

1 **Title:** Cadmium-associated differential methylation throughout the placental genome: epigenome-wide
2 association study of two US birth cohorts

3
4 **Authors:**

5 Todd M. Everson^a

6 Tracy Punshon^b

7 Brian P. Jackson^c

8 Ke Hao^d

9 Luca Lambertini^e

10 Jia Chen^e

11 Margaret R. Karagas^{f,g}

12 Carmen J. Marsit^{a,g}

13
14 **Author Details:**

15 ^a Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA,
16 USA

17 ^b Department of Biological Sciences, Dartmouth College, Hanover, NH, USA

18 ^c Department of Earth Sciences, Dartmouth College, Hanover, NH, USA

19 ^d Department of Genome Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA

20 ^e Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai,
21 New York, NY USA

22 ^a Department of Epidemiology, Geisel School of Medicine, Dartmouth College, Lebanon, NH, USA

23 ^f Children's Environmental Health and Disease Prevention Research Center at Dartmouth Geisel School
24 of Medicine, Lebanon, NH, 03756 USA

25
26 ***Corresponding Author:**

27 Carmen J. Marsit, Rollins School of Public Health at Emory University, 1518 Clifton Road, Claudia
28 Nance Rollins Room 2021, Atlanta, GA 30322; Carmen.J.Marsit@emory.edu; (404)-712-8912.

29
30 **Competing Financial Interests:**

31 The authors have no competing financial interests to declare.

32
33 **Significance:**

34 Cadmium is a toxic environmental pollutant that can impair fetal development. The mechanisms
35 underlying this toxicity are unclear, though disrupted placental functions could play an important role. In
36 this study we examined associations between cadmium concentrations and DNA methylation throughout
37 the placental genome, across two US birth cohorts. We observed cadmium-associated differential
38 methylation, and corresponding methylation-expression associations at genes involved in cellular growth
39 processes and/or immune and inflammatory signaling. This study provides supporting evidence that
40 disrupted placental epigenetic regulation of cellular growth and immune/inflammatory signaling could
41 play a role in cadmium associated reproductive toxicity in human pregnancies.

42 **Abstract (250 word-limit)**

43 **Background:** Cadmium (Cd) is a ubiquitous toxicant that during pregnancy can impair fetal
44 development. Cd sequesters in the placenta where it can impair placental function, impacting fetal
45 development. We aimed to investigate Cd-associated variations in placental DNA methylation (DNAM),
46 associations with gene expression, and identify novel pathways involved in Cd-associated reproductive
47 toxicity.

48 **Methods:** Using placental DNAM and Cd concentrations in the New Hampshire Birth Cohort Study
49 (NHBCS, n=343) and the Rhode Island Child Health Study (RICHHS, n=141), we performed an EWAS
50 between Cd and DNAM, adjusting for tissue heterogeneity using a reference-free method. Cohort-specific
51 results were aggregated via inverse variance weighted fixed effects meta-analysis, and variably
52 methylated CpGs were associated with gene expression. We then performed functional enrichment
53 analysis and tests for associations between gene expression and birth metrics.

54 **Results:** We identified 17 Cd-associated differentially methylated CpG sites with meta-analysis p-values
55 $< 1e-05$, two of which were within a 5% false discovery rate (FDR). Methylation levels at 9 of the 17 loci
56 were associated with increased expression of 6 genes (5% FDR): *TNFAIP2*, *EXOC3LA*, *GAS7*, *SREBF1*,
57 *ACOT7*, and *RORA*. Higher placental expression of *TNFAIP2* and *ACOT7*, and lower expression of
58 *RORA*, were associated with lower birth weight z-scores (p-values < 0.05).

59 **Conclusion:** Cd associated differential DNAM and corresponding DNAM-expression associations at
60 these loci are involved in inflammatory signaling and cell growth. The expression levels of genes
61 involved in inflammatory signaling (*TNFAIP2*, *ACOT7*, and *RORA*), were also associated with birth
62 metrics, suggesting a role for inflammatory processes in Cd-associated reproductive toxicity.

63

64 **Background:**

65 Cadmium (Cd) is a toxic heavy metal that is released into the environment from mining and
66 industrial processes, the application of phosphate fertilizers, and fossil fuel combustion; it accumulates in
67 soils and is readily taken up by plants (1). Human exposure primarily occurs via consumption of
68 contaminated foods, or through using tobacco products (2). Cadmium serves no biological role in
69 humans, and high concentrations can cause severe health consequences such as kidney damage, bone and
70 joint problems and various cancers (1). However, trace exposures to cadmium may also have toxic
71 effects, particularly to pregnant mothers and the developing fetus (3).

72 Recent epidemiologic studies have observed associations between Cd concentrations in maternal
73 and/or fetal tissues with restricted fetal growth and pregnancy complications (4–10). The toxic
74 mechanisms that underlie these associations are not clear, though Cd is known to induce oxidative stress,
75 interfere with cell-cycle regulation, and alter apoptotic signaling at the cellular level (11). Additionally,
76 Cd tends to sequester in the placenta throughout pregnancy, limiting the direct fetal exposure, though its
77 accumulation within the placenta is likely not without consequence (12). Animal models and *in vitro*
78 studies have suggested that maternal Cd exposure during pregnancy can perturb maternal-fetal nutrient
79 and waste transfer (13, 14), increase oxidative stress in placental tissue (15), interfere with placental
80 glucocorticoid synthesis (16), and alter trophoblast proliferation and apoptosis (17). Thus, the placenta
81 appears to be a critical target tissue for the toxic effects of Cd during human pregnancies.

82 We hypothesize that various Cd-associated perturbations to placental functions are likely linked
83 to the selective expression or repression of specific biological pathways. Likewise, maternal Cd exposure
84 during pregnancy has been associated with variations in DNA methylation (DNAM), an epigenetic
85 mechanism that regulates gene expression potential, in maternal and/or fetal blood samples (18). To date,
86 only one study has investigated associations between placental Cd and DNAM throughout the placental
87 genome in humans. Mohanty et al (2015) explored Cd-associated DNAM in a small study of 24
88 placentae, identifying multiple differentially methylated loci nearby or within genes involved in cell-
89 damage, angiogenesis, cell-differentiation, and organ development; these associations differed by fetal

90 sex (19). Studies with larger samples using recent methodological advances to estimate and adjust for
91 tissue heterogeneity are needed to understand how Cd exposure influences the placental epigenome of
92 human pregnancies, and whether such variations are related to fetal growth restriction.

93 We conducted a study of the associations between placental Cd and placental DNAM throughout
94 the epigenome across two well characterized independent cohorts while examining the potential impact of
95 tissue heterogeneity on these associations. Secondly, we investigated whether the top hits from our
96 analysis were enriched for biological pathways consistent with findings from other studies, whether
97 differentially methylated loci were associated with the expression levels of nearby genes, and whether
98 these variations in expression were associated with standardized measures of birth weight, birth length
99 and head circumference.

100

101 **Results:**

102 The EWAS for this study included 343 mother-infant pairs drawn from the New Hampshire Birth
103 Cohort (NHBCS) and 141 mother-infant pairs from the Rhode Island Child Health Study (RICHS). The
104 most notable differences between the NHBCS and RICHS samples ([Table 1](#)) were with racial diversity
105 and birth size. Because RICHS was over-sampled for LGA and SGA infants by design, we observed a
106 higher proportion of SGA and LGA infants compared to NHBCS. Placental samples from the RICHS
107 cohort also had slightly higher median Cd concentrations (4.34 ng/g) compared to NHBCS (3.13 ng/g),
108 and concentrations were approximately log-normal distributed in both cohorts ([Supplemental Figure 1](#)).

109 We regressed DNA methylation M-values (at 408,367 CpGs in NHBCS and 397,040 CpGs in
110 RICHS) on dichotomized placental Cd concentrations (median split) while adjusting for fetal sex,
111 maternal smoking during pregnancy, and maternal education level, in each cohort. We then combined
112 parameter estimates and p-values via inverse variance weighted fixed effects meta-analysis (390,583
113 CpGs overlapped between the cohorts) which yielded three loci with Cd-associations at a 5% FDR and 18
114 additional loci with p-value < 1e-05 ([Table 2](#)), in which Cd was predominantly associated with
115 hypermethylation. Full results from the meta-analysis are provided in the supplemental materials
116 ([Supplemental Table 1](#)).

117 Cell-specific DNAM is a well-recognized confounder in epigenetic epidemiology, leading many
118 studies to estimate the proportions of constituent cell-types from the DNAM array data (20), and adjust
119 for these proportions in their regression models. However, estimation of cell-type proportions requires
120 access to cell-type specific methylomes for the individual cells that constitute the heterogenous tissue
121 under study, and no reference methylomes currently exist for placental tissue. Thus we utilized a
122 reference free method that deconvoluted the major axes of variation in DNAM, which are likely
123 associated with tissue heterogeneity, to estimate the proportions of the putative constituent cell-types
124 (PCTs) in our placental samples (21). Via this method, we estimated five and four PCTs in NHBCS and
125 RICHS, respectively. We then tested whether the 10,000 CpGs with greatest variability in PCT-specific
126 methylation were enriched for biologically relevant GO terms or KEGG pathways, and checked for

127 consistency between the two cohorts. We observed 31.6% overlap among these 10,000 PCT-defining
128 CpGs, and strong concordance in the biological functions of the top 100 GO terms (60% overlap) and
129 KEGG pathways (76% overlap) across the two cohorts. The top 10 KEGG pathways from both cohorts
130 included metabolic pathways, neuroactive ligand-receptor interaction, axon guidance, calcium signaling,
131 olfactory transduction, cell adhesion molecules (CAMs), and cAMP signaling ([Supplemental Tables 2 &](#)
132 [3, Supplemental Figure 2](#)).

