

1 **Title: Parent of Origin Effects on Quantitative Phenotypes in a Founder**

2 **Population.**

3

4 **Author list:**

5 Sahar V. Mozaffari<sup>1,2\*</sup>, Jeanne M. DeCara<sup>3</sup>, Sanjiv J. Shah<sup>4</sup>, Roberto M. Lang<sup>3</sup>, Dan L.

6 Nicolae<sup>1,2,3,5</sup>, Carole Ober<sup>1,2\*</sup>

7

8 **Affiliations**

9 <sup>1</sup> Department of Human Genetics, University of Chicago, Chicago, IL 60637

10 <sup>2</sup> Committee on Genetics, Genomics, and Systems Biology, University of Chicago,

11 Chicago, IL 60637

12 <sup>3</sup> Department of Medicine, University of Chicago, Chicago, IL 60637

13 <sup>4</sup> Department of Medicine, Northwestern University Feinberg School of Medicine,

14 Chicago, IL 60611

15 <sup>5</sup> Department of Statistics, University of Chicago, Chicago, IL 60637

16

17 **Correspondence Emails**

18 \*smozaffari[at]uchicago.edu

19 \*c-ober[at]genetics.uchicago.edu

20 **ABSTRACT**

21 The impact of the parental origin of associated alleles in GWAS has been largely  
22 ignored. Yet sequence variants could affect traits differently depending on whether they  
23 are inherited from the mother or the father. To explore this possibility, we studied 21  
24 quantitative phenotypes in a large Hutterite pedigree. We first identified variants with  
25 significant single parent (maternal-only or paternal-only) effects, and then used a novel  
26 statistical model to identify variants with opposite parental effects. Overall, we identified  
27 parent of origin effects (POEs) on 11 phenotypes, most of which are risk factors for  
28 cardiovascular disease. Many of the loci with POEs have features of imprinted regions  
29 and many of the variants with POE are associated with the expression of nearby genes.  
30 Overall, our results indicate that POEs, which are often opposite in direction, are  
31 relatively common in humans, have potentially important clinical effects, and will be  
32 missed in traditional GWAS.

33

## 34 INTRODUCTION

35 Genome-wide association studies (GWAS) typically treat alleles inherited from  
36 the mother and the father as equivalent, although variants can affect traits differently  
37 depending on whether they are maternal or paternal in origin. In particular, parent of  
38 origin effects (POEs) can result from imprinting, where epigenetic modifications allows  
39 for differential gene expression on homologous chromosomes that is determined by the  
40 parental origin of the chromosome. Mutations in imprinted genes or regions can result in  
41 diseases. For example, two very different diseases, Prader-Willi Syndrome (PWS) and  
42 Angelman Syndrome (AS), are due to loss of function alleles in genes within an  
43 imprinted region on chromosome 15q11-13. Inheriting a loss of function mutation for the  
44 *SNRPN* gene from the father results in PWS but inheriting a loss of function mutation for  
45 the *UBE3A* gene from the mother results in AS<sup>1,2</sup>. Long noncoding RNA genes at this  
46 and other imprinted regions act to silence (i.e. imprint) genes in *cis*. Imprinted genes are  
47 often part of imprinted gene networks, suggesting regulatory links between these  
48 genes<sup>3-5</sup>. More than 200 imprinted loci have been described in humans<sup>6</sup>, but there are  
49 likely many other, as yet undiscovered, imprinted loci.

50 Previous studies have utilized pedigrees to test maternal and paternal alleles  
51 separately for association with phenotypes or with gene expression to uncover new  
52 imprinted loci<sup>6-9</sup>. Kong *et al.*<sup>7</sup> discovered one locus associated with breast cancer risk  
53 only when the allele is inherited from the father and another locus associated with type 2  
54 diabetes risk only when the allele is inherited from the mother. Garg *et al.* reported parent-  
55 of-origin *cis*-eQTLs with known or putative novel imprinted genes affecting gene  
56 expression<sup>8</sup>. Two additional studies by Zoledziewsk *et al.* and Benonisdottir *et al.*

57 identified opposite POEs on adult height at known imprinted loci<sup>6,10</sup>. Both studies reported  
58 associations with variants at the *KCNQ1* gene, and one showed additional opposite POEs  
59 with height at two known imprinted loci (*IGF2-H19* and *DLK1-MEG3*)<sup>6</sup>. These studies  
60 provide proof-of-principle that alleles at imprinted loci can show POEs, some with  
61 opposite effects, with common phenotypes.

62         However, no previous study has included a broad range of human quantitative  
63 phenotypes or has considered whether genome-wide variants can have different effects  
64 depending on the parent of origin. To address this possibility, we developed a statistical  
65 model that directly compares the effects of the maternal and paternal alleles to identify  
66 effects that are different, including those that are opposite. We applied this model in a  
67 study of 21 common quantitative traits that were measured in the Hutterites, a founder  
68 population of European descent for which we have phased genotype data<sup>11</sup>. We  
69 identified variants with maternally inherited or paternally inherited effects only and  
70 variants with opposite POEs. Some of the identified regions have characteristics similar  
71 to known imprinted genes. Overall, we show that this model can identify putative novel  
72 imprinted regions with POEs for a broad range of clinically relevant quantitative  
73 phenotypes.

74

## 75 **RESULTS**

### 76 **GWAS**

77 We first performed standard genome-wide association studies (GWAS) of 21 traits in  
78 the Hutterites (**Supplementary Table 1**). These studies identified one genome wide  
79 significant association ( $p < 5 \times 10^{-8}$ ) with each of five of the 21 traits: low density  
80 lipoprotein level (LDL)-cholesterol, triglycerides, carotid artery intima media thickness  
81 (CIMT), left ventricular mass index (LVMI), and monocyte count. The results of all 21  
82 GWAS are summarized in **Supplementary Table 2** and **Supplementary Figs. 1 and 2**.  
83 Results for all variants for all GWAS are deposited in dbGaP (phs000185 – submission  
84 in progress).

### 85 **Parent of Origin GWAS**

86 We considered two possible mechanisms of POEs. In the first, the effect size of one  
87 parent's allele is close to zero and the effect size of the other parent's allele is  
88 significantly different from zero. For these cases, we performed a paternal only or  
89 maternal only GWAS. In other cases, the maternal and paternal alleles may both have  
90 effect sizes different from zero, but the effects are significantly different from each other  
91 or opposite in direction. To detect these types of POEs, we developed a model that  
92 tests for differences between parental effects (see Methods). This model is especially  
93 powerful to identify variants with parental effects in opposite directions.

