *BMC Genomics*. 2019;20:620.
- 438 23. Pers TH, Karjalainen JM, Chan Y, Westra H-J, Wood AR, Yang J, et al. Biological
439 interpretation of genome-wide association studies using predicted gene functions. *Nat*
440 *Commun*. 2015;6:5890.
- 441 24. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The
442 mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*.
443 2020;581:434–443.
- 444 25. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot
445 analysis: multitissue gene regulation in humans. *Science*. 2015;348:648–660.
- 446 26. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A,
447 et al. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518:317–330.
- 448 27. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, et al. Integrative approaches
449 for large-scale transcriptome-wide association studies. *Nat Genet*. 2016;48:245–252.
- 450 28. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*.
451 2013;45:580–585.
- 452 29. Hoffman GE, Bendl J, Voloudakis G, Montgomery KS, Sloofman L, Wang Y-C, et al.
453 CommonMind Consortium provides transcriptomic and epigenomic data for Schizophrenia
454 and Bipolar Disorder. *Sci Data*. 2019;6:180.
- 455 30. Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, et al. Transcriptome-
456 wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science*.
457 2018;362.
- 458 31. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit
459 with revamped UIs and APIs. *Nucleic Acids Res*. 2019;47:W199–W205.
- 460 32. Kortte KB, Horner MD, Windham WK. The trail making test, part B: cognitive flexibility
461 or ability to maintain set? *Appl Neuropsychol*. 2002;9:106–109.
- 462 33. MacPherson SE, Allerhand M, Cox SR, Deary IJ. Individual differences in cognitive
463 processes underlying Trail Making Test-B performance in old age: The Lothian Birth
464 Cohort 1936. *Intelligence*. 2019;75:23–32.
- 465 34. Robinson-Shelton A, Malow BA. Sleep Disturbances in Neurodevelopmental Disorders.
466 *Curr Psychiatry Rep*. 2016;18:6.

UNIQUE GENETIC PATHWAYS FOR AUTISM

20

- 467 35. Mijster T, Boersma GJ, van Veen MM, Liemburg E, Cath D, Pijnenborg GHM, et al.
468 Sleep disorders in a naturalistic cohort of Dutch psychiatric outpatients: prevalence rates
469 and associations with psychopathology symptom severity and well-being. *J Sleep Res.*
470 2024;33:e14009.
- 471 36. Zhou XZ, Lu KP. The Pin2/TRF1-interacting protein PinX1 is a potent telomerase inhibitor.
472 *Cell.* 2001;107:347–359.
- 473 37. Joseph RM, Tager-Flusberg H, Lord C. Cognitive profiles and social-communicative
474 functioning in children with autism spectrum disorder. *J Child Psychol Psychiatry.*
475 2002;43:807–821.
- 476 38. Frith U, Hill EL, Tager-Flusberg H, Joseph RM. Identifying neurocognitive phenotypes in
477 autism. *Philos Trans R Soc Lond B Biol Sci.* 2003;358:303–314.
- 478 39. Wolff N, Stroth S, Kamp-Becker I, Roepke S, Roessner V. Autism Spectrum Disorder and
479 IQ - A Complex Interplay. *Front Psychiatry.* 2022;13:856084.
- 480 40. Clarke T-K, Lupton MK, Fernandez-Pujals AM, Starr J, Davies G, Cox S, et al. Common
481 polygenic risk for autism spectrum disorder (ASD) is associated with cognitive ability in the
482 general population. *Mol Psychiatry.* 2016;21:419–425.
- 483 41. Rao S, Baranova A, Yao Y, Wang J, Zhang F. Genetic Relationships between Attention-
484 Deficit/Hyperactivity Disorder, Autism Spectrum Disorder, and Intelligence.
485 *Neuropsychobiology.* 2022;81:484–496.
- 486 42. Hatoum AS, Morrison CL, Mitchell EC, Lam M, Benca-Bachman CE, Reineberg AE, et al.
487 Genome-wide Association Study Shows That Executive Functioning Is Influenced by
488 GABAergic Processes and Is a Neurocognitive Genetic Correlate of Psychiatric Disorders.
489 *Biol Psychiatry.* 2023;93:59–70.
- 490 43. Demetriou EA, Lampit A, Quintana DS, Naismith SL, Song YJC, Pye JE, et al. Autism
491 spectrum disorders: a meta-analysis of executive function. *Mol Psychiatry.* 2018;23:1198–
492 1204.
- 493 44. Hemmers J, Baethge C, Vogeley K, Falter-Wagner CM. Are Executive Dysfunctions
494 Relevant for the Autism-Specific Cognitive Profile? *Front Psychiatry.* 2022;13:886588.
- 495 45. Raets MMA, Dudink J, Govaert P. Neonatal disorders of germinal matrix. *J Matern Fetal*
496 *Neonatal Med.* 2015;28 Suppl 1:2286–2290.
- 497 46. Yoon SH, Choi J, Lee WJ, Do JT. Genetic and Epigenetic Etiology Underlying Autism
498 Spectrum Disorder. *J Clin Med Res.* 2020;9.
- 499 47. Billeci L, Tonacci A, Tartarisco G, Ruta L, Pioggia G, Gangemi S. Association Between
500 Atopic Dermatitis and Autism Spectrum Disorders: A Systematic Review. *Am J Clin*
501 *Dermatol.* 2015;16:371–388.
- 502 48. Zerbo O, Leong A, Barcellos L, Bernal P, Fireman B, Croen LA. Immune mediated
503 conditions in autism spectrum disorders. *Brain Behav Immun.* 2015;46:232–236.
- 504 49. Jameson C, Boulton KA, Silove N, Guastella AJ. Eczema and related atopic diseases are
505 associated with increased symptom severity in children with autism spectrum disorder.
506 *Transl Psychiatry.* 2022;12:415.
- 507 50. Wang B, Li Y-D, Wang Z-Y, Zhao J-Q, Zhang G-Q, Man M-Q. Alterations in Epidermal
508 Biophysical Properties in Autistic Children. *Skin Pharmacol Physiol.* 2023;36:160–164.
- 509 51. Man M-Q, Yang S, Mauro TM, Zhang G, Zhu T. Link between the skin and autism
510 spectrum disorder. *Front Psychiatry.* 2023;14:1265472.
- 511 52. Sirugo G, Williams SM, Tishkoff SA. The Missing Diversity in Human Genetic Studies.
512 *Cell.* 2019;177:26–31.