133 We then tested whether the distribution of these PCTs varied with placental Cd concentrations.
134 Higher Cd was associated with increased proportions of PCT-1 (3.4% difference, T-test p-value = 0.010)
135 and PCT-2 (4.7% difference, T-test p-value = 0.022) in NHBCS, while higher Cd was associated with
136 increased proportions of PCT-2 (7.3% difference, T-test p-value = 0.047) in RICHHS. Since these
137 associations may confound the relationships between Cd and locus-specific DNAM, we then reproduced
138 the cohort specific multivariable models and meta-analysis while also adjusting for the estimated PCT
139 proportions. These adjustments substantially reduced inflation in NHBCS ($\hat{\lambda}=1.47$ before, and $\lambda=0.96$
140 after adjustment), while only modestly increasing inflation in RICHHS ($\lambda=1.03$ before, and $\lambda=1.10$ after
141 adjustment). Furthermore, the PCT-adjustments reduced the proportion of small heterogeneity p-values (<
142 0.05) in the meta-analysis from 7.9% in the non-PCT adjusted models to 4.6% in the PCT-adjusted
143 models. Full results from the PCT-adjusted meta-analysis are provided in the supplemental materials
144 ([Supplemental Table 4](#)). The majority of Cd-DNAM associations from the original models were
145 attenuated after PCT-adjustments. For instance, the parameter estimates for our top 3 hits from the
146 original meta-analysis (cg15345741, cg16768966, and cg05888894) were reduced by 11%, 16%, and
147 45%, respectively, after adjustment for PCTs, demonstrating that cell mixture likely explained much of
148 the Cd-associated variation for cg05888894, but only a modest proportion of the Cd-DNAM associations
149 at cg15345741 and cg16768966. Additionally, the coefficients for cg15345741 and cg16768966 retained
150 highly significant p-values (< 1e-05) after PCT adjustment, due to improved standard errors around these
151 parameter estimates.

152 The PCT-adjusted top hit from NHBCS was cg20656525 on chromosome 14, downstream of
153 *TNFAIP2* and *EXOC3LA*, while the top hit from RICHHS was cg07708653 on chromosome 1, within the
154 body of the *HPDL* gene. The PCT-adjusted meta-analysis yielded 17 CpGs with Cd-association p-values
155 $< 1e-05$ (Table 3), two of which were significant at a 5% FDR (cg15345741 and cg11707084). Both of
156 the FDR significant CpGs also had significant linear associations with log-Cd (cg15345741 estimate =
157 0.17, p-value = 0.00019; cg11707084 estimate = 0.16, p-value = 0.000014) via linear mixed models. The
158 meta-analyses again revealed a clear trend towards higher methylation levels associated with higher Cd
159 concentrations; 94% of Cd-associated loci with p-values $< 1e-05$ had positive beta coefficients. Of note,
160 41% of CpGs with suggestive p-values ($< 1e-05$) were on chromosome 17, five of which were within the
161 *GAS7* gene. Given the high proportion of *GAS7* CpGs among the top hits from the meta-analysis, we
162 explored the co-methylation structure within *GAS7* and as well as all Cd-DNAM associations within and
163 around this gene. The strongest Cd-associated methylation sites tended to congregate around the first
164 exons of three of the four different *GAS7* variants (-a, -b and -d) and exhibited a strong co-methylation
165 structure as demonstrated by strong positive spearman correlation rho values (Figure 2).

166 We also investigated the underlying biology that may be affected by Cd-associated variations in
167 the placental epigenome. Since we had too few FDR-significant findings to perform an enrichment
168 analyses on only those hits, we tested for GO-term and KEGG pathway enrichment among the top 250
169 CpGs associated with placental Cd after PCT adjustment. The top GO terms included dimethylarginase
170 activity, eye and heart valve development, and heparin sulfate proteoglycan & polysaccharide
171 biosynthetic process (Supplemental Table 5), though none of these enrichments were significant at an
172 FDR threshold of 5%. On the other hand, our gene set was enriched with genes from numerous KEGG
173 pathways (Supplemental Table 6), 12 of which yielded p-values within an FDR threshold of 5% (Table
174 4); these included cell adhesion molecules (CAMs) and tight junctions, multiple cancer pathways,
175 multiple G-protein and nitric oxide signaling pathways, and cytotoxic processes.

176 To explore whether gene expression levels might be associated with variations in methylation
177 levels at Cd-associated CpGs with meta-analysis p-values $< 1e-05$, we performed an expression

178 quantitative trait methylation (eQTM) analysis. We regressed the expression levels of any gene within
179 100kb of a candidate CpG site on the methylation levels of that CpG. Three of these CpGs (cg24000528,
180 cg03917020, and cg23193177) were not within 100kb of any genes that yielded detectable RNA from our
181 placental samples, and thus 14 CpGs and 32 genes were included in this analysis; only *GAS7* was cis to
182 multiple CpGs within our set of candidates. We produced 36 linear models ranging between 1 and 7 genes
183 tested per CpG (Table 5). Of the 36 eQTM models, 10 yielded FDR-significant linear associations
184 between methylation and expression of six unique genes (Supplemental Figure 3). Of particular note,
185 were the associations between our top hit from the meta-analysis and the expression levels of *EXOC3L4*
186 ($\beta_1 = 1.95$, p-value = 0.0043) and *TNFAIP2* ($\beta_1 = 1.88$, p-value = 0.0055). Additionally, all five of the Cd-
187 associated *GAS7* CpGs (within or nearby the first exons of variants-a, -b, and -d) were associated with
188 increased *GAS7* expression (Table 5). On the other hand, DNAM surrounding the first exon of *GAS7*
189 variant-c tended to be inversely correlated with mRNA levels (Figure 2).

190 We then tested whether the expression levels of the FDR-significant genes from the eQTM
191 analysis were associated with fetal growth metrics standardized by sex and gestational age (22) (Table 6).
192 We found that the higher placental expression levels of *TNFAIP2* and *ACOT7* were modestly associated
193 with decreasing birth weight z-scores, while higher expression of *RORA* was modestly associated with
194 increased birth weight z-scores (p-values < 0.05). Higher expression of *ACOT7* was also associated with
195 decreasing birth length and head circumference z-scores (p-values < 0.05). None of the other gene
196 expression levels were associated with birth metrics.

197 Finally, we explored whether top hits from a previous small sample study (n=24) of placental Cd
198 and DNA methylation levels (19) demonstrated associations in our study. In our meta-analysis one of the
199 top six loci from that study, cg04528060 on chromosome 4 within the ADP-ribosylation factor-like 9
200 (*ARL9*) gene, demonstrated a nominally significant association with Cd ($\beta_1 = -0.49$, p-value = 0.014) with
201 the same direction of effect in our meta-analysis. However, we did not observe any evidence of sex-
202 specific effects at this, or the other 5 CpG sites. We also explored whether any of CpGs from our meta-
203 analysis with p-values < 1e-05 yielded statistically significant interactions between Cd and fetal sex

204 (Supplemental Table 7). Of the 17 models, only cg02600679 within the gene body of *FAM38A* produced
205 a significant interaction (p-value = 0.031). Stratifying by fetal sex revealed that Cd was associated with
206 increased DNAM at cg02600679 among males ($\beta_1 = 0.26$, p-value = 0.0012), while there was no
207 association among females (p-value = 0.97).
208

209 **Discussion:**

210 We conducted a large EWAS study across two cohorts to examine associations between Cd
211 concentrations and DNAM from human placentae and considered the confounding effects of placental
212 tissue heterogeneity. We further assessed eQTM and associations with birth metrics. We used identical
213 measurement techniques for both DNAM and trace metal measurement in the two cohorts, and the
214 placental concentrations of Cd were similar to the 3.53 ng/g previously reported in the full NHBCS cohort
215 (23), and towards the lower end but consistent with other cohorts that have reported concentrations
216 ranging as low as 1.2 ng/g to as high as 53.3 ng/g (24). We identified Cd-associated variations at two
217 CpGs within a 5% FDR threshold and an additional 15 CpGs with meta-analysis p-values < 1e-05. At all
218 but one of these loci, higher Cd was associated with increased methylation levels. We also found that the
219 methylation levels at these CpGs were associated with the expression patterns of six genes, and the
220 expression of three genes (*TNFAIP2*, *ACOT7*, and *RORA*) were associated with birth metrics.

221 The observed preference for Cd-associated hypermethylation, as opposed to hypomethylation, is
222 consistent with previous studies that observed higher Cd to be associated with a higher proportion of
223 hypermethylation at gene promoters from fetal and maternal blood samples (25) and with increased
224 methylation at regulatory regions for imprinted genes from cord blood samples (26). In contrast, others
225 have observed Cd-associated hypomethylation in placental DNA, however this was from a relatively
226 small study that focused on sex-specific associations (19). Cd has been shown to enhance DNA
227 methyltransferase activity and thus increase the preponderance of hypermethylated loci, though these
228 effects on DNAM may vary by timing and duration of exposure (27).

229 The top hit from our study, cg15345741 is on chromosome 14, approximately 70kb and 97kb
230 downstream of the *TNFAIP2* and *EXOC3L4* genes, respectively, which encode the *tumor necrosis factor*,
231 *alpha-induced protein 2* and the *exocyst complex component 3 like 4* protein. Higher methylation at
232 cg15345741 was associated with higher expression of both *TNFAIP2* and *EXOC3L4*, and thus may be
233 involved in the regulation of these genes or as a possible marker of their expression, though it does not
234 function through the promoter methylation paradigm. However, the location of this CpG does overlap

235 with a DNase hypersensitivity region and multiple putative transcription factor binding sites. We also
236 identified the Cd-associated loci cg11707084 on chromosome 19 which is cis to six genes (*SEPWI*,
237 *GLTSCR2*, *EHD2*, *TRPX1*, *CRX*, and *SULT2A1*), though the expression levels of these genes were not
238 detectable in placental tissue or those that were did not yield associations with DNAM at our candidate
239 loci. Thus, the potential functional role of DNAM at cg11707084 is unclear.