### 94 Maternal and Paternal GWAS

95 Using the same phenotypes, genotypes, pedigree, and criteria for significance as in the  
96 standard GWAS, we tested for maternal and paternal effects on each trait by testing each  
97 parentally inherited allele with the trait of interest, similar to previous studies<sup>7,8,10</sup>. Variants

98 were considered to have POEs if they had a p-value less than  $5 \times 10^{-8}$  in only one parent  
99 and were not significant in the standard GWAS (i.e., the LDL association on chromosome  
100 19 and the triglycerides association chromosome 11 were not considered to have POEs;  
101 see **Supplementary Table 1**). The most significant parent of origin associations are  
102 summarized in **Table 1**. All significant results of the parent of origin GWAS for all 21  
103 phenotypes are included in **Supplementary Table 5**.

104 Overall, seven phenotypes had genome-wide significant parent of origin  
105 associations: four in the maternal only GWAS and three in the paternal only GWAS. Three  
106 cardiovascular disease (CVD)-associated phenotypes (age at menarche, CIMT, LVMI)  
107 and one lung function phenotype (forced expiratory volume in one second [FEV<sub>1</sub>]) were  
108 associated with maternally-inherited alleles only.

109 A maternally inherited allele at rs7184983 (G) on chromosome 16 was associated  
110 with younger age of menarche ( $P = 3.11 \times 10^{-8}$ ) (**Fig. 1**). This SNP, rs7184983, is located  
111 upstream of the *BBS2* gene and is associated with increased expression of *OGFOD1* in  
112 transformed fibroblast cells and tibial nerve<sup>12</sup>. The maternally inherited allele at rs4077567  
113 (G) on chromosome 2 was associated with decreased CIMT ( $P = 3.02 \times 10^{-8}$ )  
114 (**Supplementary Fig. 2**). This SNP is in the intron of a long intergenic noncoding gene,  
115 *LINC00607*, that is expressed in aorta, coronary, and tibial artery, all tissues potentially  
116 relevant to CIMT and atherosclerosis<sup>12</sup>. A maternally inherited allele at rs574232282 (G)  
117 in the intron of *SCMH1* on chromosome 1 was associated with increased LVMI ( $P = 1.39$   
118  $\times 10^{-8}$ ) (**Supplementary Fig. 3**). *SCMH1* is expressed in aorta, coronary, and tibial  
119 artery<sup>12</sup>. *SCMH1* protein associates with the polycomb group multiprotein complexes  
120 required to maintain the transcriptionally repressive state of certain genes<sup>12</sup>. Lastly,

121 maternally inherited alleles at rs9849387 (A) and rs6791779 (C) on chromosome 3 were  
122 both associated with reduced FEV<sub>1</sub> (P= 4.10x10<sup>-9</sup> and 1.48x10<sup>-8</sup>, respectively)  
123 (**Supplementary Fig. 4**). The nearest gene to rs9849387 is *ROBO2* (65kb, downstream),  
124 which is expressed in the lung as well as in brain, and ovary<sup>12</sup>. The nearest gene to  
125 rs6791779 is MIR4444-1(267kb) whose expression has not been characterized.

126 Three other CVD-related phenotypes (systolic blood pressure, LDL-C, and total  
127 cholesterol) had associations with paternally inherited alleles only. The paternally  
128 inherited allele at rs12024326 (A) on chromosome 1 was associated with lower LDL-  
129 cholesterol levels (P = 8.06x10<sup>-10</sup>) (**Figure 2**). rs12024326 is in the intron of gene *ADCK3*,  
130 and the same allele was associated with increased expression of *ADCK3* in whole blood,  
131 as well as decreased expression of a neighboring gene, *CDC42BPA* in brain  
132 (cerebellum), heart (left ventricle), esophagus, and tibial artery<sup>12</sup>. The paternal G allele at  
133 rs4843650 on chromosome 16 was associated with increased LDL-C and is located in  
134 the intron of *JPH3*, which is expressed predominantly in the brain<sup>12</sup>. A SNP on  
135 chromosome 13 (rs1536182) was associated with systolic blood pressure levels when it  
136 was inherited from the father (**Supplementary Fig. 5**). The paternally inherited A allele  
137 at this SNP was associated with decreased systolic blood pressure, as well as decreased  
138 expression of its closest gene, *LINC01055*, a long intergenic noncoding gene, in testis<sup>12</sup>.  
139 A paternally inherited allele at rs113588203 (G) on chromosome 1 was associated with  
140 lower total cholesterol (P = 1.76x10<sup>-8</sup>) (**Supplementary Fig. 6**). This SNP is intergenic  
141 between *RHOU* (96kb, downstream), which is expressed across multiple tissues, and  
142 *MIR4454* (331kb), which is expressed in adipose, kidney and heart tissues<sup>12</sup>.

143 GWAS for Differential Parent of Origin Effects

144 Because some imprinted regions include genes that have both maternal or paternal  
145 specific tissue expression, we next tested for such differential effects with these 21  
146 phenotypes. In these analyses, we compared the effect and direction of the association  
147 between maternal and paternal alleles to identify variants that have different effects,  
148 including opposite effects, on the phenotype. Such loci would be completely hidden in  
149 standard GWAS in which paternally and maternally inherited alleles are combined. These  
150 opposite effect GWAS revealed 11 independent loci with opposite POEs for nine different  
151 traits, at least six of which are associated with CVD risk (**Table 3, Supplementary Fig.**  
152 **7**).

153 A locus on chromosome 16, near the *CDH8* gene (128kb, upstream), was  
154 associated with opposite POEs with age of menarche (**Fig. 3**). *CDH8* is highly expressed  
155 in the brain, as well as in the aorta artery and pituitary gland. Two loci on chromosomes  
156 5 and 6 were associated with opposite POEs on body mass index (BMI) (**Fig. 4**). The  
157 most significant variant on chromosome 5 (rs77785972) is near a long intergenic  
158 noncoding gene, *LINC01340* (409kb, downstream), whose expression has not been well  
159 characterized. The SNP on chromosome 6 (rs17605739) is also in a long intergenic  
160 noncoding gene, *RP1-209A6.1*, which is expressed in low levels in the tibial artery,  
161 bladder, spleen, lung, pituitary gland, as well as testis.

