

GWAS integrating maternal and child genetics reveals novel risk-modifying loci for fetal alcohol spectrum disorders (FASD)

Robert Carter

rcc2142@cumc.columbia.edu

Institute of Human Nutrition

Vivek Ruhela

Columbia University

Corina Lesseur

Ariadna Cilleros-Portet

Sandra W. Jacobson

Wayne State University

Joseph L. Jacobson

Wayne State University

Ernesta Meintjes

University of Cape Town Faculty of Health Sciences <https://orcid.org/0000-0002-0781-3632>

Neil Dodge

Tugba Akkaya-Hocagil

H. Eugene Hoyme

Haoxiang Cheng

Jia Chen

Ke Hao

Icahn School of Medicine at Mount Sinai <https://orcid.org/0000-0002-1815-9197>

Maya A. Deyssenroth

Columbia University

Giuseppe Tosto

Columbia University <https://orcid.org/0000-0001-7075-8245>

Article

Keywords:

Posted Date: September 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9950101/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: The authors have declared there is **NO** conflict of interest to disclose

1 **Title: GWAS integrating maternal and child genetics reveals novel risk-modifying loci for**
2 **fetal alcohol spectrum disorders (FASD)**

3 **Running title: GWAS of FASD integrating maternal-child genetics**

4 **Authors:**

5 Vivek Ruhela, PhD¹

6 Corina Lesseur, PhD²

7 Ariadna Cilleros-Portet, PhD²

8 Sandra W. Jacobson, PhD^{3,4,5}

9 Joseph L. Jacobson, PhD^{3,4}

10 Ernesta M. Meintjes, PhD⁴

11 Neil C. Dodge, PhD³

12 Tugba Akkaya-Hocagil⁶

13 H. Eugene Hoyme, MD^{7,8}

14 Haoxiang Cheng, PhD⁹

15 Jia Chen, ScD²

16 Ke Hao, PhD⁹

17 Maya A. Deyssenroth, DrPH¹⁰

18 Giuseppe Tosto, MD, PhD^{1*}

19 R. Colin Carter, MD, MMSc^{4,11*}

20 **Affiliations**

21 ¹Department of Neurology, Taub Institute for Research on Alzheimer's Disease and the Aging
22 Brain, and Gertrude H. Sergievsky Center, Columbia Vagelos College of Physicians and
23 Surgeons, New York, New York, USA.

24 ²Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount
25 Sinai, New York, NY, USA.

26 ³Department of Psychiatry and Behavioral Neurosciences, Wayne State University School of
27 Medicine, Detroit, Michigan, USA.

28 ⁴Department of Human Biology, Faculty of Health Sciences, University of Cape Town, Cape
29 Town, South Africa.

30 ⁵Department of Psychiatry and Mental Health, Faculty of Health Sciences, University of Cape
31 Town, Cape Town, South Africa.

32 ⁶Department of Biostatistics, School of Medicine, Ankara University, Ankara, Turkey

33 ⁷Sanford Children's Genomic Medicine Consortium, Sanford Health, Sioux Falls, South Dakota;
34 Department of Pediatrics, USA.

35 ⁸University of Arizona College of Medicine, Tucson, Arizona, USA.

36 ⁹Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New
37 York, NY, USA.

38 ¹⁰Department of Environmental Health Sciences, Mailman School of Public Health, Columbia
39 University, New York, NY, USA.

40 ¹¹Departments of Emergency Medicine and Pediatrics, Institute of Human Nutrition, Columbia
41 Vagelos College of Physicians and Surgeons, New York, New York, USA.

42 *Denotes shared senior authorship.

43 **Corresponding Author:**

44 R. Colin Carter, MD, MMSc

45 622 168th St

46 VC-260

47 New York, NY 10032 USA

48 +1(212)305-2995

49 rcc2142@cumc.columbia.edu

50

51

52 Abstract

53 Fetal alcohol spectrum disorders (FASD) exhibit wide phenotypic variability not explained by
54 prenatal alcohol exposure (PAE) alone. While genetic factors may modify vulnerability, there are
55 no prior FASD genome-wide association studies (GWAS). We performed genome-wide single-
56 marker and gene-based analyses of two South African prospective birth cohorts (Cape Town
57 Longitudinal Cohorts (CTLC); n=328 mother-child dyads) and the US Collaborative Initiative on
58 Fetal Alcohol Spectrum Disorders cohort (CIFASD n=429 children only). We examined main
59 genetic and gene-by-PAE interaction effects on working memory, recognition memory, height, and
60 FASD diagnosis using single-marker and gene-based analyses and the Haplotype-based
61 Transmission Disequilibrium Test (HTDT), which integrates maternal-child genetics. Genome-
62 wide genotype array data were used; TOPMed-imputed data were used to refine results. Risk-
63 modifying genes were then tested for differential mRNA expression based on PAE in placenta and
64 child blood. Gene*PAE interaction gene-based analyses identified *STX6* associated with height-
65 for-age Z-scores ($p=2.2\times10^{-6}$) and *FOXD3* with recognition memory performance ($p=4.7\times10^{-6}$) in
66 children. Single-marker analysis detected one child SNP approaching significance (*rs2833924*,
67 $p=8.8\times10^{-8}$) with Gene*PAE interaction for working memory performance. CTLC+CIFASD meta-
68 analysis in children identified one significant locus (*HTR1E*; $p_{meta}=2.75\times10^{-6}$) with Gene*PAE
69 interaction for working memory performance. The HTDT test in CTLC imputed data identified three
70 haplotypes associated with FASD diagnosis risk: *COP1* ($p_{adj}=0.029$), *WWOX* ($p_{adj}=2\times10^{-4}$),
71 *TMEM38B* ($p_{adj}=0.019$). *WWOX*, *STX6*, and *TMEM38B* were differentially expressed in placenta
72 and/or child blood based on PAE. In this first-ever genome-wide investigation of maternal and
73 child genetics in FASD, we identified multiple biologically plausible risk-modifying loci for FASD
74 neurobehavioral, growth, and FASD diagnosis outcomes.

76 1. INTRODUCTION

77 Prenatal alcohol exposure (PAE) is the are among the most common preventable causes of
78 neurodevelopmental disabilities worldwide [1, 2]. Despite prevention efforts, 14% of women in the
79 US (10% worldwide) continue to drink during pregnancy [3]. Prevalence estimates of fetal alcohol
80 spectrum disorders (FASD), the diagnostic term for children with adverse effects of PAE, range
81 from 1-5% in the US and Europe to 13.6-30.6% in other parts of the world [1, 2]. FASD manifest
82 a broad range of cognitive/behavioral deficits and growth restriction; teratogenic effects of alcohol
83 have been reported in every organ system [4–8]. Individuals with FASD exhibit lifelong learning
84 disabilities, behavioral problems, and social challenges, which affect individual quality of life and
85 place substantial demands on healthcare, education, and social systems [9]. Fetal alcohol
86 syndrome (FAS) is characterized by distinctive facial dysmorphology, small head circumference,
87 and pre- and/or postnatal growth restriction [6]. Partial FAS (PFAS) includes individuals with
88 characteristic facial anomalies and one of: small head circumference, growth restriction, or
89 neurobehavioral deficits. Children with the most prevalent form of FASD, alcohol-related
90 neurodevelopmental disorder (ARND), are the most difficult to identify, as they lack the
91 characteristic facial dysmorphology, but still have cognitive impairment that is often as severe as
92 children with FAS or PFAS.

93 Children with prenatal alcohol exposure (PAE) exhibit remarkable variability in the severity
94 of neurocognitive deficits, which is not explained by dose and timing of alcohol exposure alone;
95 only 10-15% of children with heavy PAE develop FASD, while 4-10% develop full fetal alcohol
96 syndrome (FAS) [10]. Genetic factors make large contributions to FASD phenotypic variability in
97 both human and animal studies. Animal models have shown that different genetic strains exhibit
98 different vulnerability to the teratogenic effects of alcohol [11]. In human sibling studies,
99 monozygotic twins were almost fully concordant for FASD diagnosis, while 36-44% of dizygotic
100 twins and 41 and 22% of full- and half-sibling pairs were concordant, respectively[10]. In

101 prospective longitudinal birth cohorts in Cape Town, South Africa, we recently found that although
102 ancestry was not associated with cognition in the absence of PAE, maternal and child ancestry
103 were associated with altered risk of FASD outcomes [12]. For example, for every 1% increase in
104 maternal or child East Asian ancestry, the odds of an FAS/PFAS diagnosis decreased by 15%.