240 Our top 250 sites included genes that were significantly enriched for nitric oxide (NO) and G-
241 protein signaling pathways (ras, apelin, oxytocin, and cGMP-PKG), cancerous process, cellular adhesion,
242 cellular metabolism, and cytotoxicity. These pathways are consistent with currently recognized
243 mechanisms of Cd-associated reproductive toxicity that have primarily been studied in animal or *in vitro*
244 models: altered signal transduction, impaired cellular adhesion, disruptions to cell-cycle, increased
245 oxidative stress and cytotoxic/apoptotic signaling (3). We provide evidence that these biologic processes
246 in the human placenta may be perturbed by Cd-associated differential methylation, despite overall Cd-
247 exposures being low and our study samples primarily consisting of healthy pregnancies. Furthermore, the
248 functions of the genes whose expression levels were associated with our top hits are involved in similar
249 mechanisms.

250 The proteins produced by both *TNFAIP2* and *EXOC3L4* are structurally similar to subunits of the
251 exocyst, a molecular complex that orchestrates numerous important developmental functions including
252 cellular membrane outgrowth, establishing cell polarity, and mediating cell-to-cell adhesion and
253 communication (28). Given the crucial roles that the above processes play during development, it is not
254 surprising that the exocyst is critical to proper placental development. Mouse models have demonstrated
255 that full knockout of the exocyst subunits can cause early death of the embryo (28), while disrupted
256 exocyst function may be associated with preeclampsia (29). Both *TNFAIP2* and *EXOC3L4* also serve
257 other biological roles that may be important to Cd-associated toxicity. For instance, the expression of
258 *TNFAIP2* can be up-regulated by TNF α , retinoic acid, interleukin-1 β , and other pro-inflammatory and
259 cytotoxic signals (30, 31). It is also highly expressed in some cancerous tissues, is related to poor
260 survival, and is involved in cellular motility, invasion, and metastasis (32). Additionally, *TNFAIP2*,

261 among other pro-inflammatory genes, has also been shown to be overexpressed in the hippocampus of
262 Alzheimer's disease patients (33) and may mediate cell death in motor neurons of amyotrophic lateral
263 sclerosis patients (34), suggesting a role in neurodegenerative disease. SNPs within *EXOC3L4* have been
264 implicated to affect gamma-glutamyl transferase (GGT) activity (35), and elevated levels of GGT are
265 recognized markers of oxidative stress (36), which plays a central role in Cd-toxicity. Disrupted
266 epigenetic regulation of these genes and their involvement in placental growth, development,
267 inflammation and/or oxidative stress responses present compelling potential mechanisms through which
268 Cd may elicit some of its reproductive toxicity.

269 Expression of the *GAS7* gene, which encodes the *growth arrest specific 7* protein, was strongly
270 associated with DNAM at all five candidate CpGs that were associated with Cd (p-values < 1e-05).
271 Interestingly, the strongest Cd-associations were localized nearby the first exons of transcript variants a,
272 b, and d for *GAS7*, and thus these variations may play roles in Cd-associated alternative splicing of this
273 gene. However, read-depth of our RNA-seq data was not sufficient to accurately test for variant-specific
274 expression levels, and this should be investigated further in future studies given the potential for variant-
275 specific functions described below. The *GAS7* gene, which is highly expressed in embryonic cells (37),
276 mature brain cells and Purkinje neurons, plays an essential role in neurite outgrowth (38) and inducing
277 neuronal cell death (39), specific functions may be isoform dependent. In the mouse embryo, levels of
278 *GAS7* decrease as the embryo grows, except for in neuronal-specific epidermal cells in which expression
279 remains high (37). *GAS7* expression also plays an essential role in other non-neuronal growth functions,
280 such as osteoblast differentiation and bone development (40). Furthermore, neural outgrowth and
281 arborization are shared functions between the exocyst and *GAS7* (28, 38), while *TNFAIP2* expression has
282 been associated with some neurodegenerative conditions (33, 34). Relatively few studies have
283 investigated the relationship between perinatal Cd exposure and neurodevelopment, though some
284 epidemiologic studies suggest a relationship with impaired cognition in childhood (41). This raises some
285 questions about the possible roles of the above genes in Cd-associated cognitive and neurobehavioral
286 deficits. However, the functional roles of these genes in human placental tissue are not well understood.

287 Thus it is unclear how, or whether, Cd-associated differential placental DNAM and expression may affect
288 early life cognitive function and should be the focus of additional research.

289 Genes known to contribute in neurite outgrowth are sometimes also involved in regulating
290 placental angiogenesis, as has been shown for vascular endothelial growth factor (VEGF) and
291 neurotrophins (42). Interestingly, sterol regulatory element binding proteins (SREBPs), such as *SREBF1*
292 whose expression was strongly correlated with one of our top hits, play key roles in VEGF-induced
293 angiogenesis (43). Additionally, inappropriate placental vascularization can result in fetal growth
294 restriction or other pregnancy complications (42). In our study, the expression of *GAS7*, *EXOC3L4*, and
295 *SREBF1* were not associated with standardized percentiles for birthweight, head circumference, or birth
296 length, and thus may not be directly related to fetal growth. Whereas expression of *TNFAIP2*, *ACOT7*,
297 and *RORA* were associated with fetal growth. *ACOT7*, which encodes *acyl-CoA thioesterase 7*, is
298 involved in lipid metabolism in neuronal cells and may prevent neurotoxicity (44), while it also promotes
299 pro-inflammatory activities by up-regulating arachidonic acid production (45). *RORA*, which encodes the
300 *retinoic acid receptor-related orphan receptor alpha*, is involved in numerous developmental processes
301 including innate immune system development, regulating inflammatory responses, neuronal survival and
302 growth, as well as lipid and glucose metabolism (46). One common biological thread that differentiates
303 *TNFAIP2*, *ACOT7*, and *RORA*, from *GAS7*, *EXOC3L4*, and *SREBF1*, is their role in mediating
304 inflammatory and immune responses. Thus, although we observed numerous Cd-associated variations in
305 DNAM among genes involved in neurodevelopment and cellular growth processes, variations involved in
306 inflammatory and immune signaling disruption may impact fetal growth.

307 Other epidemiologic studies of maternal Cd and birth/pregnancy outcomes, that also investigated
308 molecular mechanisms, have found that higher Cd may affect essential metals transport to the fetus (13),
309 reduce expression of a protocadherin, which are involved in neuronal development (47) and affect the
310 DNAM of genes involved in cell death, lipid metabolism, and cancer pathways (25). Similarly, electronic
311 waste (e-waste) exposure, which is associated with higher placental concentrations of toxic metals
312 including Cd, has been associated with differential expression of proteins primarily involved in immune

313 and metabolic processes (48). The top hits from a small sample EWAS of placental Cd and DNAM
314 observed differential methylation of genes involved in cellular metabolism, growth, and cell damage
315 response (19). We were able to reproduce Cd-associated DNAM variations for one top hit from this study,
316 cg04528060 in the 5'UTR of the *ARL9* gene, (meta-analysis P-value < 0.05) with the same direction of
317 effect. Interestingly, the *ARL9* gene encodes for a GTP-binding protein that is structurally similar to
318 members of the RAS superfamily (49), and RAS-signaling was the most highly enriched KEGG pathway
319 from our study.

320 The preponderance of similar biological processes associated with Cd across multiple studies
321 with different technologies, populations, and tissues, is striking, and suggests that common mechanisms
322 may be affected by Cd exposure during pregnancy. These findings should be interpreted within the
323 context of this study's limitations. This was an observational study in which placental Cd concentrations
324 and DNA-M levels were both measured in placenta at term. Thus, we cannot rule out the possibility of
325 reverse causation since exposure and outcome were measured at the same time-point, and the observed
326 associations at term may not be representative of Cd and DNA-M associations throughout development.
327 Though we adjusted for likely confounders in our study, we also cannot rule out the possibility that
328 unmeasured or residual confounding may have affected our observations. However, to our knowledge this
329 was the largest study yet to examine the relationships between placental Cd and DNAM, and the first
330 account for and evaluate associations with estimated cellular heterogeneity.

331 The loci that we did identify are involved in biological processes known to be affected by Cd-
332 toxicity and thus are strong candidates for future studies. We observed these associations in two healthy
333 populations with relatively low exposure levels. In the US, exposure to Cd has declined since the late
334 1980s, which is primarily attributed to declining smoking rates (50). However, Cd is still detectable in
335 numerous commonly consumed foods (51) and exposure prevalence remains quite high (52).
336 Additionally, e-waste recycling and disposal are emerging sources of environmental Cd-contamination
337 (and other toxic metals), particularly in developing countries (53). This raises some concerns about the
338 potential threshold for Cd-associated reproductive toxicity, and whether current dietary intake guidelines

339 and efforts to reduce tobacco smoke exposure during pregnancy adequately protect against adverse birth
340 and pregnancy outcomes. We encourage further molecular epidemiologic studies of the associations
341 between Cd and DNAM in placental tissue, particularly within populations with higher exposure levels
342 and different ethnic backgrounds, but also encourage studies of the associations between dietary intake
343 and smoking-related Cd exposure with the placental accumulation of Cd. Understanding these
344 relationships can help us determine whether changes to current guidelines and recommendations could
345 reduce Cd-associated adverse reproductive outcomes. Finally, the genes whose expression levels were
346 associated with DNAM at our top hits are multifunctional, and involved in neurodevelopmental and/or
347 neurodegenerative processes, cellular growth processes, and inflammatory/immune signaling. We also
348 found that the expression levels of three of these genes were associated with infant birth weight. Thus,
349 additional future studies should investigate whether differential placental epigenetic regulation and
350 expression of *GAS7*, *EXOC3L4*, *SREBF1*, *RORA*, *ACOT7* and *TNFAIP2*, are associated with
351 cognitive/neurobehavioral, growth, and immune outcomes in children.