162 A SNP on chromosome 16 (rs1032596) was associated with opposite POEs on  
163 LDL-cholesterol (**Supplementary Fig. 8**). This SNP lies in the intron of another long  
164 noncoding RNA gene, *LINC01081*, which has been suggested to be imprinted because  
165 its downstream genes have also been shown to have parent- and tissue-specific activity<sup>13</sup>.  
166 A region on chromosome 2 has opposite effects associated with LVMI (**Supplementary**



167 **Fig. 9).** The associated SNPs are in the intron of *XIRP2*, a cardiomyopathy associated  
168 protein that is expressed in skeletal muscle and heart left ventricle, suggesting that this  
169 gene could play a role in determining left ventricular mass<sup>12,14,15</sup>. In addition, the most  
170 significant SNP at this region, rs17616252 (and multiple SNPs in LD) is a strong eQTL (P  
171 =  $1.8 \times 10^{-13}$ ) for the gene *XIRP2* in skeletal muscle, *XIRP2-AS1* in testis, and *B3GALT1*  
172 in transformed fibroblast cells<sup>12</sup>. Four variants in a region on chromosome 1 in a  
173 microRNA gene, *MIR548F3*, were associated with opposite POEs on triglyceride levels  
174 (**Supplementary Fig. 10**). The expression of *MIR548F3* has not been characterized. SNP  
175 rs7033776 near *MELK* (27kb, downstream) on chromosome 9 was associated with  
176 opposite effects on total cholesterol (**Supplementary Fig. 11**). *MELK* is expressed in the  
177 colon and esophagus in addition to transformed lymphocytes and fibroblasts<sup>12</sup>.

178         Nine linked variants on chromosome 1 were associated with opposite POEs of  
179 blood eosinophil count (**Supplementary Fig. 12**). These variants are near the gene  
180 *IGSF21* (27kb, downstream) which is a member of the immunoglobulin superfamily and  
181 likely acts as a receptor in immune response pathways<sup>16</sup>. A variant on chromosome 3,  
182 rs12714812, was associated with opposite POEs for FEV<sub>1</sub> (**Supplementary Fig. 13**). This  
183 variant has been shown to regulate the expression of a gene *CNTN3* (45kb, upstream) in  
184 heart and brain<sup>12</sup>. Studies in mice have suggested that this gene is imprinted and  
185 maternally expressed in the murine placenta<sup>17</sup>. Variant rs142030841 in the intron of the  
186 gene *TPGS2* on chromosome 18 has opposite POEs with neutrophil levels  
187 (**Supplementary Fig. 14**). This SNP is an expression quantitative trait locus (eQTL) for  
188 the noncoding RNA gene *RP11-95O2.5* in skin, testis, breast, thyroid and adipose tissue,

189 for *CELF4* in tibial nerve and lung, and for *TPGS2* in tibial artery and transformed  
190 fibroblast cells<sup>12</sup>.

### 191 **Parent of Origin Effects on Gene Expression**

192 To determine if any of the associated variants also showed POEs on gene expression in  
193 the Hutterites, we used RNA-seq gene expression data from lymphoblastoid cell lines  
194 (LCLs) collected from 430 of the individuals in the GWAS sample. We first tested for  
195 association of maternal and paternal variants with genes detected as expressed in the  
196 LCLs and whose transcript start site was within 1Mb of each associated SNP. There  
197 were no significant associations after multiple testing correction, similar to a previous  
198 study<sup>6</sup>. However, because we considered this to be exploratory analyses, we show  
199 results for the five most significant parent of origin eQTLs (**Table 3**). We next used the  
200 opposite effect model for each SNP in Table 2 and expression of all genes that were  
201 detected as expressed in LCLs and whose transcript start site was within 1Mb of the  
202 associated SNP. This resulted in 57 tests (1 SNP for each of 8 phenotypes, and 57  
203 genes). The five most significant opposite effect eQTLs, none of which passed the  
204 Bonferroni threshold of  $8.77 \times 10^{-4}$ , are shown in **Table 4**. The most significant opposite  
205 effect eQTL was for *POLR1E* expression with the SNP on chromosome 9 (rs7033776)  
206 that was associated with total cholesterol (opposite effect eQTL  $P = 9.86 \times 10^{-4}$ )  
207 (**Supplementary Fig. 15**). *POLR1E* is involved in the purine metabolism pathway as  
208 well as DNA-directed polymerase activity. The same SNP, rs7033776, had modest  
209 opposite effects with the expression of three other genes in the region (*PAX5*, *FBXO10*,  
210 and *FRMPD1*), a signature consistent with an imprinted region. Another SNP with

211 opposite POEs on LVMI, rs16853098, was an opposite effect eQTL for *STK39*, a gene  
212 that has been previously associated with hypertension<sup>18</sup>.

## 213 **DISCUSSION**

214 In this study, we introduced a novel statistical method that allows assessment of standard  
215 GWAS signals along with measures of differential POEs on common quantitative  
216 phenotypes. Similar to previous parent of origin studies of fewer phenotypes, we tested  
217 for associations of maternally- or paternally-derived alleles with each phenotype. We then  
218 extended this method to identify variants for which maternally- and paternally-derived  
219 alleles have different, including opposite, effects on phenotypic values. The focus on 21  
220 common disease-associated phenotypes in a single large pedigree allowed us to broadly  
221 survey physiological effects of putative imprinted regions and the candidate genes at each  
222 associated locus. In contrast to previous studies, our new model can identify variants with  
223 opposite POEs that would be missed in traditional GWAS (**Table 2**).

224 Our studies of >1,000 Hutterites who are related to each other in a single pedigree  
225 allowed us to detect POEs, even when few genome-wide significant associations were  
226 detected in standard GWAS of the same phenotypes. Our method revealed parent of  
227 origin specific genome-wide significant associations for seven of the 21 phenotypes  
228 examined, with maternally-inherited alleles associated with four phenotypes, paternally-  
229 inherited alleles with three phenotypes (**Table 1**), and opposite parent of origin alleles  
230 with nine phenotypes, of which five also showed single POEs at different loci (**Table 2**).  
231 Overall, 11 of the 21 phenotypes examined showed genome-wide significant evidence of  
232 POEs with alleles at one or more loci. In contrast, standard GWAS of these same  
233 phenotypes and using the same markers in these same individuals revealed genome-  
234 wide significant association for only five traits.

235 It is notable that four of the nine significant opposite parent of origin effects (one  
236 each with LDL-C and triglycerides, and two with BMI) lie in or near long intergenic non-  
237 coding RNA genes (lincRNAs). LincRNAs are a feature of imprinted regions<sup>1</sup>, where they  
238 can silence the expression of genes on the opposite chromosome<sup>3,19</sup>. One of the variants,  
239 rs1032596, with an opposite parent of origin effect on LDL-C is located in the *LINC01081*  
240 gene. This noncoding RNA, along with *LINC01082*, regulates the *FOXF1* enhancer  
241 resulting in *FOXF1* parent- and tissue-specific activity<sup>13</sup> providing experimental support  
242 for tissue specific expression, a feature of imprinted regions.

243 Another variant with POEs in our study has been suggested to be imprinted in  
244 previously published work. The variant associated with opposite POEs for FEV<sub>1</sub> is an  
245 eQTL for the gene *CNTN3*. *CNTN3* was shown to have exclusive maternal allele-specific  
246 expression in murine placentas<sup>17</sup>, although this finding may have been due to  
247 contaminating maternal cells<sup>20,21</sup>.