105 Maternal genetic profiles are needed to identify women at risk for having a child with FASD
106 in need of intensive screening and prevention efforts. Child profiles are needed to identify children
107 at high risk for FASD for whom treatment interventions could be started early in life when they are
108 most effective. Novel risk-modifying genetic loci may also elucidate mechanisms underlying the
109 teratogenic effects of alcohol and identify novel therapeutic targets. However, despite three
110 decades of research, few genetic determinants of FASD phenotypic variability have been
111 identified. The dearth of prior knowledge regarding genetic influences in FASD may be due to
112 prior analyses being limited to candidate genes [10]. Candidate gene approaches in human
113 studies have identified important potential risk-modifying polymorphisms in genes involved in
114 alcohol metabolism [13–17], choline transport [18], retinoic acid pathways [19], and N-
115 glycosylation [20]. However, candidate gene studies require *a priori* knowledge of influential
116 genes, which is often lacking and sometimes incorrect. By contrast, GWAS studies may identify
117 completely novel genes with pathophysiologic significance. However, to our knowledge, none
118 have been employed in FASD studies to date. This may be due in part to genome-wide interaction
119 screens are especially power-intensive, making such studies particularly difficult in FASD
120 research given the relatively small sample size of cohorts with high levels of PAE and deep
121 phenotyping and the relatively low exposure levels in larger cohorts [10, 21]. Innovative
122 approaches are thus needed.

123 To our knowledge, prior studies of genetic influences in FASD have not examined both
124 maternal and child genetics. Examining both maternal and child genetics is critical given the
125 impact of both on the intrauterine environment and shared genetics between the mother and child.

126 While some genes may be relevant only in the mother (e.g., alcohol metabolism genes), others
127 may be relevant only in the child (e.g., genes impacting neurodevelopment), and others may be
128 relevant in *both* mother and child (e.g., genes in nutrient metabolism, placental development, etc.).
129 Furthermore, when examined in isolation, a variant identified as risk-modifying in the child may
130 actually be a proxy for the mother's genotype, or vice versa, due to shared genetics.

131 In this study, we sought to identify genetic loci in the mother and child that alter fetal
132 vulnerability to PAE in genome-wide analyses from two South African prospective longitudinal
133 birth cohorts and the US-based Collaborative Initiative for FASD (CIFASD) clinical cohort. To our
134 knowledge, this is the first genome-wide study of genetic influences in FASD and the first to
135 integrate maternal and child genetics. We first employed classic single-marker genome-wide
136 screens, separately for mothers and children. To reduce multiple testing burden, we then
137 employed a gene-based approach, separately for mothers and children. Recognizing that these
138 approaches do not examine maternal and child genetics simultaneously, we applied the Haplin
139 Transmission Disequilibrium Test (HTDT), which disentangles maternal and child genetic
140 contributions to disease risk.

141

142 **2. MATERIALS AND METHODS**

143 2.1 Cape Town Longitudinal Cohort (CTLC) description

144 Data are from two prospective longitudinal cohorts designed to examine the developmental
145 effects of PAE: one recruited from 1999-2002 ("CTLC 1") [8, 22]; the other from 2011-2015 ("CTLC
146 2") [23, 24]. Alcohol consumption at recruitment was ascertained in timeline follow-back interviews
147 [25, 26]. Women were invited to participate if they averaged ≥ 1.0 oz (30 mL) absolute alcohol
148 (AA)/day (≈ 1.67 standard drinks) or reported binge drinking (based on National Institute on Alcohol
149 Abuse and Alcoholism (NIAAA)) binge definition at the time of recruitment: ≥ 2.5 oz (75 mL)

150 AA/occasion for CTLC 1; (≥ 2.0 oz (60 mL) AA/ occasion for CTLC 2). Women who abstained or
151 drank only minimally (< 1 oz (30 ml) AA/day on average with no binge episodes) were enrolled as
152 controls. Exclusion criteria at recruitment included maternal age < 18 years, HIV infection, multiple
153 gestation pregnancy, and pharmacologic treatment for medical conditions. Infant exclusion criteria
154 included major chromosome anomalies, recognizable malformation syndromes (other than
155 FASD), neural tube defects, very low birthweight (< 1500 g), and extreme prematurity (gestation
156 < 29 weeks). DNA samples were unavailable for both mother and child for 25 mother-child dyads
157 with heavy exposure and 20. Nine participants failed initial quality control procedures (see 2.3
158 below). The final cohorts for the analyses presented here included 128 mother-infant dyads (76
159 with heavy PAE, 52 controls) from CTLC 1; 200 mother-infant dyads (120 with heavy PAE, 80
160 controls) from CTLC 2 (**Figure 1**).

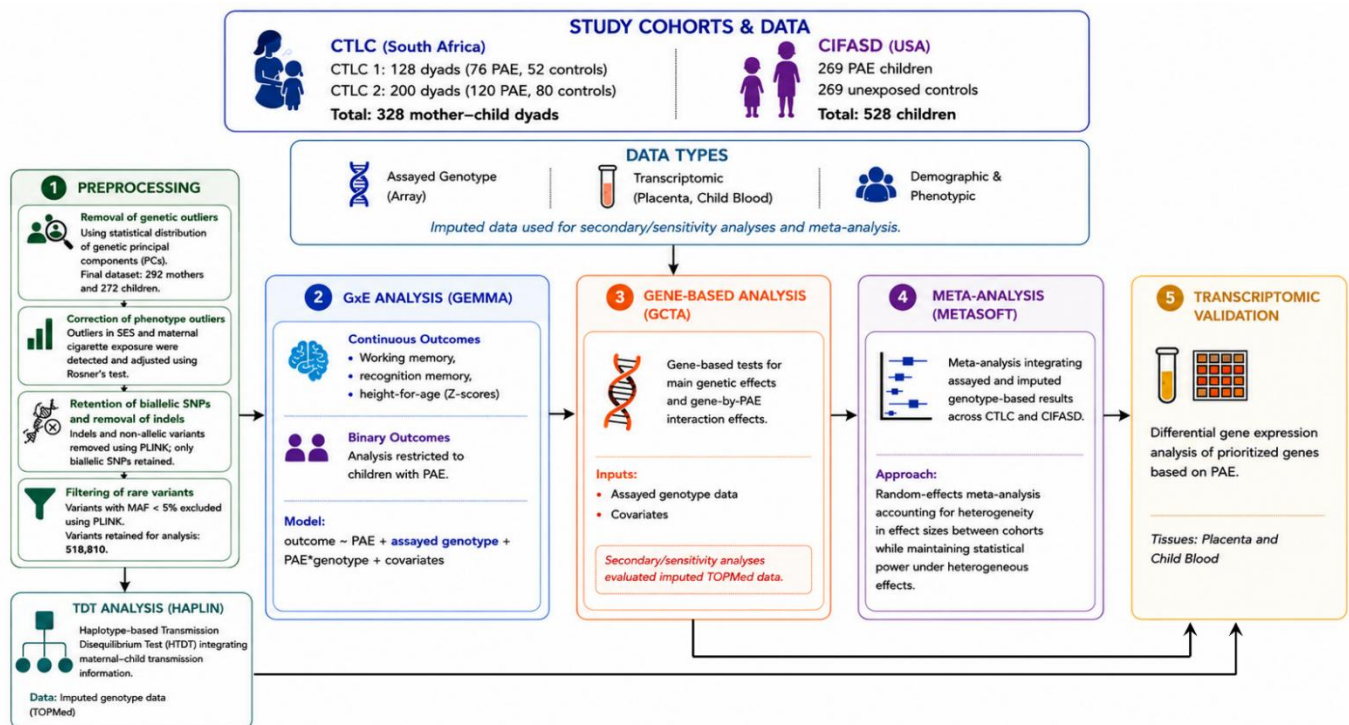
161 Each woman provided informed consent in the mother's preferred language (Afrikaans or
162 English). Women reporting alcohol or drug use were counseled by our staff about the risks of
163 prenatal alcohol and drug use, were encouraged to stop or reduce their alcohol and/or drug use
164 and were offered referrals to substance use treatment programs.

165

166 2.2 Prenatal alcohol exposure, covariates, and phenotypic characterizations

167 In timeline follow-back interviews [25] administered at recruitment and again in the 3rd
168 trimester (CTLC 1) or 4 and 12 weeks after recruitment (CTLC 2), each mother was asked about
169 her alcohol, smoking, and drug use on a day-by-day basis during the past 2 weeks, with recall
170 linked to specific daily activities. Maternal age and Hollingshead socioeconomic (SES) scores
171 were obtained in maternal interviews [27].

172 Four *a priori* outcomes were chosen to capture FASD clinical spectrum: working memory,
173 recognition memory, height-for-age (Z-scores), and FAS/PFAS diagnoses. Working memory and



174

175 **Figure 1. Overview of the study design and analytical workflow for genome-wide investigation of maternal**
 176 **and child genetic factors associated with FASD outcomes.** Genotype array, TOPMed-imputed genotype,
 177 transcriptomic, demographic, and phenotypic data from the Cape Town Longitudinal Cohorts (CTLC) and
 178 Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD) cohorts were analyzed. Following genotype
 179 quality control and preprocessing, gene-by-environment (GxE) interaction analyses were performed using GEMMA,
 180 gene-based analyses using GCTA, and cohort-level meta-analyses using METASOFT. Haplotype-based
 181 Transmission Disequilibrium Test (HTDT) analyses were conducted in CTLC mother-child dyads using Haplin.
 182 Prioritized loci were subsequently evaluated using transcriptomic validation in placenta and child blood datasets.