UNIQUE GENETIC PATHWAYS FOR AUTISM

21

- 513 53. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current
514 polygenic risk scores may exacerbate health disparities. *Nat Genet.* 2019;51:584–591.
- 515 54. Fatumo S, Chikowore T, Choudhury A, Ayub M, Martin AR, Kuchenbaecker K. A
516 roadmap to increase diversity in genomic studies. *Nat Med.* 2022;28:243–250.
- 517 55. Masi A, DeMayo MM, Glozier N, Guastella AJ. An Overview of Autism Spectrum
518 Disorder, Heterogeneity and Treatment Options. *Neurosci Bull.* 2017;33:183–193.
- 519 56. Montgomery A, Masi A, Whitehouse A, Veenstra-VanderWeele J, Shuffrey L, Shen MD, et
520 al. Identification of subgroups of children in the Australian Autism Biobank using latent
521 class analysis. *Child Adolesc Psychiatry Ment Health.* 2023;17:27.
- 522 57. Maher B. Personal genomes: The case of the missing heritability. Nature Publishing Group
523 UK. 2008. <http://dx.doi.org/10.1038/456018a>. Accessed 26 April 2024.
- 524 58. Antaki D, Guevara J, Maihofer AX, Klein M, Gujral M, Grove J, et al. A phenotypic
525 spectrum of autism is attributable to the combined effects of rare variants, polygenic risk
526 and sex. *Nat Genet.* 2022;54:1284–1292.
- 527 59. Mollon J, Almasy L, Jacquemont S, Glahn DC. The contribution of copy number variants to
528 psychiatric symptoms and cognitive ability. *Mol Psychiatry.* 2023. 3 February 2023.
529 <https://doi.org/10.1038/s41380-023-01978-4>.
- 530 60. More RP, Warriar V, Brunel H, Buckingham C, Smith P, Allison C, et al. Identifying rare
531 genetic variants in 21 highly multiplex autism families: the role of diagnosis and autistic
532 traits. *Mol Psychiatry.* 2023;28:2148–2157.

533

534

535

Figure Notes

536 **Figure 1. Genomic structural equation modeling to decompose ASD genetic variance. (A)**
537 Cholesky decomposition model producing $uASD$, a latent variable encompassing the genetic
538 variance unique to ASD independent of ADHD, and $cADHD$, which captures the residual genetic
539 variance of ASD (i.e., variance shared between ASD and ADHD) and the genetic variance of
540 ADHD. **(B)** Donut plot showing the proportion of residual genetic variance unique to ASD
541 (blue) and shared with childhood ADHD (red).