248 Other regions associated with POEs harbor genes involved in transcriptional  
249 repression (e.g., *SCMH1* with LVMI on chromosome 1) or the associated SNPs are  
250 reported as eQTLs in GTEx with expression in tissues relevant to the phenotype under  
251 investigation (e.g., the LVMI-associated SNPs are eQTLs for *XIRP2*, which is expressed  
252 in skeletal muscle and heart left ventricle)<sup>12</sup>. Overall, these patterns of expression provide  
253 additional support that the parent of origin associations in our study are flagging imprinted  
254 regions or regions involved in the regulation of gene expression. Finally, we used gene  
255 expression in LCLs from the Hutterites to directly test for parent of origin eQTLs among  
256 SNPs associated with phenotypes in the parent of origin GWAS. Although none of the  
257 parent of origin eQTLs met criteria for significance after correcting for multiple testing, the

258 SNP on chromosome 9 with opposite POEs on total cholesterol levels was borderline  
259 significant as an opposite parent of origin eQTL for *POLR1E*, and possible for three other  
260 genes at the same locus (*PAX5*, *FBXO10*, and *FRMPD1*). The presence of multiple genes  
261 with potential parent of origin expression patterns is further supportive of an imprinted  
262 locus. The availability of gene expression only in LCLs from the Hutterites limits the  
263 inferences we can draw about effects on expression because imprinted regions are often  
264 tissue-specific and sometimes developmentally regulated<sup>1,2</sup>. Despite this limitation, the  
265 fact that many of the SNPs associated with POEs on phenotypes are themselves eQTLs  
266 in relevant tissues (GTEx) and some are suggestive of having POEs on expression in  
267 LCLs from the Hutterites is generally supportive of the suggestion that some of the regions  
268 identified in this study are imprinted in humans. Additionally, our data suggest that loci  
269 with POEs influence a broad spectrum of quantitative phenotypes that are themselves  
270 risk factors for common diseases.

271 In particular, the discovery of POEs for eight traits that are associated  
272 cardiovascular disease risk is intriguing. These include metabolic phenotypes, such as  
273 BMI, total cholesterol, triglycerides, LDL, and age of menarche, that have indirect effects  
274 on cardiac health, as well as LVMI and CIMT, which more directly reflect cardiac health.  
275 Some of these phenotypes showed associations with paternally inherited alleles only  
276 (systolic blood pressure, LDL-C, total cholesterol), maternally inherited alleles only (LVMI,  
277 CIMT, and age at menarche), and/or with opposite effect variants (BMI, LDL-C,  
278 triglycerides, total cholesterol, LVMI, age at menarche). It has been suggested that  
279 genomic imprinting evolved in the mammalian lineage as a way to regulate maternally  
280 and paternally expressed genes in the placenta during pregnancy and modulate

281 metabolic functions related to growth, where the parental interests may be in conflict –  
282 paternal alleles favoring growth of the fetus at the expense of the mother while maternal  
283 alleles favor restricting resources to the fetus to ensure preservation of her nutritional  
284 needs<sup>3,19,22</sup>. Our data show some effects that are consistent with this theory. For example,  
285 three independent paternally inherited alleles on chromosome 1 are associated with  
286 increased LDL-C (**Fig. 2**) and total cholesterol (**Supplementary Fig. 7**); a paternal allele  
287 on chromosome 13 is also associated with increased systolic blood pressure  
288 (**Supplementary Fig. 6**). However, it is not always possible to interpret our results in light  
289 of this model, such as the association of maternal allele on chromosome 2 with decreased  
290 CIMT (**Supplementary Fig. 3**), or the maternal allele on chromosome 16 associated with  
291 decreased age of menarche (**Fig. 1**), which confers increased cardiovascular risk<sup>23</sup>.  
292 However, because many of the traits associated with POEs in this study were measured  
293 in adults, and none were measured in neonates, we are likely observing the downstream  
294 effects of processes that occurred *in utero*. Nonetheless, this kinship theory, or parent-  
295 conflict hypothesis, could account for the enrichment of parent of origin associations,  
296 particularly those with opposite effects, among metabolic and CVD-associated traits<sup>1</sup>.

297 Finally, we note that the parent of origin GWAS for 21 phenotypes in the Hutterites  
298 revealed overall twice as many genome-wide significant loci compared to standard  
299 GWAS of the same phenotypes in the same individuals, suggesting that variation at  
300 imprinted loci may represent some of the “missing heritability” of these phenotypes and  
301 potentially for the disease for which they confer risk. This idea is consistent with  
302 observations in mice showing that POEs contribute disproportionately to the heritability of  
303 97 traits, including those related to total cholesterol, weight, HDL, and triglycerides<sup>24</sup>.

304 Exactly how much loci with POEs in humans contribute to phenotypic variation and  
305 disease risk overall remains to be determined, but our study provides compelling  
306 evidence that it is likely to be significant for many important traits.



## 307 **SUBJECTS AND METHODS:**

### 308 **Sample Composition**

309       The individuals in this study have participated in one or more of our studies on the  
310 genetics of complex traits in the Hutterites<sup>25-27</sup>. The more than 1,500 Hutterites in our  
311 study are related to each other in a 13-generation pedigree including 3,671 individuals.

### 312 **Genotype Data**

313 Variants detected in the whole genome sequences of 98 Hutterites were previously  
314 imputed to an additional 1,317 individuals using PRIMAL, a high-accuracy pedigree  
315 based imputation method<sup>28</sup>. PRIMAL was used to phase alleles and assign parent of  
316 origin for 83% of about 12 million autosomal SNPs. For these studies, we selected SNPs  
317 that had a MAF >1% and genotype call rate > 85%. This yielded 5,891,982 autosomal  
318 SNPs. Parent of origin allele call rates differed among individuals and between  
319 phenotypes (**Supplementary Table 1**).

### 320 **Phenotype Data**

321 We included 21 quantitative phenotypes that were previously measured in the Hutterites.  
322 Descriptions for each phenotype, as well as exclusion criteria, transformations, and  
323 covariates used with each phenotype in the GWAS, are available in the Supplementary  
324 Methods (**Supplementary Table 1**). Descriptions for 18 of the 21 phenotypes can be  
325 found in Cusanovich *et al.*<sup>25</sup> The remaining three are described here. Height was  
326 measured in cm on a stadiometer with shoes removed. BMI was calculated using weight  
327 (kg, measured on scale) divided by height (m) squared. Age at menarche was collected  
328 retrospectively by interview.

### 329 **GWAS.**

330 We used a linear mixed model as implemented in GEMMA to test for genome wide  
331 association with 21 phenotypes using an additive model. We corrected for relatedness,  
332 as well as relevant covariates (**Supplementary Table 1**).