352

353 **Methods:**

354 *New Hampshire Birth Cohort Study*

355 This study included mother-infant pairs from the NHBCS, an ongoing birth cohort initiated in
356 2009. Women that were currently pregnant, between 18 and 45 years of age, receiving prenatal care from
357 one of the study clinics in New Hampshire, reporting that the primary source of drinking water at their
358 residence was from an unregulated well, and having resided in the same household since their previous
359 menstrual period with no plans to move before delivery, were enrolled in the cohort. All participants
360 provided written informed consent in accordance with the requirements of the Institutional Review Board
361 (IRB) of Dartmouth College. The sample for this study consisted of NHBCS participants that were
362 recruited between February 2012 and September 2013, for whom placenta were sampled to conduct
363 genetic and epigenetic assays (n=343). Placental gross measures such as placental diameter (cm) and
364 placental weight (g) were collected immediately after delivery. Interviewer administered questionnaires
365 and medical record abstraction were utilized to collect sociodemographic, lifestyle, and anthropometric
366 data.

367 *Rhode Island Child Health Study*

368 The RICHS enrolled mother-infant pairs with non-pathologic pregnancies at the Women and
369 Infants' Hospital in Providence, RI, USA between September 2010 and February 2013. Exclusion criteria
370 consisted of mothers younger than 18 years of age, with life threatening conditions, pregnancies with
371 gestational time < 37 weeks, or infants with congenital/chromosomal abnormalities. All protocols were
372 approved by the institutional review boards at the Women and Infants Hospital of Rhode Island and
373 Dartmouth College and all participants provided written informed consent. Infants that were born small
374 for gestational age ($\leq 10^{\text{th}}$ BW percentile) or large for gestational age ($\geq 90^{\text{th}}$ BW percentile) were
375 oversampled, then infants adequate for gestational age (between the 10^{th} and 90^{th} BW percentiles) that
376 matched on gestational age and maternal age were coincidentally enrolled. This study included all
377 mother-infant-pairs for whom placental metal concentrations and placental DNA-M arrays had been
378 conducted (n=141). Interviewer administered questionnaires were utilized to collect sociodemographic

379 and lifestyle data; anthropometric and medical history data were obtained via structured medical records
380 review.

381 *Sample Collection:*

382 Placentae samples from RICHS and NHBCS were biopsied from the fetal side adjacent to the
383 cord insertion site, within 2 hours of delivery and maternal decidua was removed. Samples were placed in
384 RNAlater (Life Technologies, Carlsbad, CA) then frozen at -80°C . Both RNA and DNA were extracted
385 (Norgen Biotek, Thorold, ON) then quantified via the Qubit Fluorometer (Life Technologies), and
386 subsequently stored at -80°C .

387 *DNA Methylation, QC, and Normalization:*

388 Both cohorts measured DNAM via Illumina Infinium HumanMethylation450K BeadArray
389 (Illumina, San Diego, CA) at the University of Minnesota Genomics Center. Bisulfite modification was
390 done with the EZ Methylation kit (Zymo Research, Irvine, CA), samples were randomized across
391 multiple batches, and data was assembled using BeadStudio (Illumina). The raw array data are available
392 via the NCBI Gene Expression Omnibus (GEO) for NHBCS and RICHS via accession numbers
393 GSE71678 and GSE75248, respectively. For QC, we excluded probes with poor detection p-values (p-
394 value > 0.001), measuring DNA-M at X- and Y-linked CpG sites, with a single nucleotide polymorphism
395 (SNP) within 10bp of the target CpG or single base extension (SBE) (and allelic frequency $> 1\%$), or
396 cross-hybridizing to multiple genomic regions (54). Background correction, dye bias, and functional
397 normalization were performed via the minfi package (55) and standardization across probe-types was
398 performed via beta mixture quantile (BMIQ) normalization (56). Batch effects were corrected for using
399 empirical Bayes via the combat function in R (57). Methylation data were reported as β -values,
400 representing the proportion of methylated alleles for each individual CpG site. For the EWAS models
401 described below, we utilized M-values, logit-transformation of the beta values, which better approximate
402 a normal distribution (58).

403 *RNA sequencing, QC, and Normalization:*

404 Transcriptome-wide sequencing was performed on 200 placental samples from RICHS. We used
405 RNeasy Mini Kit (Qiagen, Valencia, CA) to isolate total RNA, then stored at -80°C until analysis. We
406 then quantified RNA via Nanodrop Spectrophotometer (Thermo Scientific, Waltham, MA), assessed
407 integrity via the Agilent Bioanalyzer (Agilen, Santa Clara, CA), removed ribosomal RNA via Ribo-Zero
408 Kit (59), converted to cDNA using random hexamers (Thermo Scientific, Waltham, MA), and performed
409 transcriptome-wide RNA sequencing via the HiSeq 2500 platform (Illumina, San Diego, CA) (60). Raw
410 reads have been deposited at the NCBI sequence read archive (SRP095910). Quality control was
411 performed in FastQC, then reads were mapped to the human reference genome (h19) using the Spliced
412 Transcripts Alignment to a Reference (STAR) aligner. Low-expressed genes were excluded, read counts
413 were adjusted for GC content (61), then normalized via the trimmed mean of m-values (TMM) (62); final
414 data are normalized log₂ counts per million (logCPM) reads.

415 *Cadmium Quantification:*

416 Trace element concentrations were quantified in RICHS and NHBCS placental samples at the
417 Dartmouth Trace Elements Analysis Core using inductively coupled plasma mass spectrometry (ICP-
418 MS); details of the processing are described elsewhere (23). None of the RICHS samples received non-
419 detectable Cd concentrations, while only four of the NHBCS samples received non-detects which were
420 assigned a value equal to the 0.5 percentile of detectable values. For EWAS models, we dichotomized
421 cadmium concentrations by the cohort-specific median concentrations (NHBCS median = 3.13 ng/g) and
422 (RICHS median = 4.34 ng/g) for the EWAS analysis.

423 *Estimation of placental tissue heterogeneity*

424 We addressed the potential issue of confounding due methylation-associated tissue heterogeneity
425 via the RefFreeEWAS package in R (21). This method utilized non-negative matrix factorization to
426 estimate the proportions of putative cellular mixtures and has been demonstrated to yield reliable
427 estimates of the constituent cell types and the underlying methylomes that define them (21). Utilizing the
428 10,000 CpGs with the greatest variability in methylation and we identified four and five putative
429 constituent cell mixtures in the RICHS and NHBCS cohorts respectively, then used the full set of CpGs

430 for each cohort (NHBCS $p=408,367$ and RICHs $p=397,040$) to estimate the relative proportions of each
431 putative cell-type per placental sample. We examined whether the 10,000 CpGs with the largest variance
432 across putative cell-type specific methylation values were enriched for gene ontology (GO) terms or
433 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, via the `gometh` function in the
434 `missMethyl` package in R (63). This method adjusts for the potential bias introduced by some genes
435 having greater probability of being included in the gene-set due to having greater numbers of CpGs on the
436 450K array. We examined the consistency of the CpGs with high cell-type specific methylation, as well
437 as consistency in the top 100 GO-terms and KEGG pathways associated with those highly variable CpGs.
438 *Epigenome-wide association study (EWAS)*

439 Robust linear regressions (`rlm`) were used to estimate the associations between CpG-specific
440 DNA methylation levels and high-vs-low placental concentrations of Cd. We produced two covariate-
441 adjusted models for each CpG within each cohort, first regressing M-values on dichotomized Cd-
442 concentrations (cohort-specific median as the reference) while adjusted for infant sex, maternal smoking
443 during pregnancy, and highest achieved maternal education level. Second, we produced the same models
444 while also adjusted for the estimated proportions of putative placental cell-types. Genomic inflation was
445 examined via the genomic inflation factor (λ) which was calculated with the `genABEL` package in R (64).
446 We selected the 250 CpGs with the smallest Cd-associated meta-analysis p-values and assessed these for
447 functional enrichment with GO-terms and KEGG pathways as described above. Statistical significance
448 for enrichment was determined at a 5% FDR.

449 *Meta Analyses*

450 We used METAL to meta-analyze the cohort-specific Cd-associations via inverse variance
451 weighted fixed-effects models, and evaluated interstudy heterogeneity via Cochran's Q-test p -values $<$
452 0.05. (65). We produced FDR-adjusted p -values via the `qvalue` package in R. Statistical significance was
453 determined at an FDR of 5%. Volcano plots were produced to visualize overall magnitude and direction
454 of associations, while Manhattan plots were produced via the `qqman` package to visualize the genomic
455 distribution of significant associations. Plots for specific regions of the genome were visualized using the

456 coMET package (66). Top hits from the study were also evaluated for dose-response relationships via
457 modeling associations between log-transformed continuous Cd and DNAM, via linear mixed models,
458 allowing for random intercepts by study.

459 *Expression Quantitative Trait Methylation (eQTM) Analysis*

460 We then tested whether any of our top Cd-associated CpGs could be expression quantitative trait
461 methylation (eQTM) loci within the 200 RICHS samples, 85 of which overlap with the samples utilized
462 in the EWAS. We tested for cis-gene-CpG associations (within 100kb of the target CpG site) by
463 regressing gene expression levels ($\log_2(\text{CPM})$) on CpG β -values using robust linear models. Statistical
464 significance for eQTM associations was determined at a 5% FDR.