542

543 **Figure 2. Genetic correlations between $uASD$ and external traits.** Genetic correlations
544 between $uASD$ (blue) and external traits for domains of cognition and education **(A)**, psychiatric
545 disorders and symptoms **(B)**, and interpersonal traits **(C)**. Traits are sorted top to bottom by
546 ascending p -value for the $uASD$ correlation. **(D)** Genetic correlations between accelerometer-
547 based average total hourly movement within the 24-hour day beginning at midnight (i.e., hour 1)
548 and $uASD$ and $cADHD$. Correlations are also shown between external traits and ASD (pink) as
549 well as $cADHD$ (red). Error bars represent 95% confidence intervals. Translucent points and
550 error bars represent genetic correlations that did not surpass the Bonferroni-adjusted significance
551 threshold. Panel D depicts a LOESS regression line used to visualize overall trends across
552 individual point estimates. Shaded region around the regression line represents a 95% confidence
553 interval. Dashed pink line represents the LOESS regression line for genetic correlations with
554 ASD. EA = educational attainment, EF = executive function, MDD = major depressive disorder,
555 OCD = obsessive-compulsive disorder, PTSD = post-traumatic stress disorder.

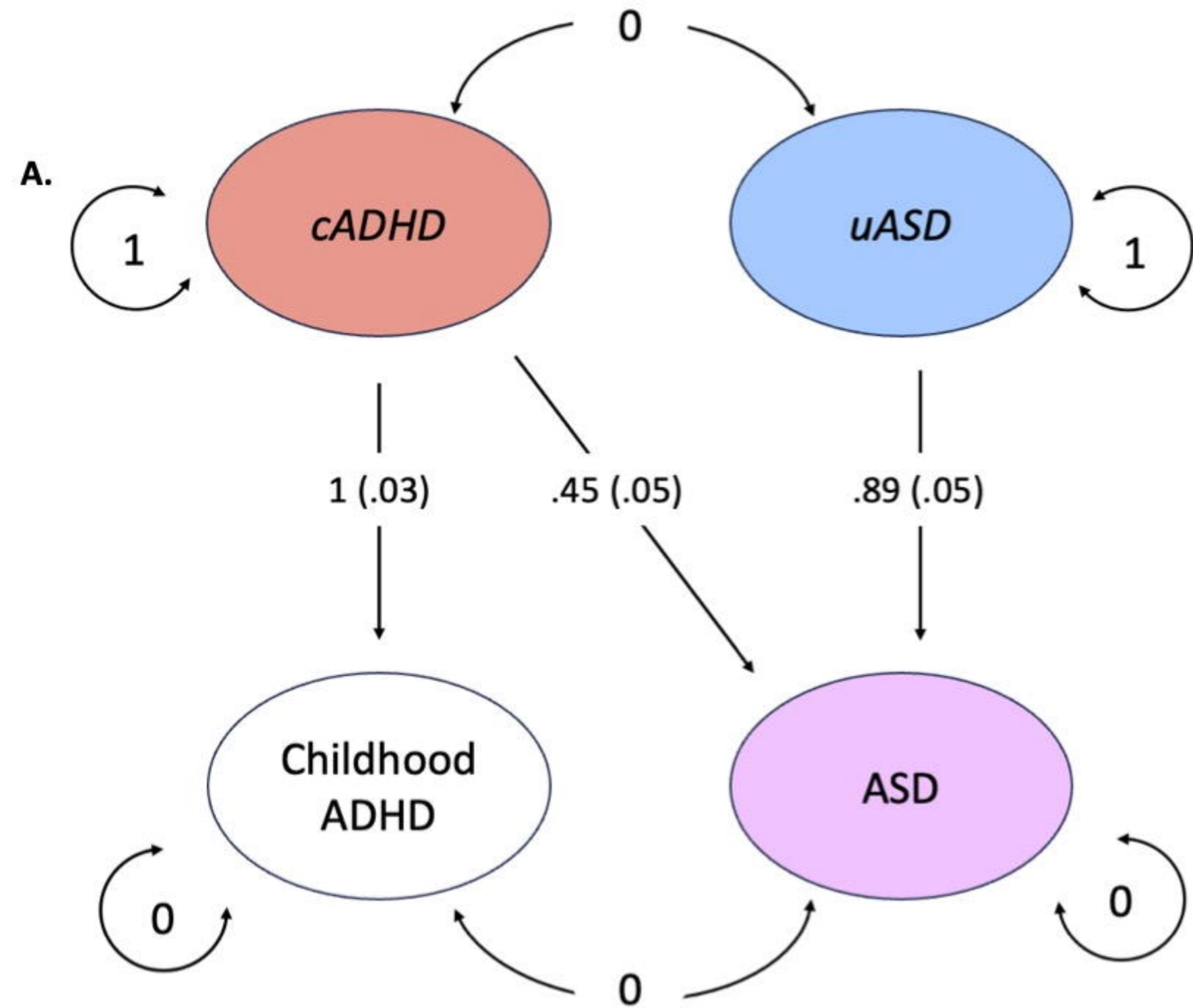
556

557 **Figure 3. Genetic enrichment of $uASD$ for functional annotations.** Functional annotations are
558 arranged by ascending p -value. Enrichment is measured as the ratio of the proportion of genome-
559 wide relative risk represented by the size of that annotation relative to the entire genome. Null
560 enrichment value is 1.0 (visualized by dashed vertical line), in which the genetic variance
561 captured by that annotation is proportional to the expected genetic variance based on annotation
562 size. Significant enrichments at an FDR threshold are represented by solid blue bars, and error
563 bars represent the 95% CI around the enrichment estimate. For visualization purposes,
564 enrichment values are also provided for ASD (pink) and $cADHD$ (red). Single asterisk (*)
565 indicates conserved mammalian genes defined by GERP score. Double asterisk (**) indicates
566 conserved mammalian genes defined by Lindblad-Toh et al. (2011).