183

184 recognition memory assessments were the California Verbal Learning Test-Children's version
 185 (Pearson Clinical Assessments, USA) at age 10 years in CTLC 1 and the Wide Range
 186 Assessment of Memory and Learning, 2nd Edition (Pearson Clinical Assessments, USA) at age 7
 187 years for CTLC 2. Height was measured by trained staff, and age- and sex-specific Z-scores
 188 calculated [28] (ages 1 and 5 years in both cohorts and ages 9 and 16 years in CTLC1). In
 189 diagnostic clinics in 2005, 2009, 2013, and 2016, children were independently examined for

190 growth and FASD-related dysmorphology by expert FASD dysmorphologist HE Hoyme MD and
191 other trained FASD dysmorphologists, all blinded to PAE and prior diagnoses. FASD diagnoses
192 were determined in case conferences using the Hoyme criteria [6]; consistent with these criteria,
193 “cases” for neurobehavioral outcomes were defined as scoring <1.5 SD below the mean for
194 unexposed children and <10th percentile for height.

195

196 2.3 Genotyping Assays

197 Please see Supplementary Materials for sample collection and DNA extraction details. DNA from
198 whole blood or saliva samples were assayed using the Illumina (San Diego, USA) Infinium
199 Expanded Multi-Ethnic Genotyping Array (MEGA^{EX}), which contains >2 million SNPs and covers
200 65.7% of GWAS catalog SNPs. Genotyped data were cleaned with standard QC measures using
201 PLINK (v1.9). In brief, SNPs were excluded if Hardy-Weinberg equilibrium p -value <1e-06 and
202 missing call rate <95%; individuals were removed if genotype missingness ≥2%. Principal
203 component analysis (PCA) was conducted using *king* software [29] to account for variance in
204 population stratification and to identify unrelated set of samples (defined as kinship <0.09).

205 To identify genetic outliers, we performed PCA analysis using PLINK (version 1.9.10) [30, 31] on
206 individual SNP genotype data and compared the distribution of the top principal components
207 (PCs). Individuals having PC values beyond the range of (Mean of all PCs) ± (four standard
208 deviations) were excluded (4 children and 16 mothers). Covariate outliers were identified
209 employing the Rosner test [32]: for a) cohort, b) mothers' age, c) SES, d) cigarette exposure, and
210 e) weekly gestation, only SES and cigarette exposure showed outliers and were thus adjusted.
211 Values exceeding the upper outlier threshold identified by the Rosner test were winsorized by
212 replacing them with the corresponding upper-bound cutoff value, a commonly used robust
213 statistical approach for minimizing the influence of extreme observations while preserving sample

size and distributional stability [33]. 256 mother-child dyads had complete genotype data for the Haplin transmission disequilibrium test (HTDT; see below). Rare SNPs (those with minor allele frequency (MAF) <5%) were filtered out. The pre-processing workflow is shown in Supplementary Methods (Figure S1).

2.4 Gene Expression Assays

Please see Supplementary Materials for sample collection and RNA extraction details. Using bulk-tissue RNA extracted from placental and child blood (age 5 years for CTLC1; 9 years for CTLC2) samples, libraries were sequenced to a depth of approximately 50 million 100-bp paired-end reads per sample. Raw sequencing reads were pre-processed to remove adapter sequences and filtered to eliminate low-quality reads (Phred score < 20). Reads were aligned to the GRCh38.v33 human reference genome using the STAR aligner, and gene-level quantification was performed. Only protein-coding genes with raw counts >10 in at least 50% of the samples were retained, resulting in a final expression matrix of 14,867 genes. One placenta sample showing >5% estimated maternal cell contamination, identified through prior integration of maternal-infant genotyping and transcriptomic data, was excluded. After additional quality control filtering, including removal of low-RNA Integrity Number (RIN) samples, the final dataset included RNA-seq data from 62 high-quality placental samples and 308 high-quality blood samples.

2.7 Imputation

To increase power to detect signals associated with SNPs that might be missing from the genotype array, we performed genotype imputation for CTLC data using the TOPMed Imputation Server [34]. Variants with an imputation quality score (R^2) below 80% were filtered out and the resulting high-quality data were converted into PLINK format for processing and further genetic association and analytical workflows (**Figure 1**).

238 2.6 CIFASD cohort dataset

239 The CIFASD data resource includes data from 269 children with PAE (defined as >4
240 drinks/occasion at least weekly or >13 drinks/week) and 269 normotypically developing
241 unexposed (<1 drink/week and <2 drinks/occasion) children with child genetic, demographic, and
242 phenotypic data [35]. Fully de-identified data are available to researchers after application to and
243 approval by the CIFASD steering committee. 107 individuals were removed due to lack of
244 exposure data. No genetic outliers were detected based on PCA. Children ages 7-18 yr were
245 recruited from 2007-2012 and 2014-2017 from clinical and research sites at San Diego State
246 University, UCLA, University of Minnesota, and Emory University. Exclusion criteria included:
247 significant head injury with loss of consciousness, significant or psychiatric (e.g. psychosis)
248 disability that would preclude participation and, for the exposed group, other known causes of
249 cognitive deficiency (e.g. congenital hypothyroidism, neurofibromatosis, chromosomal
250 abnormalities). 17% of exposed children have full FAS. The cohort is 32% African American, 19%
251 White, and 49% Hispanic. Outcomes examined included age-/sex-specific height Z-scores,
252 working memory (Behavioral Rating Inventory of Executive Function-Preschool (BRIEF-P)),
253 recognition memory (Differential Ability Scales (2nd Edition; DAS-II) Recall of Designs (within-
254 cohort Z-score) or, if DAS-II was not administered, Cambridge Neuropsychological Test
255 Automated Battery Delayed Matching to Sample (within-cohort Z-score)), and clinical FAS
256 diagnosis.

257 Samples were genotyped on the OmniExpress genome array (Illumina) or Multi-Ethnic
258 Genotyping Array [36][37], and imputed using the Michigan Imputation Server [34] using the HG19
259 build 1000 Genomes Phase 3, b37 reference panel [38]. The final dataset consisted of 4,000,362
260 SNPs with MAF ≥ 0.05 , genotype rate ≥ 0.99 , and Hardy–Weinberg equilibrium $P \geq 0.000001$
261 **(Figure 1).**

262 2.8 Statistical Analysis

263 All statistical analyses were performed using R (version 4.3.1) [39]. The complete workflow
264 overview is shown in **Figure 1**.

265 For working memory, recognition memory, height-for-age (Z-scores), we employed the following
266 model:

$$267 \quad \text{outcome} \sim \text{PAE} + \text{genotype} + \text{PAE} * \text{genotype} + \text{covariates}.$$

268 For FAS/PFAS diagnosis – since it can only occur in the setting of PAE - we limited the analyses
269 to children with PAE and employed the following model:

$$270 \quad \text{Diagnosis} \sim \text{genotype} + \text{covariates}.$$

271 Based on our prior studies, covariates included cohort (CTLC1 vs CTLC2), maternal age at
272 conception and SES for all outcomes and, additionally for height-for-age and FAS/PFAS
273 diagnosis, prenatal cigarette exposure (cigarettes/day) and weeks gestation at delivery. In
274 CIFASD, SES, maternal age at conception, and prenatal cigarette exposure data were not
275 available. All analyses included the top three PCs.

276 Separately for mothers and children in CTLC and children in CIFASD, we employed the
277 genome-wide efficient mixed-model association using the GEMMA [40] for single-marker analyses
278 and the Genome-wide Complex Trait Analysis (GCTA) [41] for gene-based analyses (see
279 Supplementary for full options employed). For binary outcomes, GEMMA models the phenotype
280 using a linear mixed-model approximation rather than a generalized linear mixed model with a
281 logistic link function. For continuous outcomes, linear mixed models with an identity link function
282 were employed. To increase statistical power and evaluate consistency of genetic effects across
283 cohorts, we performed a meta-analysis using the MetaSoft framework under the Han and Eskin
284 random-effects model, which accounts for heterogeneity in effect sizes between cohorts while

285 maintaining statistical power under heterogeneous effects [42]. Meta-analysis was performed
286 separately for each neurobehavioral and growth outcome across cohorts.