465 *Associations with fetal growth*

466 We tested for associations between birth metrics and gene expression using Kendal correlations,
467 which is a non-parametric correlation test that allows for ties. Since birth metric z-scores were
468 standardized by fetal sex and by weeks of gestation, no additional adjustments were made to these
469 association tests.

470 *Replication of previous studies and sex-specific effects*

471 We then investigated whether we could reproduce the results from a previous study of sex-
472 specific associations between placental Cd and placental DNAM (19). Because multiple studies of various
473 fetal and maternal tissues have suggested that associations between Cd and DNAM may differ by fetal
474 sex (7, 19, 67), we also screened the CpGs from our study with meta-analysis p-values $< 1e-05$ for
475 interactions with fetal sex. For these analyses, we utilized linear mixed models with DNAM M-values as
476 the outcome, allowing for random intercepts by study, and included an interaction-term between DNAM
477 levels and fetal sex while adjusting for maternal education and maternal smoking. We utilized the Wald
478 test to produce p-values, and determined statistically significant interactions at a p-value threshold of
479 0.05. Any models producing interaction p-values < 0.05 were then re-run, stratified by fetal sex.

480

481 Table 1: Maternal and offspring characteristics for the NHBCS (n=343) and the RICHs (n=141) samples.

Maternal Characteristics		NHBCS*	RICHs*
Smoking During Pregnancy	Any	21 (6.1%)	19 (13.5%)
	None	322 (93.9%)	122 (86.5%)
Educational Attainment	High-School or less	37 (10.8%)	34 (24.1%)
	More than High-School	306 (89.2%)	107 (75.9%)
Ethnicity	Not White	6 (1.7%)	36 (25.5%)
	White	337 (98.3%)	107 (74.5%)
Infant Characteristics			
Gestational Age (weeks)		38.95 (1.54)	39.33 (1.00)
Birth Weight (grams)		3435.14 (510.95)	3623.09 (716.03)
Size for Gestational Age			
	Small (SGA)	15 (4.4%)	21 (14.9%)
	Adequate (AGA)	292 (86.1%)	68 (48.2%)
	Large (LGA)	32 (9.4%)	52 (36.9%)
Sex	Female	159 (46.4%)	68 (48.2%)
	Male	184 (53.6%)	73 (51.8%)
Placental Cd (ng/g)[†]			
	Overall	3.13 (2.61)	4.37 (2.71)
	Below cohort-specific median	2.07 (1.05)	2.78 (1.35)
	Above cohort-specific median	4.68 (2.32)	5.49 (2.44)

482 * Means and standard deviations are presented for continuous variables, whereas frequencies and
 483 percentages are presented for categorical variables.

484 [†] Medians and IQRs presented due to skewed distributions of Cd concentrations.

485

486 Table 2: Results from cohort-specific and meta-analyses of Cd-associated DNA methylation, for probes
 487 with meta-analysis p-values < 1e-05; adjusted for maternal smoking during pregnancy, maternal
 488 education and fetal sex.

CpG Annotations			Meta-Analysis			NHBCS (n=343)		RICHS (n=141)	
CpG ID	Ch.	Gene	β_1	p-value	FDR	β_1	p-value	β_1	p-value
cg16034168	1	<i>ACOT7</i>	0.39	1.92E-06	0.107	0.48	4.47E-05	0.31	7.39E-03
cg18384460	1	-	0.35	2.22E-06	0.107	0.43	1.44E-05	0.25	2.42E-02
cg08705382	1	<i>GALNT2</i>	0.34	9.51E-06	0.151	0.34	8.06E-04	0.35	3.79E-03
cg26262055	4	-	0.22	8.94E-06	0.151	0.29	1.15E-02	0.21	2.12E-04
cg26653990	5	-	0.23	8.75E-06	0.151	0.25	1.41E-04	0.19	1.87E-02
cg00766220	5	-	0.14	4.37E-06	0.147	0.13	3.32E-04	0.14	4.13E-03
cg09391066	6	-	0.19	6.10E-06	0.151	0.21	1.10E-03	0.18	1.65E-03
cg03797906	7	<i>TRIL</i>	0.17	9.62E-06	0.151	0.21	2.48E-04	0.14	7.97E-03
cg11157440	8	<i>CSGALNACT1</i>	0.29	7.64E-06	0.151	0.28	6.19E-04	0.30	3.91E-03
cg21437157	8	-	0.23	4.12E-06	0.147	0.19	1.53E-03	0.33	3.32E-04
cg05888894	8	<i>EXT1</i>	0.33	2.81E-07	0.046	0.33	5.56E-03	0.33	1.54E-05
cg03509329	14	<i>LRRC16B</i>	0.15	7.09E-06	0.151	0.19	9.15E-04	0.12	1.47E-03
cg15345741	14	-	0.28	4.10E-07	0.046	0.28	2.52E-04	0.29	4.64E-04
cg17232357	15	<i>SMAD6</i>	-0.19	2.16E-06	0.107	-0.17	1.04E-02	-0.21	5.85E-05
cg12642651	17	<i>GAS7</i>	0.38	5.81E-06	0.151	0.27	2.78E-02	0.48	3.09E-05
cg03762242	17	<i>GAS7</i>	0.26	3.38E-06	0.142	0.26	7.76E-04	0.26	1.34E-03
cg16768966	17	<i>GAS7</i>	0.37	4.05E-07	0.046	0.34	1.55E-03	0.40	6.93E-05
cg15285733	17	<i>GAS7</i>	0.17	7.32E-07	0.062	0.16	2.42E-03	0.18	8.78E-05
cg02393525	17	<i>NAGLU</i>	0.23	8.90E-06	0.151	0.16	3.33E-02	0.31	3.13E-05
cg06707369	19	<i>TYROBP</i>	0.17	7.49E-06	0.151	0.19	2.35E-04	0.15	8.89E-03
cg11707084	19	<i>CRX</i>	0.23	7.79E-06	0.151	0.25	1.37E-04	0.19	1.68E-02

489 CpG: cytosine-phosphate-guanine methylation site; Ch.: chromosome; FDR: false discovery rate adjusted
 490 p-value; NHBCS: New Hampshire Birth Cohort; RICHS: Rhode Island Child Health Study.

491 Table 3: Results from cohort-specific and meta-analyses of Cd-associated DNA methylation, for probes
 492 with meta-analysis p-values < 1e-05; adjusted for maternal smoking during pregnancy, maternal
 493 education, fetal sex, and estimated tissue heterogeneity.

CpG Annotations			Meta-Analysis			NHBCS (n=343)		RICHS (n=141)	
CpG ID	CHR	Gene	β_1	p-value	FDR	β_1	p-value	β_1	p-value
cg16034168*	1	<i>ACOT7</i>	0.32	2.80E-06	0.171	0.32	3.11E-04	0.32	2.78E-03
cg24000528	2	<i>MIR10B</i>	0.16	7.59E-06	0.182	0.18	1.52E-03	0.16	1.50E-03
cg24696183	11	<i>KCNQ1DN</i>	-0.16	4.31E-06	0.171	-0.20	5.67E-04	-0.14	1.71E-03
cg23193177	11	-	0.24	4.68E-06	0.171	0.22	4.94E-04	0.27	2.70E-03
cg15345741*	14	-	0.25	1.49E-07	0.040	0.25	9.69E-05	0.24	4.22E-04
cg06100161	15	<i>RORA</i>	0.20	1.97E-06	0.171	0.12	6.28E-02	0.28	1.99E-06
cg03917020	16	-	0.14	5.52E-06	0.174	0.13	3.62E-04	0.14	4.85E-03
cg02600679	16	<i>FAM38A</i>	0.25	4.92E-06	0.171	0.17	2.55E-02	0.34	2.02E-05
cg04267691	17	<i>GAS7</i>	0.18	6.98E-06	0.178	0.12	3.09E-02	0.24	2.57E-05
cg10517535	17	<i>GAS7</i>	0.31	6.46E-06	0.177	0.16	6.97E-02	0.59	3.08E-07
cg03762242*	17	<i>GAS7</i>	0.23	3.33E-06	0.171	0.20	2.08E-03	0.26	4.12E-04
cg16768966*	17	<i>GAS7</i>	0.31	1.92E-06	0.171	0.23	7.31E-03	0.42	3.12E-05
cg10618704	17	<i>GAS7</i>	0.36	5.91E-06	0.174	0.24	2.03E-02	0.54	1.65E-05
cg11851174	17	<i>RAI1</i>	0.19	3.69E-06	0.171	0.20	1.07E-04	0.18	1.10E-02
cg19944656	17	<i>TIMP2</i>	0.18	2.62E-06	0.171	0.17	8.06E-05	0.18	9.82E-04
cg11707084*	19	<i>CRX</i>	0.22	2.09E-07	0.040	0.23	3.10E-05	0.22	1.93E-03
cg03566291	21	-	0.23	9.46E-06	0.201	0.21	9.57E-04	0.29	2.38E-03

494 * CpG Sites that yielded meta-analysis p-values < 1e-05 in models before and after heterogeneity
 495 adjustments.

496 CpG: cytosine-phosphate-guanine methylation site; CHR: chromosome; FDR: false discovery rate
 497 adjusted p-value; NHBCS: New Hampshire Birth Cohort; RICHS: Rhode Island Child Health Study.

498 Table 4: Significantly enriched KEGG pathways (FDR 5%) associated with for genes annotated to the top
499 250 CpG sites from the PCT-adjusted meta-analysis; N represents the total genes in the KEGG path, DE
500 represents the number of genes within our top 250 sites.