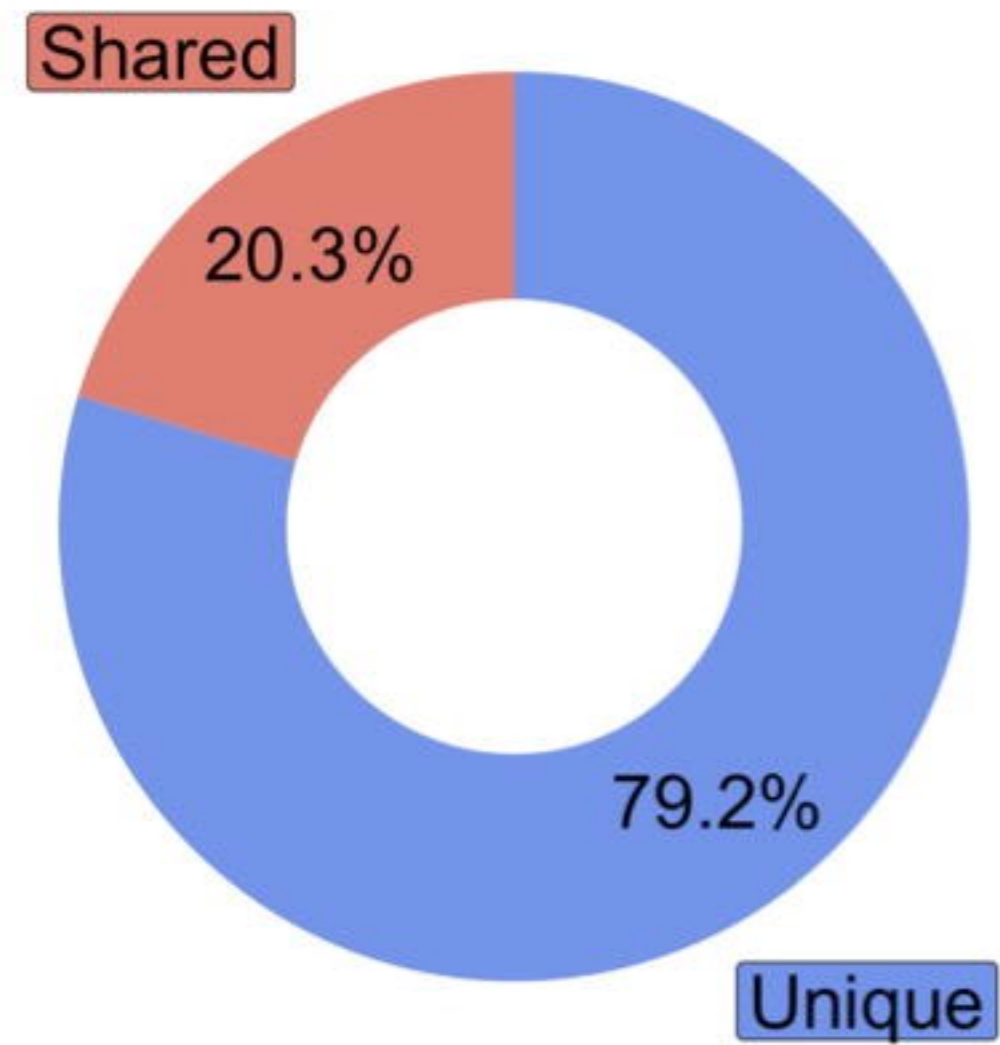
567

568 **Figure 4. Miami plot of gene expression hits for $uASD$.** Upper and lower bounds represent the
569 FDR-adjusted significance threshold ($p < 9.37E-5$). Genes surpassing the upper bound are
570 upregulated and those below the lower bound are downregulated. Significant hits are colored red
571 and labeled with gene ID names.