287 For mother–child dyads, we conducted haplotype-based association testing using the
288 Haplotype-based Transmission Disequilibrium Test (HTDT) implemented in the Haplin R package
289 [43, 44]. Haplin applies maximum likelihood estimation to phased haplotypes and accommodates
290 uncertain phase and missing parental genotypes. We analyzed 4-SNP sliding-window haplotypes
291 under a multiplicative dose–response model, which estimates per-allele dosage effects. For each
292 haplotype, Haplin produces separate maternal and child effect estimates, including relative risks
293 (RR) with 95% confidence intervals, single and double dose p-values, and population haplotype
294 frequencies, allowing us to distinguish whether associations arise from maternal genotype and
295 fetal genotype patterns. Because sliding-window haplotypes overlap extensively, the resulting
296 tests are not independent, and standard FDR procedures (e.g., unweighted Benjamini–Hochberg)
297 can be overly conservative. Therefore, we applied a frequency-weighted Benjamini–Hochberg
298 (BH) correction, where haplotypes with higher population frequency receive proportionally greater
299 weight. This approach improves statistical power while maintaining false discovery rate control by
300 incorporating independent covariate information (here, haplotype frequency) into the multiple-
301 testing procedure, as implemented in the Independent Hypothesis Weighting (IHW) framework
302 [45].

303 For genes identified in the above analyses, differential gene expression was examined using
304 DESeq2 [46] based on exposure (yes vs. no) was examined, adjusting for surrogate variables,
305 the covariates described above, and gene expression-inferred cell-type proportions for all cell
306 types with mean proportion >2% in the cohort [47] using the model:

307 *Gene expression ~ exposure + tissue-specific cell-type composition + technical covariates*

308

309 3. RESULTS

310 3.1 Sample characteristics.

311 As we have previously reported [12], mothers were 27.4 years old at conception on average, and
312 heavy drinking mothers were older than controls (**Table 1**). Heavy drinking and control mothers
313 completed 8-10 years of school, on average, with heavy drinking women averaging fewer years
314 of school than controls. Heavy drinking mothers in both cohorts exhibited a weekend-binge pattern
315 of drinking, averaging 7.0 standard drinks/occasion on 1.4 days per week in CTLC 1 and 6.7
316 standard drinks/occasion on 1.4 days per week in CTLC 2. Most control women abstained from
317 drinking during pregnancy, with 12 women reporting infrequent, light consumption (no binges (i.e.,
318 <4 standard drinks/occasion), with drinking occasions <2/month for 10, <1/week for 2). Cigarette
319 smoking was more common among heavy drinking women than controls in CTLC 1, but the
320 number of cigarettes smoked per day was similarly low in both groups. 14.5% of heavy drinking
321 women vs. 3.8 % of controls in CTLC 1 and 24.2% of heavy drinking women vs. 10.0 % of controls
322 in CTLC 2 reported prenatal marijuana use. 11.7% vs. 15.0% of control women in CTLC 2 reported
323 prenatal methamphetamine use; no women in CTLC 1 reported methamphetamine use, as this
324 was uncommon in the area in the early 2000's. Methaqualone, cocaine, and opiate use were not
325 reported by any participant. Most women (82.0% in CTLC 1; 86.5% in CTLC 2) delivered at term.
326 47.4% of exposed children in CTLC 1 and 30.8% of children in CTLC 2 met diagnostic criteria for
327 FAS or PFAS. The CIFASD database has been extensively previously described [35]. Detailed
328 CTLC sample characteristics are provided in **Supplementary Table S1**.

329

330 3.2 Gene-by-PAE interaction analyses for neurobehavior and growth outcomes examining child 331 and maternal genetic data separately

332

333 **Table 1.** Sample characteristics in the Cape Town Longitudinal Cohorts

		Controls	Heavy drinking mothers	<i>p</i> ^a
		<i>M</i> (SD) / <i>n</i> (%)	<i>M</i> (SD) / <i>n</i> (%)	
N		132	196	
Maternal characteristics				
Maternal age at conception (yr)		25.84 (5.23)	27.33 (6.07)	0.018
Education (yr school completed)		9.88 (1.75)	8.71 (2.22)	< 0.001
Socioeconomic status score ^b		22.82 (7.64)	19.84 (7.67)	0.006
Alcohol consumption	oz AA/day	0.0 (0.0)	0.90 (1.1)	< 0.001
	oz AA/drinking day	0.10 (0.35)	4.20 (2.42)	< 0.001
<i>n</i> (%) reporting cigarette smoking		80 (60.6)	169 (86.22)	< 0.001
Cigarettes/day among smokers		6.04 (5.40)	7.11 (4.71)	0.1293
Child characteristics				
Weeks gestation at delivery		38.97 (2.05)	38.69 (2.05)	0.2437
Infant sex, no. female (%)		65 (49.24)	100 (51.02)	0.7531
FAS or PFAS diagnosis, <i>n</i> (%)		0 (0)	72 (36.73)	< 0.001

334

335 FAS = fetal alcohol syndrome using Hoyme 2005 criteria [6]; PFAS = partial FAS.

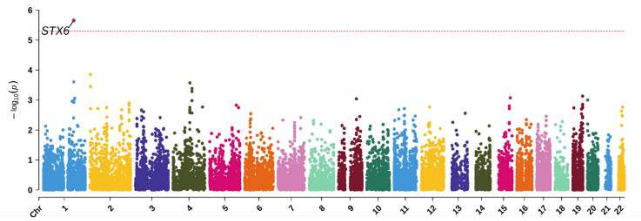
336 ^aFrom independent samples t-tests for continuous variables and chi-square for categorical variables.

337

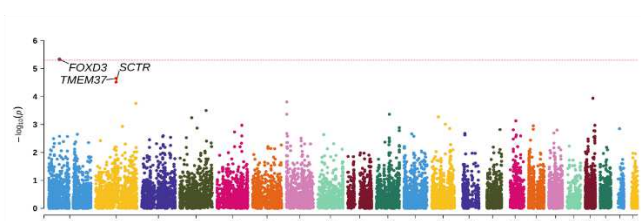
338 In CTLC's single marker SNP*PAE interaction analyses of child data, one SNP approaching
339 GWAS significance (*rs2833924*, *p*=8.8x10⁻⁸) was associated with working memory performances,
340 with several additional SNPs with marginal significance (*p* <10⁻⁵) reported in **Supplementary**
341 **Table S2.**

342 In single marker SNP*PAE interaction analyses of maternal data, we found one SNP approaching
343 GWAS significance (*rs17050152*, non-synonymous exonic variant in *SYNPO2*, *p*=7.7x10⁻⁷)

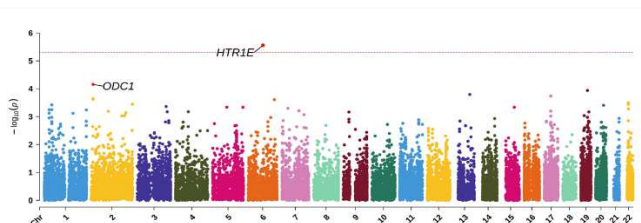
(A) Gene-based Genotype*PAE Analysis of Child Genotype data for Reduced Child Height (CTLC Cohort)



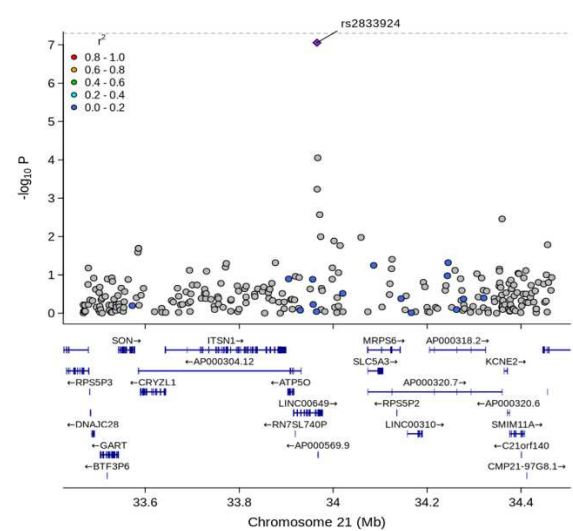
(B) Gene-based Genotype*PAE Analysis of Child Genotype data for Recognition Discrimination Performance (CTLC Cohort)



(C) Gene-based Genotype*PAE Meta Analysis of Child Genotype data for Working Memory Performance (CTLC and CIFASD Cohort)



(D) Regional Association Plot for rs2833924 SNP (CTLC Cohort)



344

345 **Figure 2:** Genome-wide association results (genotype data) for neurobehavioral and growth-related traits in CLTC data. (A) Gene-
346 based association analysis for child height-for-age z-scores adjusted for postnatal covariates and Heavy_YN exposure ($\lambda = 1.075$).
347 The Manhattan plot highlights one genome-wide significant gene (*STX6*). (B) Gene-based association analysis for child recognition
348 memory adjusted for neurobehavioral covariates and Heavy_YN exposure ($\lambda = 1.05$). Three genes (*FOXD3*, *TMEM37*, *CTR*) show
349 marginal associations. (C) Gene-based meta-analysis results for child working memory in CTLC and CIFASD. (D) Regional
350 Association plot for approaching genome-wide significant rs2833924 SNP in working memory performance. The Manhattan plot
351 highlights *HTR1E* gene as genome-wide significant. The QQ plots for all these Manhattan plots are shown in **Supplementary**
352 **Figure S2**.