### 333 **Maternal and Paternal GWAS.**

334 To evaluate the evidence for POEs, we tested maternal and paternal alleles separately  
335 with each phenotype, comparing phenotypic differences between the maternally inherited  
336 alleles and between the paternally inherited alleles. We used a linear mixed model as  
337 implemented in GEMMA, which allows us to correct for relatedness as a random effect,  
338 as well as sex, age, and other covariates as fixed effects<sup>29</sup>. The linear mixed model for  
339 the parent of origin GWAS for testing maternal alleles and paternal alleles is shown in  
340 Equation 1 and Equation 2, respectively.

$$341 \quad Y = \mathbf{W}\boldsymbol{\alpha} + X_M\beta_M + \mathbf{g} + \boldsymbol{\varepsilon} \quad (1)$$

$$342 \quad Y = \mathbf{W}\boldsymbol{\alpha} + X_P\beta_P + \mathbf{g} + \boldsymbol{\varepsilon} \quad (2)$$

343  $n$  is the number of individuals,  $Y$  is an  $n \times 1$  vector of quantitative traits,  $\mathbf{W}$  is an  $n \times c$   
344 matrix of covariates (fixed effects) including intercept 1.  $\boldsymbol{\alpha}$  is a  $c \times 1$  vector of covariate  
345 coefficients.  $X_M$  is an  $n \times 1$  vector of maternal alleles, and  $X_P$  an  $n \times 1$  vector of paternal  
346 alleles.  $\beta_M$  and  $\beta_P$  are the effect sizes of maternal and paternal alleles, respectively.  $\mathbf{g}$  is  
347 a vector of genetic effects with  $\mathbf{g} \sim N(0, \mathbf{A}\sigma_g^2)$  where  $\mathbf{A}$  is the genetic relatedness matrix;  
348  $\boldsymbol{\varepsilon}$  is a vector of non-genetic effects with  $\boldsymbol{\varepsilon} \sim N(0, \mathbf{I}\sigma_e^2)$ .

### 349 **Differential Effect GWAS (PO-GWAS).**

350 To test for a difference in the same allele inherited from each parent, including opposite  
351 effects, we re-parameterized the test model (Equation 3) from Garg *et al.*<sup>8</sup>. The null model  
352 (Equation 4) is a standard GWAS model, ignoring parent of origin of alleles. The test

353 model (Equation 3) is more significant when maternal and paternal alleles have  
354 differential effects on gene expression.

$$355 \quad Y = \mathbf{W}\boldsymbol{\alpha} + X_M\beta_M + X_P\beta_P + \mathbf{g} + \varepsilon \quad (3)$$

$$356 \quad Y = \mathbf{W}\boldsymbol{\alpha} + X_{PM}\beta_{PM} + \mathbf{g} + \varepsilon \quad (4)$$

357 This new model allows us to measure the difference in parental effect of the same allele  
358 when the genotype is a covariate in Equation 5.

$$359 \quad Y = \mathbf{W}\boldsymbol{\alpha} + \frac{(X_P - X_M)}{2}(\beta_P - \beta_M) + X_{PM}\frac{(\beta_P + \beta_M)}{2} + \mathbf{g} + \varepsilon \quad (5)$$

360  $X_{PM}$  is a  $n \times 1$  vector of genotypes with possible values [ 0, 1, 2 ], equivalent to  $X_P + X_M$ .  
361  $(\beta_P - \beta_M)$  is the difference in parental effect size. If the difference in parental effect size  
362 is large and significantly different from 0 it suggests a parent of origin effect exists at this  
363 variant.  $\frac{(X_P - X_M)}{2}$  is a  $n \times 1$  vector of genotypes with possible values [-1, 0, 1].  $\frac{(\beta_P + \beta_M)}{2}$  is  
364 the average parental effect size that is captured in normal GWAS using genotypes. The  
365 average genotypes are added in as a covariate, with the average parental effect size the  
366 corresponding covariate coefficient. This differential effect GWAS was tested in GEMMA  
367 using BIMBAM format to use average genotype values<sup>30</sup>.

### 368 **Parent of Origin eQTL Studies**

369 RNA-seq data from LCLs were available from a previous study in the Hutterites<sup>25</sup>. For this  
370 study, sequencing reads were reprocessed as follows. Reads were trimmed for adaptors  
371 using Cutadapt (with reads <5 bp discarded) then remapped to hg19 using STAR indexed  
372 with gencode version 19 gene annotations<sup>31,32</sup>. To remove mapping bias, reads were  
373 processed using WASP mapping pipeline<sup>33</sup>. Gene counts were collected using HTSeq-  
374 count<sup>34</sup>. VerifyBamID was used to identify sample swaps to include individuals that were  
375 previously excluded<sup>35</sup>. Genes mapping to the X and Y chromosome were removed; genes

376 with a Counts Per Million (CPM) value of 1 (expressed with less than 20 counts in the  
377 sample with lowest sequencing depth) were also removed. Limma was used to normalize  
378 and convert counts to log transformed CPM values<sup>36</sup>. Technical covariates that showed  
379 a significant association with any of the top principal components were regressed out  
380 (RNA Integrity Number and RNA concentration).

### 381 **Maternal and Paternal Parent of Origin eQTL**

382 LCL RNA-seq data was used to test the single parent model for the most significant SNP  
383 from the maternal or paternal only GWAS for each phenotype. We selected all genes  
384 detected as expressed in the LCLs and residing within 1Mb of each most significant  
385 associated SNP. Summary of the SNPs and genes tested are in **Supplementary Table**  
386 **3**.

### 387 **Differential Parent of Origin eQTL**

388 LCL RNA-seq data was used to test the opposite effect model for the most significant  
389 SNP in each region that was associated with a phenotype in the parent of origin opposite  
390 effects GWAS. We selected all genes detected as expressed in the LCLs and residing  
391 within 1Mb of each associated SNP. Summary of the SNPs and genes tested are in  
392 **Supplementary Table 4**.

393

394

395 **Conflict of Interest: None**

396 **Supplemental Data description**

397 **Supplementary Table 1.** Phenotypes and sample composition

398 **Supplementary Table 2.** GWAS results for all variants with p-value  $< 5 \times 10^{-08}$

399 **Supplementary Table 3.** Candidate Genes for Parent of Origin eQTL

400 **Supplementary Table 4.** Candidate Genes for Parent of Origin Differential Effect eQTL

401 **Supplementary Table 5a-c.** Parent of Origin GWAS results with p-value  $< 5 \times 10^{-08}$ .