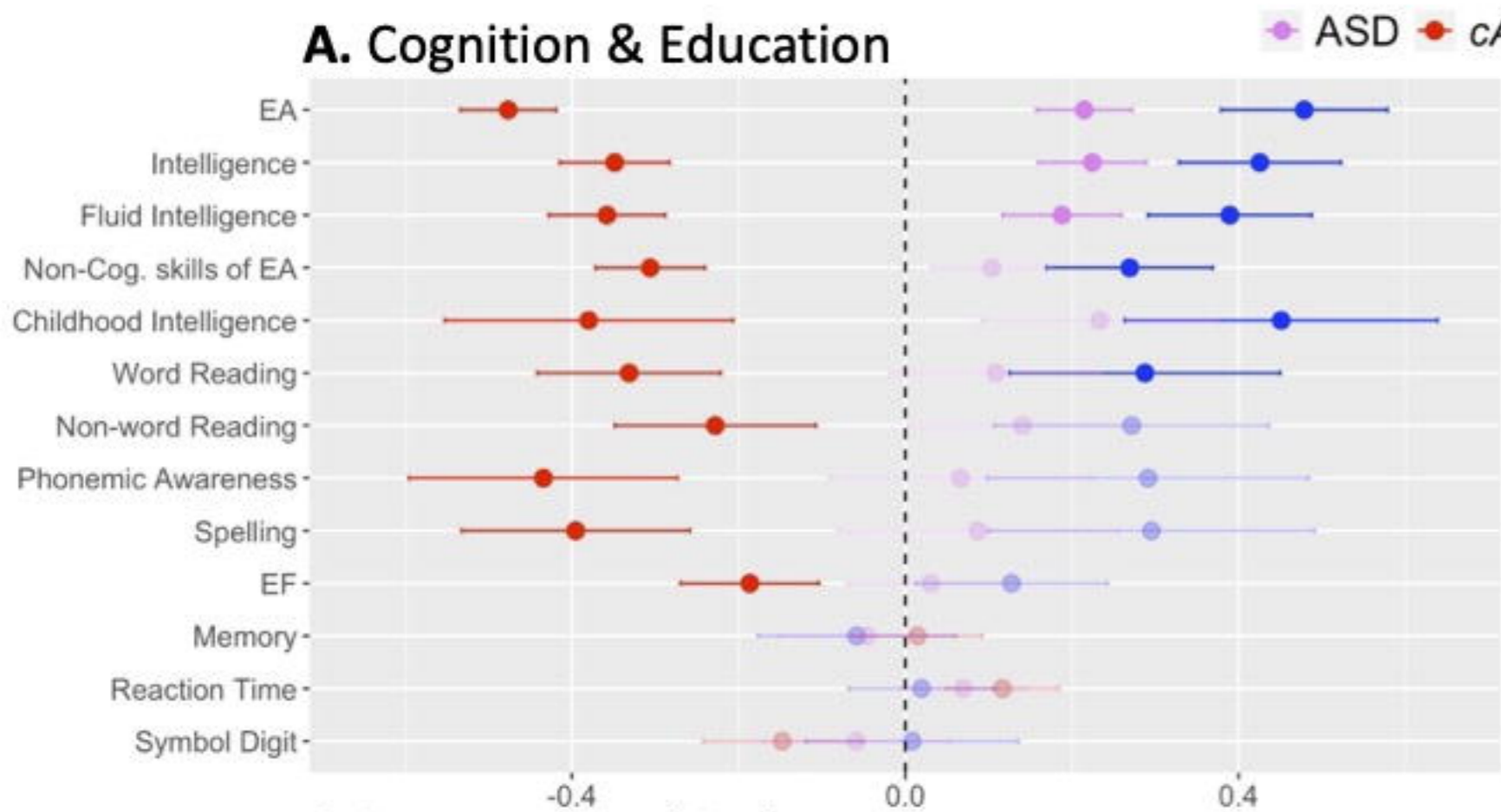
572