353 associated with working memory (additional suggestive ($p < 10^{-5}$) results in **Supplementary Table**
354 **S2**).

355 Top findings from gene-based gene-by-PAE analyses of child data identified a significant locus,
356 *STX6* ($p = 2.2 \times 10^{-6}$) associated with reduced child height (**Figure 2A**) and a second locus
357 significant at trend, *FOXD3* ($p = 4.7 \times 10^{-6}$), associated with recognition memory performance
358 (**Figure 2B**). No significant loci were identified in mothers. Suggestive genes ($p < 10^{-4}$) for working

memory, recognition memory, and FAS diagnosis in child and, for working memory, recognition memory and height, in mothers are reported in **Supplementary Table S3**.

361

3.3 Genetic main effects on FAS/PFAS diagnosis (exposed only) examining child and maternal genetic data separately

No significant loci were observed in single-marker nor gene-based analyses in either maternal or child strata for FAS/PFAS diagnosis. However, four genes showed marginal associations (**Supplementary Table S3**). Similarly, in the CIFASD dataset, no SNP or gene reached genome-wide significance, with six single markers and six genes showing marginal significance (**Supplementary Table S4 and S5**).

369

3.4 Meta-analysis of Single-Marker and Gene-Based Association Results (child genetic data only)

Meta-analysis of the CTLC and CIFASD cohorts single-marker association using the Han and Eskin random-effects (RE2) model identified several loci showing suggestive evidence of association across multiple neurocognitive and FASD-related phenotypes. For working memory, the most significant association was observed at locus rs12090855 (chr1:28874969 HG19) ($P_{RE2}=1.38 \times 10^{-6}$), with additional suggestive signals at loci rs2125398 (chr4:75208519) and rs1990171 (chr7:86885119) ($P_{RE2} < 1 \times 10^{-5}$). For recognition memory, the strongest association signal was detected at locus rs1048274 (chr9:6256292) ($P_{RE2}=1.65 \times 10^{-6}$), with additional consistent signals at loci including rs4349637 (chr4:182065498) and rs12805702 (chr11:25117751) ($P_{RE2} < 1 \times 10^{-5}$ across multiple recognition memory model definitions. Here, P_{RE2} is the p-value derived from the Han and Eskin random-effects (RE2) meta-analysis model,

381 which is designed to maintain statistical power under heterogeneous effect sizes across cohorts
382 while accounting for between-study heterogeneity.

383 For FASD-related diagnostic and postnatal phenotypes, the top association signals were
384 observed at loci rs10105040 (chr8:94711093), rs7866848 (chr9:97505608), and rs946488
385 (chr1:215415487), with P_{RE2} values ranging from 7.65×10^{-6} to 1.20×10^{-5} . Notably, these loci
386 were consistently identified in both neurobehavioral diagnostic and postnatal height analyses,
387 suggesting shared genetic influences across FASD-related phenotypes. Finally, no genome-wide
388 results were observed in single-marker analyses. The meta-analysis of gene-based association
389 results found a significant genome-wide association in the *HTR1E* gene associated with working
390 memory performance ($p=2.75 \times 10^{-6}$) (**Figure 2C**).

391

392 3.5 Haplin Transmission Disequilibrium Test

393 In CTLC, gene-by-PAE HTDT test found no significant maternal or child haplotypes after multiple
394 testing correction for neurobehavior and growth outcomes. For FAS/PFAS diagnosis, no
395 significant main effect was found.

396 We followed up multiple nominally significant haplotypes from the genotype array data
397 (**Supplementary Table S6, S7**) employing imputed data: we found three haplotypes showing
398 significant main effects after frequency weighted FDR correction on FAS/PFAS diagnosis (**Table**
399 **2**). A haplotype in children lying within the *COP1* gene ($p_{adj}=0.0295$), and two maternal haplotypes
400 lying within the *WWOX* gene ($p_{adj}=0.0002$) and within an intergenic region between the
401 *TMEM38B* and *MIR8081* genes ($p_{adj}=0.0192$). The risk ratios associated with the significant
402 haplotypes in offspring and mother are shown in **Figure 2**. Comparison of effect sizes of the
403 significant imputed haplotypes in child and mother showed the opposite direction in the confidence
404 intervals. This regional enrichment of nominally associated haplotypes suggests that these loci

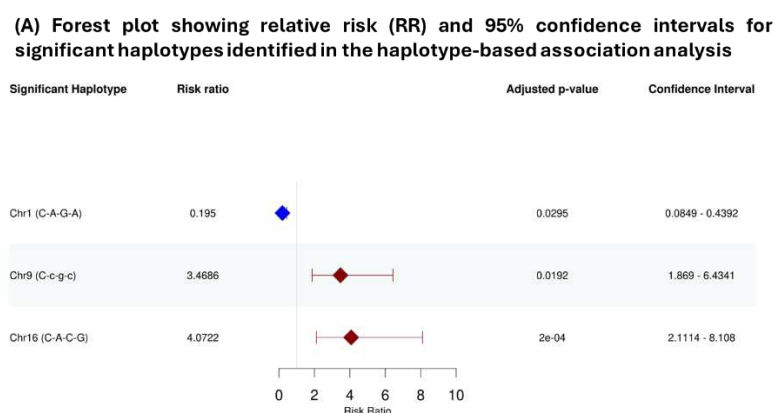
405 **Table 2.** Genetic haplotypes with main effects on FAS/PFAS diagnoses (imputed data)

Haplotype	OUTCOME	Risk Ratio (95% CI)	Haplotype Frequency	adjusted <i>p</i>	Function	Gene
CHILD						
chr1:175992527:T:C chr1:175992552:G:A chr1:175992588:G:C chr1:175992852:A:G	FAS/PFAS	0.19 (0.08–0.44)	0.22	0.0295	Intronic	<i>COP1</i>
MOTHERS						
chr9:109190024:C:T chr9:109190198:G:C chr9:109190646:C:G chr9:109191014:G:C	FAS/PFAS	3.47 (1.87–6.43)	0.12	0.0192	Intergenic	<i>TMEM38B</i> ; <i>MIR8081</i>
chr16:78493391:T:C chr16:78493396:T:A chr16:78493458:C:G chr16:78493612:A:G		4.07 (2.11–8.11)	0.1170	0.0002	Intronic	<i>WWOX</i>

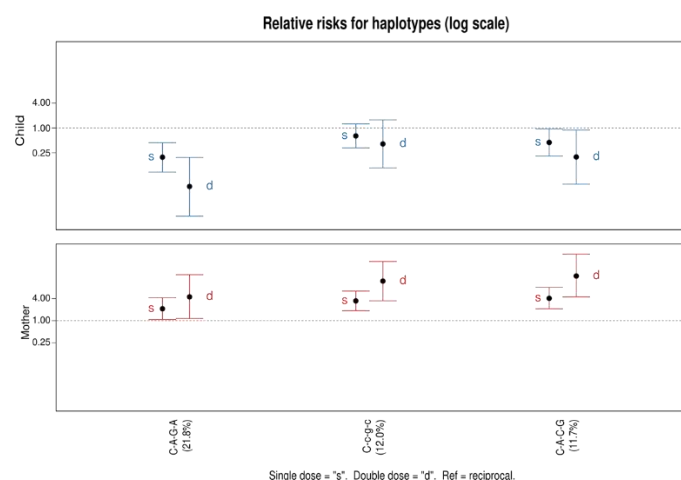
406 ^a Haplin Transmission Disequilibrium Test

407 harbor underlying association signals that were more effectively captured after imputation due to
408 increased marker density and improved haplotype resolution.

409 **Figure 3** presents the forest plot of relative risk (RR) and 95% confidence intervals for the
410 haplotypes that remained significant after multiple testing correction in the imputed dataset. The
411 Chr1 (C-A-G-A) haplotype showed a protective effect in the offspring genome (RR = 0.19, 95%



(B) Comparison of effect estimates for significant imputed haplotypes and nearby nominally associated haplotypes identified in the genotype dataset across maternal and child genomes.



412

413 **Figure 3:** Haplotype association results and relative risk (RR) with 95% confidence interval (CI) for FAS/PFAS diagnosis at Chr1,
 414 Chr9, and Chr16 loci. The top panel shows a forest plot of RR and 95% CI for haplotypes that remained significant after multiple
 415 testing correction in the imputed dataset. The bottom panel compares the RR of these significant haplotypes in mother and
 416 offspring for single (s) and double dose (d). Blue indicates offspring (child) haplotypes and red indicates maternal haplotypes.
 417 Error bars represent 95% confidence intervals.