KEGG Pathway Description	N	DE	FDR
Ras signaling pathway	217	6	0.0058
Tight junction	163	5	0.0058
Platelet activation	122	4	0.019
Cell adhesion molecules (CAMs)	133	4	0.019
Taste transduction	79	3	0.019
Apelin signaling pathway	136	4	0.019
Pancreatic cancer	61	3	0.026
Oxytocin signaling pathway	150	4	0.026
cGMP-PKG signaling pathway	157	4	0.026
Viral carcinogenesis	191	4	0.030
Metabolic pathways	1189	8	0.034
Natural killer cell mediated cytotoxicity	107	3	0.042

501

502 Table 5: Expression quantitative trait methylation (eQTM) results for CpGs with PCT-adjusted meta-
 503 analysis p-values < 1e-05; analysis conducted with RICHS placental samples for which RNAseq and
 504 Illumina Infinium 450K were performed (n=200).

CpG Site	Gene	β_1	p-value	FDR
cg15345741	<i>EXOC3L4</i>	2.28	5.22E-03	0.019
cg15345741	<i>TNFAIP2</i>	2.56	4.97E-03	0.019
cg11707084	<i>EHD2</i>	-1.76	4.12E-02	0.11
cg11707084	<i>GLTSCR2</i>	0.68	4.16E-01	0.66
cg11707084	<i>SEPWI</i>	0.89	3.76E-01	0.66
cg16768966	<i>GAS7</i>	4.18	3.15E-06	2.27E-05
cg06100161	<i>RORA</i>	4.18	1.04E-05	6.22E-05
cg19944656	<i>CYTH1</i>	0.77	4.20E-01	0.66
cg19944656	<i>USP36</i>	0.99	3.40E-01	0.64
cg19944656	<i>TIMP2</i>	-0.16	8.91E-01	0.97
cg19944656	<i>LGALS3BP</i>	-0.01	9.93E-01	0.99
cg16034168	<i>RPL22</i>	-0.21	7.18E-01	0.97
cg16034168	<i>RNF207</i>	0.94	1.74e_01	0.37
cg16034168	<i>ICMT</i>	-0.52	4.18E-01	0.66
cg16034168	<i>GPR153</i>	-0.05	9.38E-01	0.97
cg16034168	<i>ACOT7</i>	2.20	1.87E-03	0.0096
cg03762242	<i>GAS7</i>	4.24	1.19E-06	1.07E-05
cg11851174	<i>RAI1</i>	0.14	9.08E-01	0.97
cg11851174	<i>SREBF1</i>	4.19	2.21E-03	0.0099
cg11851174	<i>TOMIL2</i>	2.35	1.05E-01	0.25
cg24696183	<i>KCNQ1</i>	0.25	8.03E-01	0.97
cg24696183	<i>CDKN1C</i>	-1.90	7.62E-02	0.20
cg24696183	<i>SLC22A18</i>	-1.75	1.22E-01	0.28
cg24696183	<i>PHLDA2</i>	-2.74	2.00E-02	0.065
cg24696183	<i>NAP1L4</i>	-0.24	8.34E-01	0.97
cg02600679	<i>RNF166</i>	0.17	8.49E-01	0.97
cg02600679	<i>CTU2</i>	0.42	5.51E-01	0.78
cg02600679	<i>FAM38A</i>	1.77	2.27E-02	0.068
cg02600679	<i>CDT1</i>	-0.72	4.38E-01	0.66
cg02600679	<i>APRT</i>	-0.06	9.41E-01	0.97
cg02600679	<i>GALNS</i>	1.17	1.89E-01	0.38
cg02600679	<i>TRAPPC2L</i>	0.15	8.65E-01	0.97
cg10618704	<i>GAS7</i>	3.84	7.08E-14	1.27E-12
cg10517535	<i>GAS7</i>	3.16	8.86E-10	1.06E-08
cg04267691	<i>GAS7</i>	9.17	8.97E-16	3.23E-14
cg03566291	<i>SIK1</i>	0.17	8.87E-01	0.97

505

506 Table 6: Associations between expression levels with growth metrics at birth in RICHS (n=200); birth
507 weight (BW), length (BL), and head circumference (HC) were sex-specific z-scores that have been
508 adjusted for gestational age, Tau represents the Kendall correlation coefficient.

Gene	BW Tau (p-value)	BL Tau (p-value)	HC Tau (p-value)
<i>TNFAIP2</i>	-0.099 (0.039)	-0.057 (0.24)	-0.092 (0.061)
<i>EXOC3L4</i>	-0.024 (0.62)	0.005 (0.91)	-0.030 (0.54)
<i>GAS7</i>	0.012 (0.80)	0.032 (0.51)	0.032 (0.51)
<i>RORA</i>	0.106 (0.026)	0.068 (0.16)	0.077 (0.12)
<i>ACOT7</i>	-0.134 (0.0048)	-0.106 (0.029)	-0.145 (0.0032)
<i>SREBF1</i>	0.032 (0.50)	0.003 (0.95)	-0.051 (0.30)

509

510 Acknowledgements:

511 This work was supported by the National Institutes of Health [NIH-NIMH R01MH094609, NIH-NIEHS
512 R01ES022223, NIH-NIGMS P20 GM104416 and NIH-NIEHS P01 ES022832] and by the United States
513 Environmental Protection Agency [US EPA grant RD83544201]. Its contents are solely the responsibility
514 of the grantee and do not necessarily represent the official views of the US EPA. Further, the US EPA
515 does not endorse the purchase of any commercial products or services mentioned in the presentation.
516

517 References:

- 518
- 519 1. Faroon O, et al. Toxicological Profile for Cadmium. *Agency Toxic Subst Dis Regist Atlanta*
520 (G(2012)).
 - 521 2. Järup L, Åkesson A (2009) Current status of cadmium as an environmental health problem.
522 *Toxicol Appl Pharmacol* 238(3):201–208.
 - 523 3. Thompson J, Bannigan J (2008) Cadmium: Toxic effects on the reproductive system and the
524 embryo. *Reprod Toxicol* 25:304–315.
 - 525 4. Al-Saleh I, et al. (2015) Interaction between cadmium (Cd), selenium (Se) and oxidative stress
526 biomarkers in healthy mothers and its impact on birth anthropometric measures. *Int J Hyg Environ*
527 *Health* 218(1):66–90.
 - 528 5. Al-Saleh I, Shinwari N, Mashhour A, Rabah A (2014) Birth outcome measures and maternal
529 exposure to heavy metals (lead, cadmium and mercury) in Saudi Arabian population. *Int J Hyg*
530 *Environ Health* 217(2–3):205–218.
 - 531 6. Kippler M, et al. (2012) Environmental exposure to arsenic and cadmium during pregnancy and
532 fetal size: A longitudinal study in rural Bangladesh. *Reprod Toxicol* 34(4):504–511.
 - 533 7. Kippler M, et al. (2013) Sex-specific effects of early life cadmium exposure on DNA methylation
534 and implications for birth weight. *Epigenetics* 8(5):494–503.
 - 535 8. Menai M, et al. (2012) Association between maternal blood cadmium during pregnancy and birth
536 weight and the risk of fetal growth restriction: The EDEN mother–child cohort study. *Reprod*
537 *Toxicol* 34(4):622–627.
 - 538 9. Laine JE, et al. (2015) Placental cadmium levels are associated with increased preeclampsia risk.
539 *PLoS One* 10(9):1–9.
 - 540 10. Wang H, et al. (2016) Maternal serum cadmium level during pregnancy and its association with
541 small for gestational age infants: a population-based birth cohort study. *Sci Rep* 6(22631):1–7.
 - 542 11. Valko M, Jomova K, Rhodes CJ, Kuča K, Musílek K (2016) Redox- and non-redox-metal-induced
543 formation of free radicals and their role in human disease. *Arch Toxicol* 90:1–37.
 - 544 12. Gundacker C, Hengstschlger M (2012) The role of the placenta in fetal exposure to heavy
545 metals. *Wiener Medizinische Wochenschrift* 162(9–10):201–206.
 - 546 13. Kippler M, et al. (2010) Accumulation of cadmium in human placenta interacts with the transport
547 of micronutrients to the fetus. *Toxicol Lett* 192:162–168.
 - 548 14. Wang H, et al. (2016) Maternal cadmium exposure reduces placental zinc transport and induces
549 fetal growth restriction in mice. *Reprod Toxicol* 63:174–182.
 - 550 15. Wang Z, et al. (2012) Cadmium-induced teratogenicity: Association with ROS-mediated
551 endoplasmic reticulum stress in placenta. *Toxicol Appl Pharmacol* 259(2):236–247.
 - 552 16. Wang F, et al. (2014) Preeclampsia induced by cadmium in rats is related to abnormal local
553 glucocorticoid synthesis in placenta. *Reprod Biol Endocrinol* 12(77):1–9.
 - 554 17. Erboga M, Kanter M (2016) Effect of Cadmium on Trophoblast Cell Proliferation and Apoptosis
555 in Different Gestation Periods of Rat Placenta. *Biol Trace Elem Res* 169(2):285–293.
 - 556 18. Vilahur N, Vahter M, Broberg K (2015) The Epigenetic Effects of Prenatal Cadmium Exposure.
557 *Curr Envir Heal Rpt* 2:195–203.
 - 558 19. Mohanty AF, et al. (2015) Infant sex-specific placental cadmium and DNA methylation
559 associations. *Environ Res* 138:74–81.