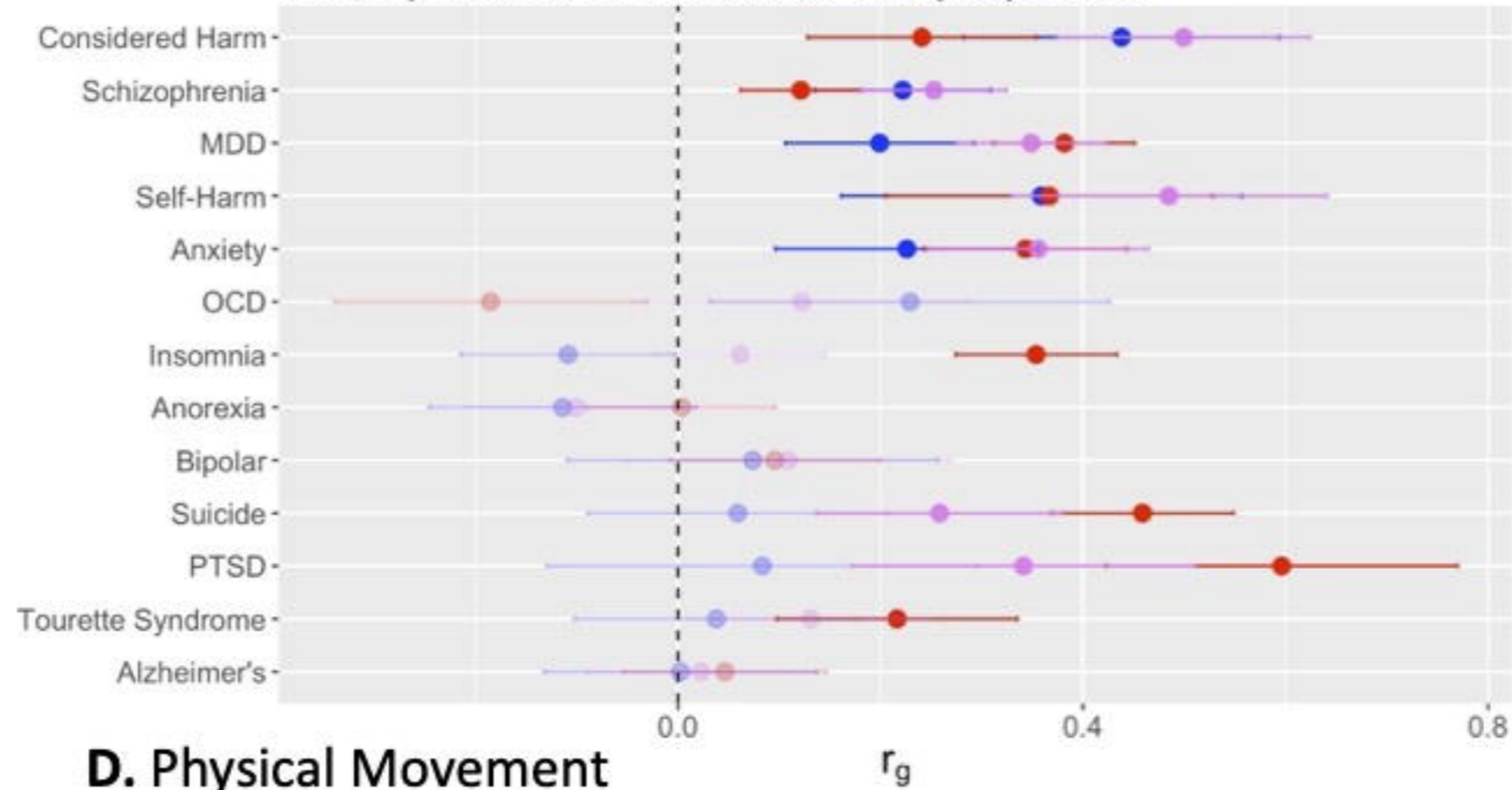
B.