418

419 CI: 0.08–0.44, adjusted p = 0.029), whereas the Chr9 (C-c-g-c) and Chr16 (C-A-C-G) haplotypes
 420 showed significant risk effects in the maternal genome, with relative risks of 3.47 (95% CI: 1.87–
 421 6.43, adjusted p = 0.019) and 4.07 (95% CI: 2.11–8.11, adjusted p = 2×10^{-4}), respectively. These
 422 results indicate distinct maternal and offspring haplotype effects at different loci, suggesting that

both maternal genetic background and offspring genotype may contribute to FASD-related phenotypes.

425

3.6 Differential gene expression results

Table 3 displays the results of differential gene expression analyses based on PAE status (exposed vs. unexposed) in CTLC for risk-modifying genes described above, with adjustment for tissue-specific cell-type composition and technical confounders (using SVA). The cell type distribution for placenta and child blood is shown in **Supplementary Figure S3**. *HTR1E*, implicated in FAS/PFAS diagnosis, and *FOXD3*, implicated in recognition memory, were not expressed in placenta or child blood. *STX6*, which had evidence of gene*PAE interaction effects on child height, showed increased expression in child blood. For FAS/PFAS diagnoses, *COP1*, which had evidence of child genetic main HTDT effects, showed no differential expression. Two genes with evidence of maternal genetic HTDT main effects exhibited notable patterns: *TMEM38B* showed reduced expression in child blood, and *WWOX* demonstrated decreased expression in both placenta and child blood. Collectively, these transcriptomic findings provide functional support for several genetic signals and highlight tissue-specific regulatory patterns that may contribute to fetal vulnerability to prenatal alcohol exposure.

440

4. DISCUSSION

Integrating child and maternal genetic data from two prospective longitudinal birth cohorts and child genetic data from a clinical cohort of exposed and unexposed children, we identified several novel genetic loci that may modify the risk of FASD. To our knowledge, this is the first FASD study to utilize genome-wide approaches and the first to integrate maternal and child genetics.

446 **Table 3.** Differential gene expression^a based on alcohol exposure (yes vs. no) for potential risk-
 447 modifying genes

			Placenta		Child blood	
Outcome	Gene prioritized by GWAS	Strata	logFC	<i>p</i>	logFC	<i>p</i>
Height	<i>STX6</i>	Child	0.3124	0.2131	0.7480	0.0230
FAS/PFAS	<i>COP1</i>	Child	-0.00549	0.9761	0.2641	0.4729
	<i>TMEM38B</i>	Mother	-0.9751	0.211	-1.397	0.0129
	<i>WWOX</i>	Mother	-0.7567	0.0358	-0.8289	0.0500

448 N.E. = not expressed.

449 ^aDifferential gene expression expressed as log(fold-change) [FC]

450

451 We employed single-marker and gene-based approaches and the Haplin Transmission
 452 Disequilibrium Test. Our gene-based and Haplin analyses were substantially higher yield than
 453 single-marker analyses in identifying potential risk-modifying child and maternal loci, likely due to
 454 the reduced multiple testing burden in gene-based analyses; genome-wide screens and gene-by-
 455 environment interaction analyses are both very power intensive. It should be noted that a large
 456 number of loci were found to have marginal significance in both single marker and gene-based
 457 analyses, greatly exceeding standard power estimations. This yield likely reflects the high PAE
 458 levels and unique genetic architecture in our cohorts and the potent teratogenicity of alcohol.

459 The key strength of the Haplin Transmission Disequilibrium Test over standard approaches
 460 is that it examines maternal and child genetics simultaneously, thereby disentangling the shared
 461 genetics between mother and child. The multiple-testing burden of a haplotype approach (2-3
 462 million haplotypes), and use of binary “case” variables rather than continuous quantitative traits

463 make this approach especially power intensive. Furthermore, as in gene-based approaches,
464 strong gene-by-environment interaction signals may be diluted and thus obscured by markers
465 with weaker effect sizes in the same haplotype/region. Notably, our HTDT examination of main
466 effects of genotype on risk for child FAS/PFAS diagnosis among exposed children revealed one
467 child locus and two maternal loci. The precision of FAS/PFAS diagnostic criteria and the specificity
468 of those diagnoses to PAE likely enhanced power in these analyses.

469 The detection of significant haplotypes in the imputed dataset but not in the genotype-only
470 dataset is likely attributable to the increased marker density provided by imputation, which
471 improves haplotype reconstruction and enables the capture of untyped variants in linkage
472 disequilibrium with causal loci. Consistent with this, we observed multiple nominally significant
473 haplotypes in the surrounding genomic regions in the genotype dataset, suggesting the presence
474 of underlying regional association signals that became statistically significant only after imputation
475 increased genomic coverage and refined haplotype structure. The significant haplotype signals
476 detected for FAS/PFAS in the imputed CTLC HTDT analysis likely arose because the informative
477 SNPs defining these haplotypes were not present in the directly genotyped CTLC dataset and
478 therefore became analyzable only after imputation. These SNPs and haplotypes were also absent
479 in the CIFASD dataset, limiting replication across cohorts. Additional differences between findings
480 from directly genotyped and imputed datasets may reflect differences in variant coverage, linkage
481 disequilibrium structure, and ancestry composition across cohorts, particularly within African and
482 admixed populations where haplotype diversity is substantial. Such factors can influence the
483 transferability and consistency of association signals across datasets.

484 In contrast, analyses based on directly genotyped data produced more consistent and
485 interpretable findings across most phenotypes, suggesting that imputation quality and population
486 mismatch were important limiting factors in the imputed dataset. These findings highlight the need

487 for improved population-specific reference panels and optimized imputation strategies when
488 analyzing genetically diverse cohorts such as CTLC.

489 The child and maternal genetic loci identified in our analyses have remarkable biologic
490 plausibility for contributing to FASD pathophysiology. Child markers included rs2833924, which is
491 in the 'Down syndrome critical region' in a binding site for several transcription factors near
492 cytokine genes (*IFNA*, *IFNB*, *IL10RB*) relevant for neuroinflammation, which is implicated in
493 FASD; *RUNX1*, which helps regulate hematopoiesis and neurodevelopment [48]; and several
494 genes involved in mitochondrial function, which has also been implicated in FASD [49]. *STX6*
495 mutations are associated with CEDNIK Syndrome, in which children exhibit developmental delay,
496 absence of the corpus callosum, hypomyelination, and white matter abnormalities [50], all of which
497 we have previously demonstrated in FASD [23, 51, 52]. Notably, in our data, *STX6* blood
498 expression was increased among exposed vs. control children. *FOXD3* is down-regulated in
499 response to PAE in animal models and plays critical roles in differentiation of neural crest cells, a
500 key target in alcohol teratogenesis [53, 54]. *HTR1E* encodes the 5-hydroxytryptamine-1E G-
501 protein-coupled serotonin receptor, which is highly expressed in the brain and promotes neuronal
502 survival against oxidative and neurotoxic stress [55]. Maternal genetic loci included *COP1*, which
503 encodes a ubiquitin ligase that suppresses neuroinflammation, especially in microglia [56].
504 Notably, in our data, *COP1* blood expression was increased among children with FAS or PFAS
505 vs. non-syndromal children. Maternal *rs17050152* approached significance and is in *SYNPO2*, a
506 tumor suppressor gene that positively regulates actin filament bundle assembly and cell migration.
507 Alterations in *SYNPO2* expression have been associated with ethanol exposure in adult rodent
508 models and with alcohol use in humans [57, 58]. *TMEM38B* encodes a calcium channel important
509 for Type I collagen synthesis and is maternal blood biomarker for disrupted placental
510 development/preeclampsia [59]. *WWOX* is tumor suppressor gene that is important for placental
511 glucose control/metabolism in mother and neurodevelopment in child [60]; mutations lead to

512 microcephaly, epilepsy, developmental delay, early death [61]. In our data, placental *WWOX*
513 expression was increased in exposed vs. control children. It should be noted that genes not found
514 to be differentially expressed based on PAE or phenotype may still play important roles in the
515 pathophysiology of FASD; they may be differentially expressed in other tissues or in specific
516 developmental time windows not captured in our study design.

517 A limitation of these analyses is the small sample size for genome-wide approaches. To
518 address this limitation, we employed gene-based approaches to reduce the multiple testing
519 burden. Power boosting strengths of our South African cohorts include the very heavy PAE levels
520 and the unique genetic architecture of participants, who identify as South African Cape Coloured.
521 In data from these cohorts, 29,847 monomorphic variants in the gnomAD European reference
522 population [62] were polymorphic in our cohorts, and 90% of these were *common* ($\geq 1\%$ frequency)
523 in our cohorts. We have also demonstrated a high proportion ($>2/3$) of non-synonymous (i.e.,
524 functionally relevant) exonic variants, unusually high homozygosity (6%), and high proportions of
525 homozygous genetic segments that ancestrally harbor a higher burden of deleterious variants
526 [63]. Future studies are needed to validate/replicate our findings in larger cohorts, with the caveat
527 that most existing large cohorts have lower levels of PAE than our South African cohorts.
528 Differences in findings between our South African cohort and the CIFASD database may be due,
529 in part, to cohort differences such as allele frequencies that vary by geographic ancestry. However,
530 gene-based analyses are less dependent on allele frequencies of specific variants within a gene
531 that may vary across populations. Another possible source of these differences in findings is the
532 different neurobehavioral measures utilized across datasets. Future studies are needed that
533 incorporate multiple domain-specific outcome assessments into latent variables that are
534 comparable across cohorts. Lastly, the MEGA^{ex} array does not provide complete genome
535 coverage, and as noted above, imputation in the South African Cape Coloured population is prone
536 to errors. Thus, our analyses may have missed important genetic variants that were not assayed,

537 such as rare variants that often have large effect sizes. Future studies are needed that utilize
538 whole genome sequencing to capture such variants.