402 Includes Maternal and Paternal GWAS and Differential Effect GWAS

403 **Supplementary Figure 1.** Manhattan and QQ Plots from Standard GWAS of 21

404 Quantitative Phenotypes

405 **Supplementary Figures 2-6.** Maternal and Paternal GWAS results for CIMT, LVMI,

406 FEV<sub>1</sub>, LVMI, SBP, and Total Cholesterol

407 **Supplementary Figure 7.** Manhattan and QQ Plots from Differential Effect GWAS of 21

408 Quantitative Phenotypes

409 **Supplementary Figures 8-14.** Differential Parent of Origin GWAS results for LDL,

410 LVMI, Triglycerides, Total Cholesterol, Blood Eosinophil Count, FEV<sub>1</sub>, and Neutrophil

411 Count

412 **Supplementary Figure 15.** Differential Effect eQTL for rs7033776

413

414 **Acknowledgments**

415 We thank Catherine Stanhope for help with processing phenotype data, Mark Abney

416 and members of the Ober lab for useful discussions, Joe Urbanski and Lorenzo Pesce

417 for assistance using Beagle, the many members of our field trip teams for help in

418 phenotyping and collecting and processing samples, and the Hutterites for their  
419 continued support of our studies. This work was supported by NIH grants HL085197  
420 and HD21244; and in part by NIH through resources provided by the Computation  
421 Institute and the Biological Sciences Division of the University of Chicago and Argonne  
422 National Laboratory, under grant 1S10OD018495-01. S.V.M was supported by NIH  
423 Grant T32 GM007197 and the Ruth L. Kirschstein NRSA Award F31HL134315.

424 **Web Resources.**

425 Code for PO-GWAS: [https://github.com/smozaffari/PO\\_GWAS](https://github.com/smozaffari/PO_GWAS)

426 **Author Contribution**

427 S.V.M., D.L.N., and C.O. designed the study and wrote the paper. J.M.D., S.J.S., and  
428 R.M.L provided clinical data. S.V.M. performed analyses. All authors discussed results  
429 and commented on the manuscript.

430 **REFERENCES**

- 431 1. Peters, J. The role of genomic imprinting in biology and disease: an expanding  
432 view. *Nature Reviews Genetics* **15**, 517–530 (2014).
- 433 2. Falls, J. G., Pulford, D. J., Wylie, A. A. & Jirtle, R. L. Genomic imprinting:  
434 implications for human disease. *Am. J. Pathol.* **154**, 635–647 (1999).
- 435 3. Patten, M. M., Cowley, M., Oakey, R. J. & Feil, R. Regulatory links between  
436 imprinted genes: evolutionary predictions and consequences. *Proc. Biol. Sci.* **283**,  
437 (2016).
- 438 4. Gabory, A. *et al.* H19 acts as a trans regulator of the imprinted gene network  
439 controlling growth in mice. *Development* **136**, 3413–3421 (2009).
- 440 5. Varrault, A. *et al.* *Zac1* regulates an imprinted gene network critically involved in  
441 the control of embryonic growth. *Developmental Cell* **11**, 711–722 (2006).
- 442 6. Benonisdotir, S. *et al.* Epigenetic and genetic components of height regulation.  
443 *Nat Comms* **7**, 13490 (2016).
- 444 7. Kong, A. *et al.* Parental origin of sequence variants associated with complex  
445 diseases. *Nature Publishing Group* **462**, 868–874 (2009).
- 446 8. Garg, P., Borel, C. & Sharp, A. J. Detection of Parent-of-Origin Specific  
447 Expression Quantitative Trait Loci by Cis-Association Analysis of Gene  
448 Expression in Trios. *PLoS ONE* **7**, e41695 (2012).
- 449 9. Perry, J. R. B. *et al.* Parent-of-origin-specific allelic associations among 106  
450 genomic loci for age at menarche. *Nature Publishing Group* **514**, 92–97 (2014).
- 451 10. Zoledziewska, M. *et al.* Height-reducing variants and selection for short stature in  
452 Sardinia. *Nat Genet* **47**, 1352–1356 (2015).
- 453 11. Livne, O. E. *et al.* PRIMAL: Fast and Accurate Pedigree-based Imputation from  
454 Sequence Data in a Founder Population. *PLoS Computational Biology* **11**,  
455 e1004139–14 (2015).
- 456 12. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx)  
457 pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660  
458 (2015).
- 459 13. Szafranski, P. *et al.* Pathogenetics of alveolar capillary dysplasia with  
460 misalignment of pulmonary veins. *Human Genetics* **135**, 569–586 (2016).
- 461 14. Wang, Q., Lin, J. L.-C., Erives, A. J., Lin, C.-I. & Lin, J. J.-C. New insights into the  
462 roles of Xin repeat-containing proteins in cardiac development, function, and  
463 disease. *Int Rev Cell Mol Biol* **310**, 89–128 (2014).
- 464 15. Nilsson, M. I. *et al.* Xin is a marker of skeletal muscle damage severity in  
465 myopathies. *Am. J. Pathol.* **183**, 1703–1709 (2013).
- 466 16. O'Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current  
467 status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*  
468 **44**, D733–45 (2016).
- 469 17. Brideau, C. M., Eilertson, K. E., Hagarman, J. A., Bustamante, C. D. & Soloway,  
470 P. D. Successful computational prediction of novel imprinted genes from  
471 epigenomic features. *Mol. Cell. Biol.* **30**, 3357–3370 (2010).
- 472 18. Wang, Y. *et al.* From the Cover: Whole-genome association study identifies  
473 STK39 as a hypertension susceptibility gene. *Proceedings of the National*  
474 *Academy of Sciences of the United States of America* **106**, 226–231 (2009).