- 560 20. Houseman EA, et al. (2012) Open Access DNA methylation arrays as surrogate measures of cell
561 mixture distribution. *BMC Bioinformatics* 13(86).
- 562 21. Houseman EA, et al. (2016) Reference-free deconvolution of DNA methylation data and
563 mediation by cell composition effects. *BMC Bioinformatics* 17(1):259.
- 564 22. Fenton TR, Kim JH (2013) A systematic review and meta-analysis to revise the Fenton growth
565 chart for preterm infants. *BMC Pediatr* 13(59). doi:10.1186/1471-2431-13-59.
- 566 23. Punshon T, et al. (2016) Placental Metal Concentrations in Relation to Maternal and Infant
567 Toenails in a U.S. Cohort. *Environ Sci Technol* 50(3):1587–1594.
- 568 24. Esteban-Vasallo MD, Aragonés N, Pollan M, López-Abente G, Perez-Gomez B (2012) Mercury,
569 cadmium, and lead levels in human placenta: A systematic review. *Environ Health Perspect*
570 120(10):1369–1377.
- 571 25. Sanders AP, et al. (2014) Cadmium exposure and the epigenome: Exposure-associated patterns of
572 DNA methylation in leukocytes from mother-baby pairs. *Epigenetics* 9(2):212–221.
- 573 26. Vidal AC, et al. (2015) Maternal cadmium, iron and zinc levels, DNA methylation and birth
574 weight. *BMC Pharmacol Toxicol* 16(20):1–9.
- 575 27. Takiguchi M, Achanzar WE, Qu W, Li G, Waalkes MP (2003) Effects of cadmium on DNA-
576 (Cytosine-5) methyltransferase activity and DNA methylation status during cadmium-induced
577 cellular transformation. *Exp Cell Res* 286(2):355–365.
- 578 28. Martin-Urdiroz M, Deeks MJ, Horton CG, Dawe HR, Jourdain I (2016) The Exocyst Complex in
579 Health and Disease. *Front Cell Dev Biol* 4(April):1–22.
- 580 29. Gonzalez IM, Ackerman WE, Vandre DD, Robinson JM (2014) Exocyst complex protein
581 expression in the human placenta. *Placenta* 35(7):442–449.
- 582 30. Sarma V, Wolf FW, Marks RM, Shows TB, Dixit VM (1992) Cloning of a novel tumor necrosis
583 factor-alpha-inducible primary response gene that is differentially expressed in development and
584 capillary tube-like formation in vitro. *J Immunol* 148(10):3302–3312.
- 585 31. Rusiniak ME, Yu M, Ross DT, Tolhurst EC, Slack JL (2000) Identification of B94 (TNFAIP2) as
586 a potential retinoic acid target gene acute promyelocytic leukemia. *Cancer Res* 60(7):1824–1829.
- 587 32. Chen L-C, et al. (2011) A novel role for TNFAIP2: its correlation with invasion and metastasis in
588 nasopharyngeal carcinoma. *Mod Pathol* 24(2):175–184.
- 589 33. Colangelo V, et al. (2002) Gene expression profiling of 12633 genes in Alzheimer hippocampal
590 CA1: Transcription and neurotrophic factor down-regulation and up-regulation of apoptotic and
591 pro-inflammatory signaling. *J Neurosci Res* 70(3):462–473.
- 592 34. Brohawn DG, O'Brien LC, Bennett JP (2016) RNAseq analyses identify tumor necrosis factor-
593 mediated inflammation as a major abnormality in ALS spinal cord. *PLoS One* 11(8):1–25.
- 594 35. Middelberg RP, et al. (2012) Loci affecting gamma-glutamyl transferase in adults and adolescents
595 show age \times SNP interaction and cardiometabolic disease associations. *Hum Mol Genet*
596 21(2):446–455.
- 597 36. Fentiman IS (2012) Gamma-glutamyl transferase: risk and prognosis of cancer. *Br J Cancer*
598 106(9):1467–8.
- 599 37. Moorthy PP, Kumar AA, Devaraj H (2005) Stem Cells of Mice. 670:664–670.
- 600 38. You JJ, Lin-Chao S (2010) Gas7 functions with N-WASP to regulate the neurite outgrowth of
601 hippocampal neurons. *J Biol Chem* 285(15):11652–11666.
- 602 39. Hung FC, Chao CCK (2010) Knockdown of growth-arrest-specific gene 7b (gas7b) using short-
603 hairpin RNA desensitizes neuroblastoma cells to cisplatin: Implications for preventing apoptosis
604 of neurons. *J Neurosci Res* 88(16):3578–3587.
- 605 40. Chao CCK, Hung FC, Chao JJ (2013) Gas7 is required for mesenchymal stem cell-derived bone
606 development. *Stem Cells Int* 2013. doi:10.1155/2013/137010.
- 607 41. Sanders AP, Henn BC, Wright RO, Wright RO (2015) Perinatal and Childhood Exposure to
608 Cadmium, Manganese, and Metal Mixtures and Effects on Cognition and Behavior: A Review
609 of Recent Literature Bayley Scales of Infant Development. 284–294.
- 610 42. Sahay AS, Sundrani DP, Joshi SR (2015) Regional changes of placental vascularization in

- 611 preeclampsia: A review. *IUBMB Life* 67(8):619–625.
- 612 43. Zhou R-H, et al. (2004) Vascular Endothelial Growth Factor Activation of Sterol Regulatory
613 Element Binding Protein: A Potential Role in Angiogenesis. *Circ Res* 95(5):471–478.
- 614 44. Ellis JM, Wong GW, Wolfgang MJ (2013) Acyl coenzyme A thioesterase 7 regulates neuronal
615 fatty acid metabolism to prevent neurotoxicity. *Mol Cell Biol* 33(9):1869–82.
- 616 45. Brocker C, Carpenter C, Nebert DW, Vasiliou V (2010) Evolutionary divergence and functions of
617 the human acyl-CoA thioesterase gene (ACOT) family. *Hum Genomics* 4(6):411–420.
- 618 46. Cook DN, Kang HS, Jetten AM (2015) Retinoic Acid-Related Orphan Receptors (RORs):
619 Regulatory Functions in Immunity, Development, Circadian Rhythm, and Metabolism. *Nucl*
620 *Recept Res* 2:101185.
- 621 47. Everson TM, et al. (2016) Maternal cadmium, placental PCDHAC1, and fetal development.
622 *Reprod Toxicol* 65:263–271.
- 623 48. Xu L, et al. (2016) Differential proteomic expression of human placenta and fetal development
624 following e-waste lead and cadmium exposure in utero. *Sci Total Environ* 550:1163–1170.
- 625 49. Louro R, et al. (2004) RASL11A, member of a novel small monomeric GTPase gene family, is
626 down-regulated in prostate tumors. *Biochem Biophys Res Commun* 316(3):618–627.
- 627 50. Tellez-Plaza M, et al. (2012) Reduction in cadmium exposure in the United States population,
628 1988–2008: The contribution of declining smoking rates. *Environ Health Perspect* 120(2):204–
629 209.
- 630 51. Awata H, Linder S, Mitchell LE, Delclos GL (2016) Association of Dietary Intake and Biomarker
631 Levels of Arsenic, Cadmium, Lead, and Mercury among Asian Populations in the U.S.: NHANES
632 2011–2012. *Environ Health Perspect* 314(June):314–323.
- 633 52. Satarug S, Vesey DA, Gobe GC (2017) Health Risk Assessment of Dietary Cadmium Intake: Do
634 Current Guidelines Indicate How Much is Safe? *Environ Health Perspect* 125(3):284–288.
- 635 53. Sthiannopkao S, Wong MH (2013) Handling e-waste in developed and developing countries:
636 Initiatives, practices, and consequences. *Sci Total Environ* 463–464:1147–1153.
- 637 54. Chen YA, et al. (2013) Discovery of cross-reactive probes and polymorphic CpGs in the Illumina
638 Infinium HumanMethylation450 microarray. *Epigenetics* 8(2):203–209.
- 639 55. Fortin J-P, et al. (2014) Functional normalization of 450k methylation array data improves
640 replication in large cancer studies. *Genome Biol* 15(12):503.
- 641 56. Teschendorff AE, et al. (2013) A beta-mixture quantile normalization method for correcting probe
642 design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* 29(2):189–196.
- 643 57. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data
644 using empirical Bayes methods. *Biostatistics* 8(1):118–27.
- 645 58. Du P, et al. (2010) Comparison of Beta-value and M-value methods for quantifying methylation
646 levels by microarray analysis. *BMC Bioinformatics* 11(1):587.
- 647 59. Huang R, et al. (2011) An RNA-seq strategy to detect the complete coding and non-coding
648 transcriptome including full-length imprinted macro ncRNAs. *PLoS One* 6(11).
649 doi:10.1371/journal.pone.0027288.
- 650 60. Bentley DR, et al. (2008) Accurate whole human genome sequencing using reversible terminator
651 chemistry. *Nature* 456(7218):53–9.
- 652 61. Risso D, et al. (2011) GC-Content Normalization for RNA-Seq Data. *BMC Bioinformatics*
653 12(1):480.
- 654 62. Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: A Bioconductor package for differential
655 expression analysis of digital gene expression data. *Bioinformatics* 26(1):139–140.
- 656 63. Phipson B, Maksimovic J, Oshlack A (2015) MissMethyl: An R package for analyzing data from
657 Illumina’s HumanMethylation450 platform. *Bioinformatics* 32(2):286–288.
- 658 64. Aulchenko YS, De Koning DJ, Haley C (2007) Genomewide rapid association using mixed model
659 and regression: A fast and simple method for genomewide pedigree-based quantitative trait loci
660 association analysis. *Genetics* 177(1):577–585.
- 661 65. Willer CJ, Li Y, Abecasis GR (2010) METAL: Fast and efficient meta-analysis of genomewide