539

540 **5. CONCLUSIONS**

541 In summary, in this genome-wide screen for risk modifying maternal and child genetic loci, we
542 identified several putative genes that may have physiologic significance in the teratogenic effects
543 of alcohol and may inform future work to identify specific markers that identify pregnancies and/or
544 children at risk for FASD. Our findings also highlight the benefits of utilizing multi-tissue
545 transcriptomics in genetics research. Furthermore, our findings support the need for genome-wide
546 approaches in FASD research and underscore the importance of examining both maternal and
547 child genetics when studying FASD and other developmental diseases.

548

549 **Acknowledgments:**

550 We thank our UCT and Wayne State University research staff including Maggie September,
551 Beverley Arendse, Patricia O’Leary, and Patricia Solomon for their work on subject recruitment
552 and maintenance; and Renee Sun, for her assistance with data processing. We thank Catherine
553 O’Leary, Nadine Lindinger, and Natalia Berghoff for their work on child neurobehavioral
554 assessments. We also thank the dysmorphologists Luther Robinson, and Nathaniel Khaole for
555 their work in examining the children. We thank Yula Ma (Mt. Sinai) and Qian Li (Columbia
556 University) for their work on DNA isolation. We also thank Susan Fawcus, MD, Head of
557 Department of Obstetrics, Mowbray Maternity Hospital; the nursing and records department staff
558 at the Hanover Park and Retreat Midwife Obstetric Units, Mowbray Maternity Hospital, Somerset
559 Hospital, and Groote Schuur Hospital, where the mothers were recruited and the infants were

560 born. Lastly, we extend our deep appreciation to the Cape Town study participants for their
561 contributions to this study.

562 Data used in the preparation of this article were obtained from CIFASD consortium
563 (<https://cifasd.org/data-sharing/>; <https://doi.org/10.5967/ntw9-h991>), supported by the National
564 Institute of Alcohol Abuse and Alcoholism (NIAAA). This manuscript reflects the views of the
565 authors and may not reflect the opinions or views of NIAAA or CIFASD consortium investigators.
566 This research was funded by NIH/NIAAA (R01AA016781, R01AA027916, R01AA023503,
567 K23AA020516, U01AA014790), NIH/NIA (U01AG081817, RF1AG082009, U19AG074865), and
568 the Lycaki-Young Fund (State of Michigan).

569

570 **Conflicts of Interest:** The authors declare no conflict of interest.

571

572 Supplementary information is available at MP's website.

573

574

575

576

577

578

579

580

581

582 **References**

- 583 1. May PA, Gossage JP, Kalberg WO, Robinson LK, Buckley D, Manning M, et al. Prevalence and
584 epidemiologic characteristics of FASD from various research methods with an emphasis on
585 recent in-school studies. *Dev Disabil Res Rev.* 2009;15:176–192.
- 586 2. May PA, Marais AS, De Vries MM, Buckley D, Kalberg WO, Hasken JM, et al. The prevalence,
587 child characteristics, and maternal risk factors for the continuum of fetal alcohol spectrum
588 disorders: A sixth population-based study in the same South African community. *Drug Alcohol*
589 *Depend.* 2021;218:108408.
- 590 3. Gosdin LK, Deputy NP, Kim SY, Dang EP, Denny CH. Alcohol Consumption and Binge Drinking
591 During Pregnancy Among Adults Aged 18-49 Years - United States, 2018-2020. *MMWR Morb*
592 *Mortal Wkly Rep.* 2022;71:10–13.
- 593 4. Colin Carter R, Jacobson JL, Molteno CD, Jiang H, Meintjes EM, Jacobson SW, et al. Effects of
594 Heavy Prenatal Alcohol Exposure and Iron Deficiency Anemia on Child Growth and Body
595 Composition through Age 9 Years. *Alcohol Clin Exp Res.* 2012;36:1973–1982.
- 596 5. Carter RC, Wainwright H, Molteno CD, Georgieff MK, Dodge NC, Warton F, et al. Alcohol,
597 Methamphetamine, and Marijuana Exposure Have Distinct Effects on the Human Placenta.
598 *Alcohol Clin Exp Res.* 2016;40:753–764.
- 599 6. Hoyme HE, May PA, Kalberg WO, Kodituwakku P, Gossage JP, Trujillo PM, et al. A practical
600 clinical approach to diagnosis of fetal alcohol spectrum disorders: clarification of the 1996
601 Institute of Medicine criteria. *Pediatrics.* 2005;115:39–47.
- 602 7. Jacobson SW. Specificity of neurobehavioral outcomes associated with prenatal alcohol
603 exposure. *Alcohol Clin Exp Res.* 1998;22:313–320.
- 604 8. Carter RC, Jacobson JL, Molteno CD, Dodge NC, Meintjes EM, Jacobson SW. Fetal alcohol
605 growth restriction and cognitive impairment. *Pediatrics.* 2016;138.
- 606 9. Popova S, Lange S, Probst C, Gmel G, Rehm J. Estimation of national, regional, and global
607 prevalence of alcohol use during pregnancy and fetal alcohol syndrome: a systematic review
608 and meta-analysis. *Lancet Glob Health.* 2017;5:e290–e299.
- 609 10. Sambo D, Goldman D. Genetic Influences on Fetal Alcohol Spectrum Disorder. *Genes (Basel).*
610 2023;14.
- 611 11. Eberhart JK, Parnell SE. The Genetics of Fetal Alcohol Spectrum Disorders (FASD). *Alcohol Clin*
612 *Exp Res;*40.
- 613 12. Carter RC, Yang Z, Akkaya-Hocagil T, Jacobson SW, Jacobson JL, Dodge NC, et al. Genetic
614 admixture predictors of fetal alcohol spectrum disorders (FASD) in a South African population.
615 *Gene.* 2024;931.

13. Das UG, Cronk CE, Martier SS, Simpson PM, McCarver DG. Alcohol dehydrogenase 2*3 affects alterations in offspring facial morphology associated with maternal ethanol intake in pregnancy. *Alcohol Clin Exp Res*. 2004;28:1598–1606.
14. Dodge NC, Jacobson JL, Jacobson SW. Protective effects of the alcohol dehydrogenase-ADH1B*3 allele on attention and behavior problems in adolescents exposed to alcohol during pregnancy. *Neurotoxicol Teratol*. 2014;41:43–50.
15. Jacobson SW, Carr LG, Croxford J, Sokol RJ, Li TK, Jacobson JL. Protective effects of the alcohol dehydrogenase-ADH1B allele in children exposed to alcohol during pregnancy. *J Pediatr*. 2006;148:30–37.
16. Mccarver DG, Thomasson HR, Martier SS, Sokol RJ, Li TK. Alcohol dehydrogenase-2*3 allele protects against alcohol-related birth defects among African Americans. *Journal of Pharmacology and Experimental Therapeutics*. 1997;283:1095–1101.
17. Viljoen DL, Carr LG, Foroud TM, Brooke L, Ramsay M, Li TK. Alcohol dehydrogenase-2*2 allele is associated with decreased prevalence of fetal alcohol syndrome in the mixed-ancestry population of the Western Cape Province, South Africa. *Alcohol Clin Exp Res*. 2001;25:1719–1722.
18. Smith SM, Virdee MS, Eckerle JK, Sandness KE, Georgieff MK, Boys CJ, et al. Polymorphisms in SLC44A1 are associated with cognitive improvement in children diagnosed with fetal alcohol spectrum disorder: an exploratory study of oral choline supplementation. *Am J Clin Nutr*. 2021;114:617–627.
19. McKay L, Petrelli B, Pind M, Reynolds JN, Wintle RF, Chudley AE, et al. Risk and Resilience Variants in the Retinoic Acid Metabolic and Developmental Pathways Associated with Risk of FASD Outcomes. *Biomolecules*. 2024;14.
20. De La Morena-Barrio ME, Ballesta-Martínez MJ, López-Gálvez R, Antón AI, López-González V, Martínez-Ribot L, et al. Genetic predisposition to fetal alcohol syndrome: association with congenital disorders of N-glycosylation. *Pediatric Research* 2018 83:1. 2017;83:119–127.
21. Carter RC. The critical need for novel precision medicine methodologies in fetal alcohol spectrum disorders. *American Journal of Clinical Nutrition*. 2024;119:5–6.
22. Jacobson SW, Stanton ME, Molteno CD, Burden MJ, Fuller DS, Hoyme HE, et al. Impaired eyeblink conditioning in children with fetal alcohol syndrome. *Alcohol Clin Exp Res*. 2008;32:365–372.
23. Jacobson SW, Jacobson JL, Molteno CD, Warton CMR, Wintermark P, Hoyme HE, et al. Heavy prenatal alcohol exposure is related to smaller corpus callosum in newborn MRI scans. *Alcohol Clin Exp Res*. 2017;41:965–975.