- 475 19. Barlow, D. P. & Bartolomei, M. S. Genomic imprinting in mammals. *Cold Spring*  
476 *Harb Perspect Biol* **6**, (2014).
- 477 20. Okae, H. *et al.* Re-investigation and RNA sequencing-based identification of  
478 genes with placenta-specific imprinted expression. *Human Molecular Genetics* **21**,  
479 548–558 (2012).
- 480 21. Proudhon, C. & Bourc'his, D. Identification and resolution of artifacts in the  
481 interpretation of imprinted gene expression. *Brief Funct Genomics* **9**, 374–384  
482 (2010).
- 483 22. Haig, D. The kinship theory of genomic imprinting. *Annual review of ecology and*  
484 *systematics* (2000).
- 485 23. Canoy, D. *et al.* Age at menarche and risks of coronary heart and other vascular  
486 diseases in a large UK cohort. *Circulation* **131**, 237–244 (2015).
- 487 24. Mott, R. *et al.* The Architecture of Parent-of-Origin Effects in Mice. *Cell* **156**, 332–  
488 342 (2014).
- 489 25. Cusanovich, D. A. *et al.* Integrated analyses of gene expression and genetic  
490 association studies in a founder population. *Human Molecular Genetics* **25**, 2104–  
491 2112 (2016).
- 492 26. Weiss, L. A., Abney, M., Cook, E. H. & Ober, C. Sex-specific genetic architecture  
493 of whole blood serotonin levels. *The American Journal of Human Genetics* **76**,  
494 33–41 (2005).
- 495 27. Abney, M., McPeck, M. S. & Ober, C. Broad and narrow heritabilities of  
496 quantitative traits in a founder population. *The American Journal of Human*  
497 *Genetics* **68**, 1302–1307 (2001).
- 498 28. Livne, O. E. *et al.* PRIMAL: Fast and Accurate Pedigree-based Imputation from  
499 Sequence Data in a Founder Population. *PLoS Computational Biology* **11**,  
500 e1004139–14 (2015).
- 501 29. Zhou, X. & Stephens, M. Genome-wide efficient mixed-model analysis for  
502 association studies. *Nat Genet* **44**, 821–824 (2012).
- 503 30. Servin, B. & Stephens, M. Imputation-based analysis of association studies:  
504 candidate regions and quantitative traits. *PLoS Genet* **3**, e114 (2007).
- 505 31. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–  
506 21 (2013).
- 507 32. Martin, M. Cutadapt removes adapter sequences from high-throughput  
508 sequencing reads. *EMBnet.journal* **17**, pp. 10–12 (2011).
- 509 33. van de Geijn, B., McVicker, G., Gilad, Y. & Pritchard, J. K. WASP: allele-specific  
510 software for robust molecular quantitative trait locus discovery. *Nat Meth* **12**,  
511 1061–1063 (2015).
- 512 34. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-  
513 throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
- 514 35. Jun, G. *et al.* Detecting and estimating contamination of human DNA samples in  
515 sequencing and array-based genotype data. *American journal of human genetics*  
516 **91**, 839–848 (2012).
- 517 36. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-  
518 sequencing and microarray studies. *Nucleic Acids Research* **43**, e47 (2015).  
519  
520



521 **TABLES**

522

523

Phenotype	rsid (Effect allele/ Other allele)	chr:loc	Variant Location	Nearest Gene	MAF	N	Beta (SE)	Paternal GWAS p-value	Maternal GWAS p-value	Standard GWAS p-value
<b>A. Maternal Associations</b>										
<b>Age at menarche</b>	rs7184983 (A/G)	16:56554709	Upstream	<i>BBS2</i>	0.059	336	0.862 (0.154)	5.01E-01	3.11E-08	6.75E-03
<b>CIMT</b>	rs4077567 (G/A)	2:216703202	intronic	<i>LINC00607*</i>	0.30	429	0.047 (0.008)	5.72E-01	3.02E-08	4.21E-06
<b>FEV<sub>1</sub></b>	rs9849387 (A/G)	3:77764243	Intergenic	<i>ROBO2</i>	0.39	1029	-0.089 (0.015)	3.87E-01	4.10E-09	4.38E-04
	rs6791779 (C/G)	3:74996505	Intergenic	<i>MIR4444-1*</i>	0.24	879	-0.102 (0.021)	6.88E-02	1.48E-08	4.52E-02
<b>LVMI</b>	rs574232282 (G/A)	1:41662388	Intronic	<i>SCMH1</i>	0.018	537	0.239 (0.042)	5.52E-01	1.39E-08	1.05E-03
<b>B. Paternal Associations</b>										
<b>LDL</b>	rs12024326 (A/G)	1:227146433	Intronic	<i>ADCK3</i>	0.175	686	-0.295 (0.048)	8.06E-10	4.21E-01	4.24E-05
	rs4843650 (A/G)	16:87683486	Intronic	<i>JPH3</i>	0.448	621	0.211 (0.036)	6.57E-09	2.21E-01	1.50E-04
<b>SBP</b>	rs1536182 (A/G)	13:46275415	Upstream	<i>LINC01055*</i>	0.2	684	-0.028 (0.005)	1.53E-08	1.78E-01	6.93E-04
<b>Total cholesterol</b>	rs113588203 (G/T)	1:228979156	Intergenic	<i>RHOV</i>	0.099	703	-0.341 (0.060)	1.76E-08	7.43E-02	8.08E-03

524 **Table 1. Phenotypes with significant single parent of origin associations. The**

525 most significant variant ( $P < 5 \times 10^{-8}$ ) at each locus for the (A) maternal and (B) paternal

526 associations associated with each phenotype is shown.

527 \*non-coding RNA genes

Phenotype	rsid	chr:loc	Variant Location	Nearest Gene	MAF	$\beta_M - \beta_P$ (SE)	Opposite Effect GWAS	Paternal GWAS		Maternal GWAS		Standard GWAS p-value
								P-value	Beta (SE)	P-value	Beta (SE)	
<b>Age of menarche</b>	rs12447191	16:62199299	intergenic	<i>CDH8</i>	0.17	-0.654 (0.109)	5.27E-09	5.20E-06	0.391 (0.085)	1.85E-05	-0.368 (0.085)	8.68E-01
<b>BMI</b>	rs77785972	5:97415767	intergenic	<i>LINC01340*</i>	0.025	0.154 (0.025)	5.12E-10	5.84E-07	-0.094 (0.019)	1.58E-05	0.081 (0.019)	5.39E-01
	rs17605739	6:22962798	intronic	<i>RP1-209A6.1*</i>	0.17	0.053 (0.010)	3.01E-08	6.99E-05	-0.032 (0.008)	1.42E-06	0.034 (0.007)	1.56E-01
<b>eosinophil</b>	rs2355879	1:18732860	intergenic	<i>IGSF21</i>	0.14	0.091 (0.016)	1.69E-08	5.83E-08	-0.065 (0.012)	5.59E-04	0.043 (0.012)	2.53E-01
<b>FEV<sub>1</sub></b>	rs12714812	3:74813002	intergenic	<i>CNTN3</i>	0.45	-0.119 (0.021)	4.52E-08	1.78E-03	0.052 (0.017)	6.35E-06	-0.073 (0.016)	9.58E-01
<b>LDL</b>	rs1032596	16:86281537	Intronic	<i>LINC01081*</i>	0.30	-0.310 (0.056)	3.69E-08	1.05E-06	0.201 (0.041)	4.56E-04	-0.148 (0.042)	2.71E-01
<b>LVMI</b>	rs16853098	2:168013281	intronic	<i>XIRP2</i>	0.12	-0.091 (0.053)	4.18E-08	5.29E-06	0.064 (0.014)	2.04E-04	-0.048 (0.013)	9.26E-01
<b>neutrophils</b>	rs142030841	18:34371947	intronic	<i>TPGS2</i>	0.042	-0.224 (0.041)	4.40E-08	2.25E-03	0.078 (0.025)	1.30E-07	-0.188 (0.035)	5.77E-01
<b>Triglycerides</b>	rs7525463	1:218860879	intronic	<i>MIR548F3*</i>	0.16	-0.401 (0.071)	2.51E-08	1.14E-03	0.195 (0.060)	5.52E-08	-0.267 (0.049)	2.84E-02
<b>Total cholesterol</b>	rs7033776	9:36704465	intergenic	<i>MELK</i>	0.41	0.230 (0.041)	4.12E-08	5.60E-08	-0.183 (0.034)	2.28E-03	0.099 (0.032)	6.70E-02

528 **Table 2. Significant Opposite Parent of Origin Effect GWAS Associations.** The

529 most significant variant at each locus for each phenotype is shown.  $\beta_M - \beta_P$  represents

530 difference in parental effect size.