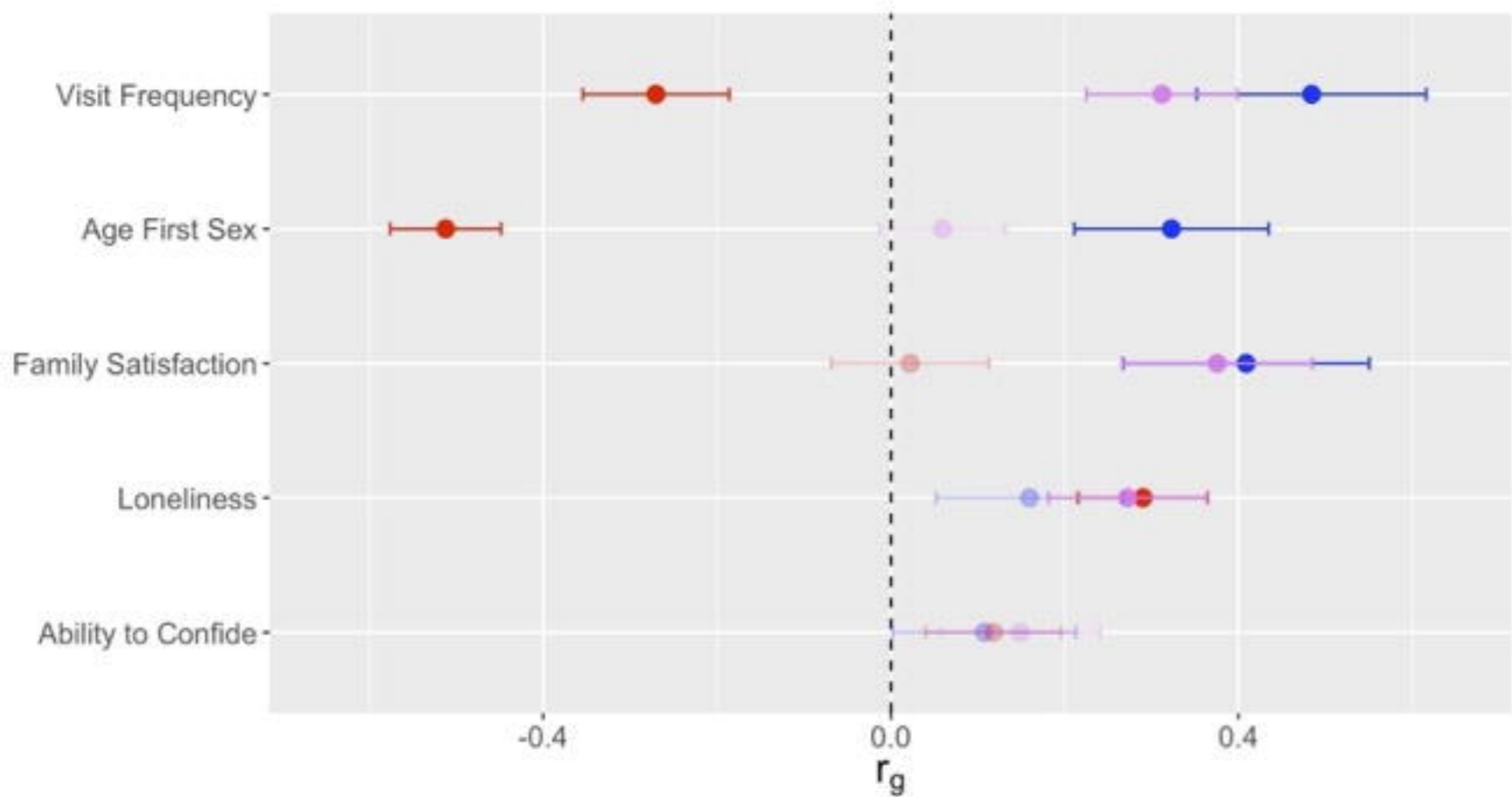
A. Cognition & Education



B. Psychiatric Disorders & Symptoms



C. Interpersonal Traits



D. Physical Movement