24. Carter RC, Senekal M, Dodge NC, Bechard LJ, Meintjes EM, Molteno CD, et al. Maternal Alcohol Use and Nutrition During Pregnancy: Diet and Anthropometry. *Alcohol Clin Exp Res*. 2017;41:2114–2127.
25. Jacobson SW, Chiodo LM, Sokol RJ, Jacobson JL. Validity of maternal report of prenatal alcohol, cocaine, and smoking in relation to neurobehavioral outcome. *Pediatrics*. 2002;109:815–825.
26. Sokol RJ, Martier S, Ernhart C. Identification of alcohol abuse in the prenatal clinic. In: Chang NC, Chao M, editors. *Early Identification of Alcohol Abuse*, vol. Monograph No. 17, Rockville, MD: Alcohol, Drug Abuse, and Mental Health Administration Research; 1983.
27. Hollingshead AB. Four factor index of social status. *Yale J Sociol*. 2011;8:21–51.
28. Robert J Kuczmarski, Cynthia L Ogden, Shumei S Guo, Laurence M Grummer-Strawn, Katherine M Flegal, Zuguo Mei, et al. 2000 CDC Growth Charts for the United States: methods and development - PubMed. *Vital Health Stat 11*. 2002;246:1–190.
29. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26:2867–2873.
30. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*;4.
31. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–575.
32. Millard SP. *EnvStats: An R Package for Environmental Statistics*. Springer; 2013.
33. Bramwell M. The future of data analysis. *Scientific Computing and Instrumentation*. 2001;18:20–22.
34. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. *Nature Genetics* 2016 48:10;48.
35. Mattson SN, Foroud T, Sowell ER, Jones KL, Coles CD, Fagerlund A, et al. Collaborative initiative on fetal alcohol spectrum disorders: methodology of clinical projects. *Alcohol*. 2010;44:635–641.
36. Dou X, Menkari C, Mitsuyama R, Foroud T, Wetherill L, Hammond P, et al. L1 coupling to ankyrin and the spectrin-actin cytoskeleton modulates ethanol inhibition of L1 adhesion and ethanol teratogenesis. *FASEB J*. 2018;32:1364–1374.
37. Schwantes-An TH, Darlay R, Mathurin P, Masson S, Liangpunsakul S, Mueller S, et al. Genome-wide Association Study and Meta-analysis on Alcohol-Associated Liver Cirrhosis Identifies Genetic Risk Factors. *Hepatology*. 2021;73:1920–1931.

683 38. Fairley S, Lowy-Gallego E, Perry E, Flicek P. The International Genome Sample Resource (IGSR)
684 collection of open human genomic variation resources. *Nucleic Acids Res.* 2020;48:D941–
685 D947.

686 39. Team R. R: A language and environment for statistical computing. *MSOR Connections.* 2014.
687 2014.

688 40. Zhou X, Stephens M, Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for
689 association studies. *Nature Genetics* 2012 44:7;44.

690 41. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A Tool for Genome-wide Complex Trait
691 Analysis. *Am J Hum Genet*;88.

692 42. Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of
693 genome-wide association studies. *Am J Hum Genet.* 2011;88:586–598.

694 43. Gjessing HK. Haplin: Analyzing case-parent triad and/or case-control data with SNP
695 haplotypes. *R Package Version.* 2010;3.

696 44. Gjerdevik M, Jugessur A, Haaland OA, Romanowska J, Lie RT, Cordell HJ, et al. Haplin power
697 analysis: A software module for power and sample size calculations in genetic association
698 analyses of family triads and unrelated controls. *BMC Bioinformatics.* 2019;20.

699 45. Ignatiadis N, Klaus B, Zaugg JB, Huber W. Data-driven hypothesis weighting increases
700 detection power in genome-scale multiple testing. *Nature Methods* 2016 13:7. 2016;13:577–
701 580.

702 46. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq
703 data with DESeq2. *Genome Biol.* 2014;15:550.

704 47. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune
705 Cells with CIBERSORT. *Methods Mol Biol.* 2018;1711:243–259.

706 48. Rozen EJ, Ozeroff CD, Allen MA. RUN(X) out of blood: emerging RUNX1 functions beyond
707 hematopoiesis and links to Down syndrome. *Human Genomics* 2023 17:1. 2023;17:1–20.

708 49. Darbinian N, Darbinyan A, Merabova N, Kassem M, Tatevosian G, Amini S, et al. In utero ethanol
709 exposure induces mitochondrial DNA damage and inhibits mtDNA repair in developing brain.
710 *Front Neurosci.* 2023;17:1214958.

711 50. Vinayagamani S, Sekar S, Thomas B, Kesavadas C. Neuroimaging in CEDNIK Syndrome: A Rare
712 Neuro-Ichthyosis. *Neurol India.* 2022;70:818–819.

713 51. Spottiswoode BS, Meintjes EM, Anderson AW, Molteno CD, Stanton ME, Dodge NC, et al.
714 Diffusion Tensor Imaging of the Cerebellum and Eyeblink Conditioning in Fetal Alcohol
715 Spectrum Disorder. *Alcohol Clin Exp Res.* 2011;35:2174.

716 52. Fan J, Jacobson SW, Taylor PA, Molteno CD, Dodge NC, Stanton ME, et al. White matter deficits
717 mediate effects of prenatal alcohol exposure on cognitive development in childhood. *Hum*
718 *Brain Mapp.* 2016;37:2943.

719 53. Chen S yu, Kannan M. Neural crest cells and fetal alcohol spectrum disorders: Mechanisms
720 and potential targets for prevention. *Pharmacol Res.* 2023;194.

721 54. Zhang P, Wang G, Lin Z, Wu Y, Zhang J, Liu M, et al. Alcohol exposure induces chick craniofacial
722 bone defects by negatively affecting cranial neural crest development. *Toxicol Lett.*
723 2017;281:53–64.

724 55. Sharma VK, Loh YP. The discovery, structure, and function of 5-HTR1E serotonin receptor. *Cell*
725 *Commun Signal.* 2023;21:235.

726 56. Ndoja A, Reja R, Lee SH, Webster JD, Ngu H, Rose CM, et al. Ubiquitin Ligase COP1 Suppresses
727 Neuroinflammation by Degrading c/EBP β in Microglia. *Cell.* 2020;182:1156-1169.e12.

728 57. McClintick JN, McBride WJ, Bell RL, Ding ZM, Liu Y, Xuei X, et al. Gene expression changes in
729 the ventral hippocampus and medial prefrontal cortex of adolescent alcohol-preferring (P) rats
730 following binge-like alcohol drinking. *Alcohol.* 2018;68:37–47.

731 58. Baker JA, Li J, Zhou D, Yang M, Cook MN, Jones BC, et al. Analyses of differentially expressed
732 genes after exposure to acute stress, acute ethanol, or a combination of both in mice. *Alcohol.*
733 2017;58:139–151.

734 59. Sufriyana H, Salim HM, Muhammad AR, Wu YW, Su ECY. Blood biomarkers representing
735 maternal-fetal interface tissues used to predict early-and late-onset preeclampsia but not
736 COVID-19 infection. *Comput Struct Biotechnol J.* 2022;20:4206–4224.

737 60. Iacomino M, Baldassari S, Tochigi Y, Kośła K, Buffelli F, Torella A, et al. Loss of Wwox Perturbs
738 Neuronal Migration and Impairs Early Cortical Development. *Front Neurosci.* 2020;14:540673.

739 61. Dong X sheng, Wen X jun, Zhang S, Wang D gang, Xiong Y, Li Z ming. Identification of compound
740 heterozygous deletion of the WWOX gene in WOREE syndrome. *BMC Medical Genomics* 2023
741 16:1. 2023;16:1–7.

742 62. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational
743 constraint spectrum quantified from variation in 141,456 humans. *Nature.* 2020;581:434–443.

744 63. Szpiech ZA, Mak ACY, White MJ, Hu D, Eng C, Burchard EG, et al. Ancestry-Dependent
745 Enrichment of Deleterious Homozygotes in Runs of Homozygosity. *The American Journal of*
746 *Human Genetics.* 2019;105:747–762.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [GeneticInfluencesFASDMolPsychSuppMatJune052026.pdf](#)
- [GeneticInfluencesFASDMolPsychSuppMatJune102026.pdf](#)