- 662 association scans. *Bioinformatics* 26(17):2190–2191.
- 663 66. Martin TC, Yet I, Tsai P-C, Bell JT (2015) coMET: visualisation of regional epigenome-wide
664 association scan results and DNA co-methylation patterns. *BMC Bioinformatics* 16(1):131.
- 665 67. Virani S, et al. (2016) DNA methylation is differentially associated with environmental cadmium
666 exposure based on sex and smoking status. *Chemosphere* 145:284–290.
- 667
- 668
- 669

670 Figure Legends

671

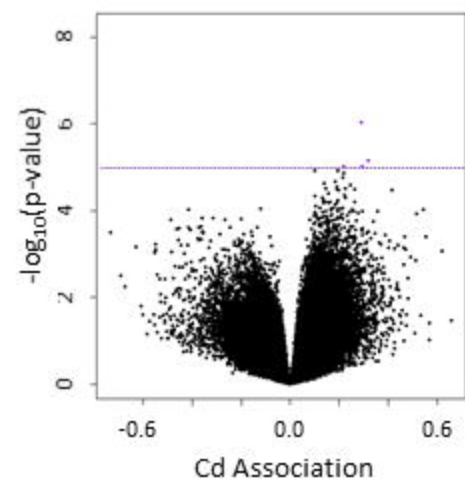
672 Figure 1: Volcano plots from each cohort-specific EWAS and the meta-analysis, and Manhattan plot of
673 meta-analysis results; all models adjusted for fetal sex, maternal smoking during pregnancy, mother's
674 highest achieved education level, and putative placental cell types; dashed line represents the FDR 5%
675 threshold, dotted line represents suggestive association (p-values < 1e-05).

676

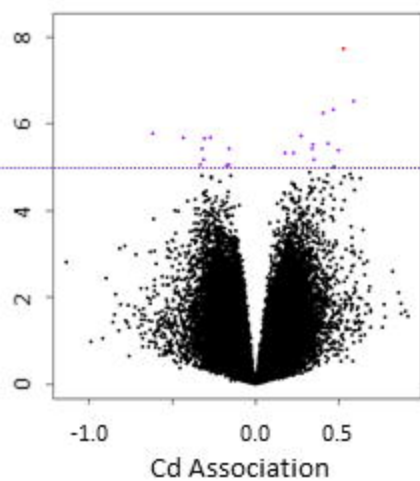
677 Figure 2: Manhattan plot of meta-analysis results surrounding the *GAS7* gene, annotated with Ensembl
678 gene names and RefSeq IDs for *GAS7* variants, Spearman correlations between *GAS7* CpG sites, and
679 CpG-expression correlations for *GAS7*. Cd-associated CpG outside of the *GAS7* coding region
680 ($\pm 1,500$ bp), have orange circles, where CpGs within *GAS7* have red, white, or blue circles corresponding
681 to positive, no, or negative associations, respectively.

682

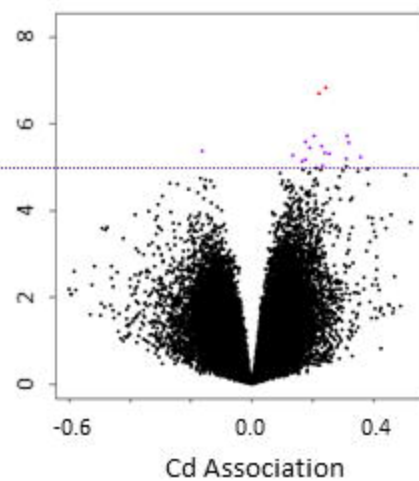
NHBCS



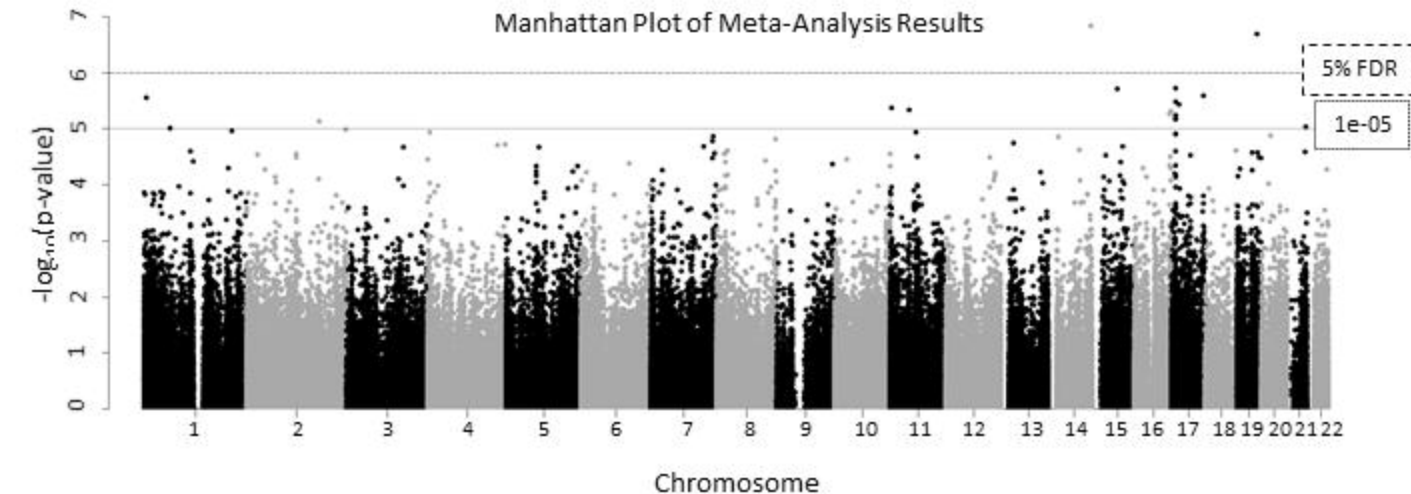
RICHs



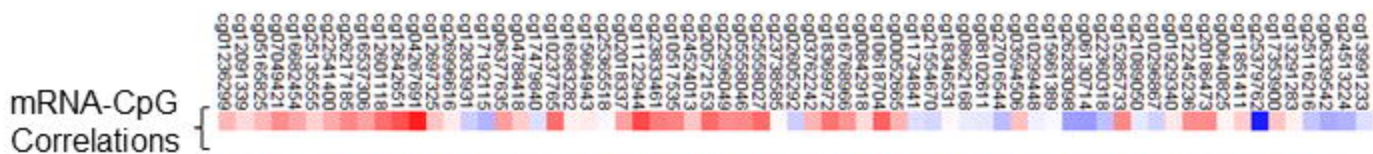
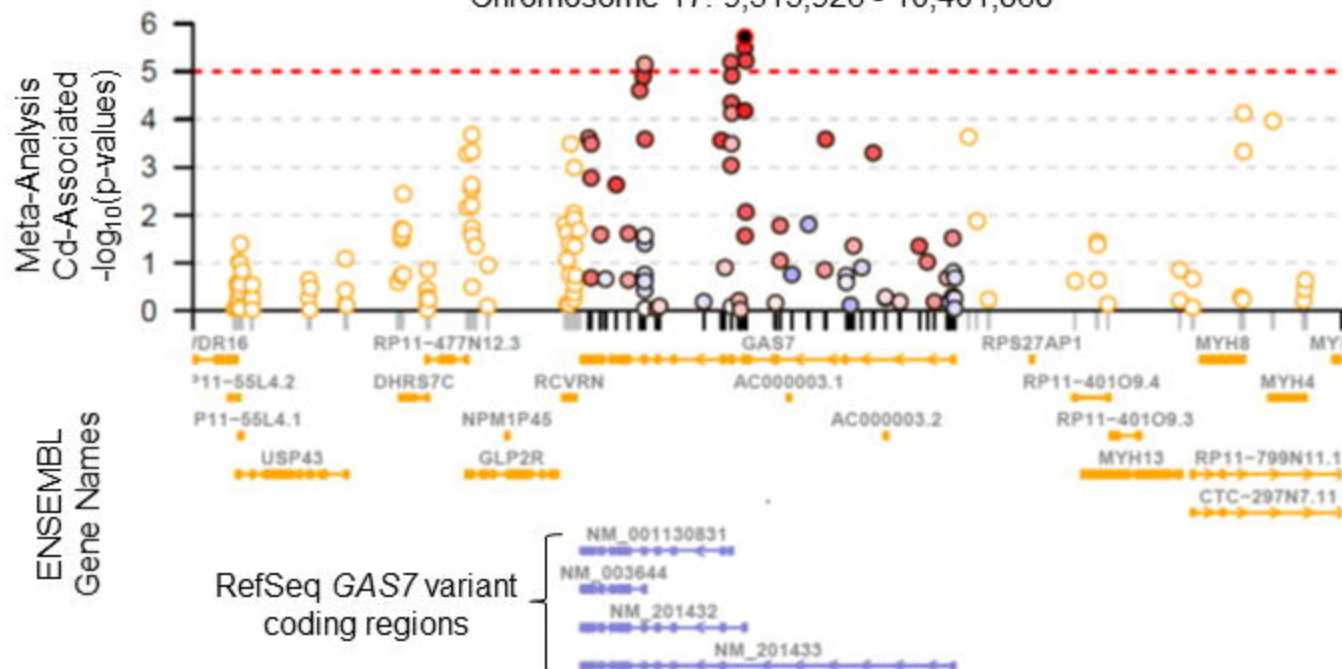
Meta-Analysis



Manhattan Plot of Meta-Analysis Results

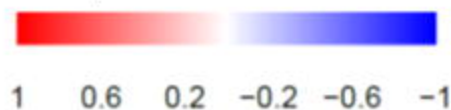


Chromosome 17: 9,513,926 - 10,401,868



Between-CpG
Correlations

Spearman's rho Values



● Outside
● *GAS7* DNAM
● cg16768966