531 \*non-coding RNA genes

Phenotype	Sample Size	rsid	chr:loc	Gene	Beta (SE)	Maternal eQTL p-value	Paternal eQTL p-value
<b>A. Maternal Associations</b>							
<b>CIMT</b>	334	rs4077567	1:216703202	<i>ABCA12</i>	0.039 (0.017)	2.14E-02	1.53E-02
<b>Age at menarche</b>	336	rs7184983	16:56554709	<i>POLR2C</i>	-0.085 (0.039)	2.91E-02	7.93E-01
<b>Age at menarche</b>	336	rs7184983	16:56554709	<i>SLC12A3</i>	-0.064 (0.031)	3.77E-02	2.28E-01
<b>CIMT</b>	334	rs4077567	1:216703202	<i>RPL37A</i>	0.030 (0.016)	5.72E-02	5.90E-01
<b>LVM</b>	457	rs74232282	1:41662388	<i>SMAP2</i>	1.40 (0.159)	5.82E-02	1.12E-01
<b>B. Paternal Associations</b>							
<b>Total cholesterol</b>	352	rs113588203	1:228979165	<i>HIST3H2A</i>	0.560 (0.308)	8.81E-01	6.85E-02
<b>Total cholesterol</b>	352	rs113588203	1:228979165	<i>SPHAR</i>	0.073 (0.047)	6.01E-01	1.20E-01
<b>LDL</b>	352	rs1110603	16:87687317	<i>MAP1LC3B</i>	-0.024 (0.015)	4.35E-01	1.25E-01
<b>LDL</b>	352	rs1110603	16:87687317	<i>FBXO31</i>	-0.027 (0.018)	1.56E-01	1.36E-01
<b>Total cholesterol</b>	357	rs113588203	1:228979165	<i>RAB4A</i>	-0.039 (0.028)	6.16E-01	1.59E-01

532

533 **Table 3. Parent of Origin eQTLs in LCLs.** The most significant SNP for each  
534 phenotype (**Table 1**) was tested for association with gene expression for genes with  
535 TSS within 1Mb of the SNP. The effect sizes correspond to the maternal (A) or paternal  
536 (B) effect sizes.

537

Phenotype	Sample Size	rsid	chr:loc	Gene	$\beta_M - \beta_P$ (SE)	Opposite Effect p-value
Total cholesterol	381	rs7033776	9:36704465	<i>POLR1E</i>	0.0603 (0.399)	9.86E-04
Total cholesterol	381	rs7033776	9:36704465	<i>PAX5</i>	0.0608 (0.0253)	0.0162
Total cholesterol	381	rs7033776	9:36704465	<i>FBXO10</i>	0.0789 (0.0337)	0.019
LVMI	355	rs16853098	2:168013281	<i>STK39</i>	-0.238 (0.124)	0.055
Total cholesterol	381	rs7033776	9:36704465	<i>FRMPD1</i>	0.185 (0.0988)	0.060

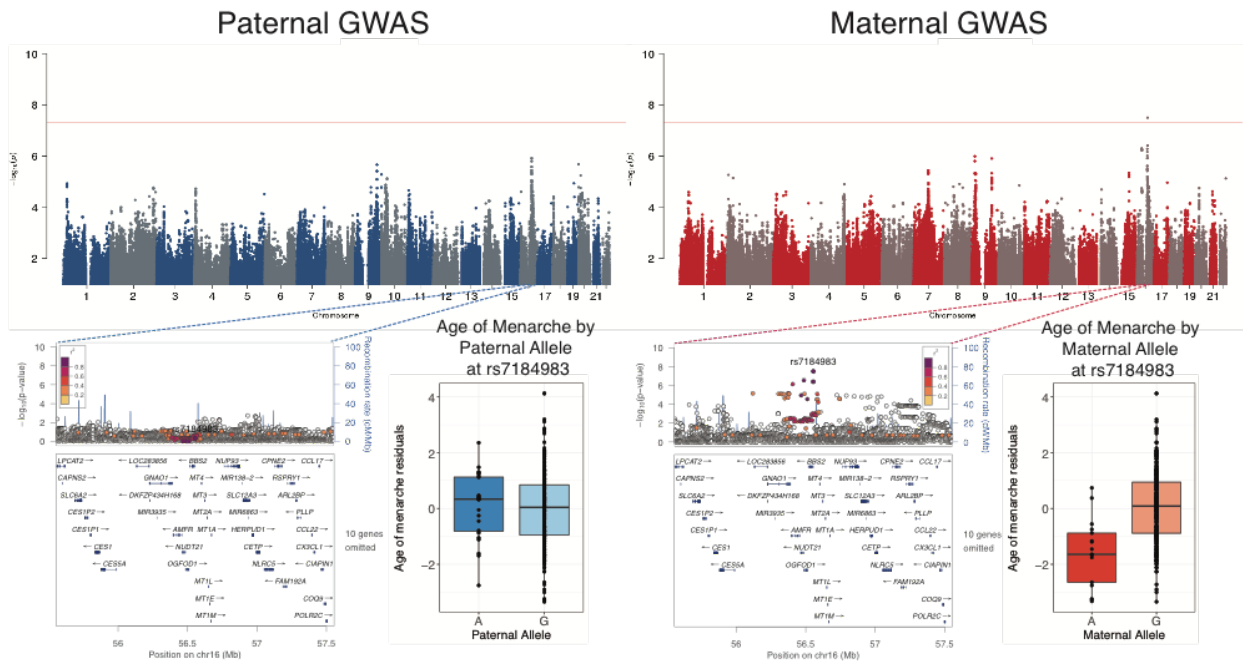
538

539 **Table 4. Opposite Parent of Origin eQTLs in LCLs.** The most significant SNP for  
540 each phenotype (**Table 2**) was tested for opposite effect eQTLs with genes with TSS  
541 within 1Mb of the SNP. The effect size corresponds to the difference in maternal and  
542 paternal effect sizes.

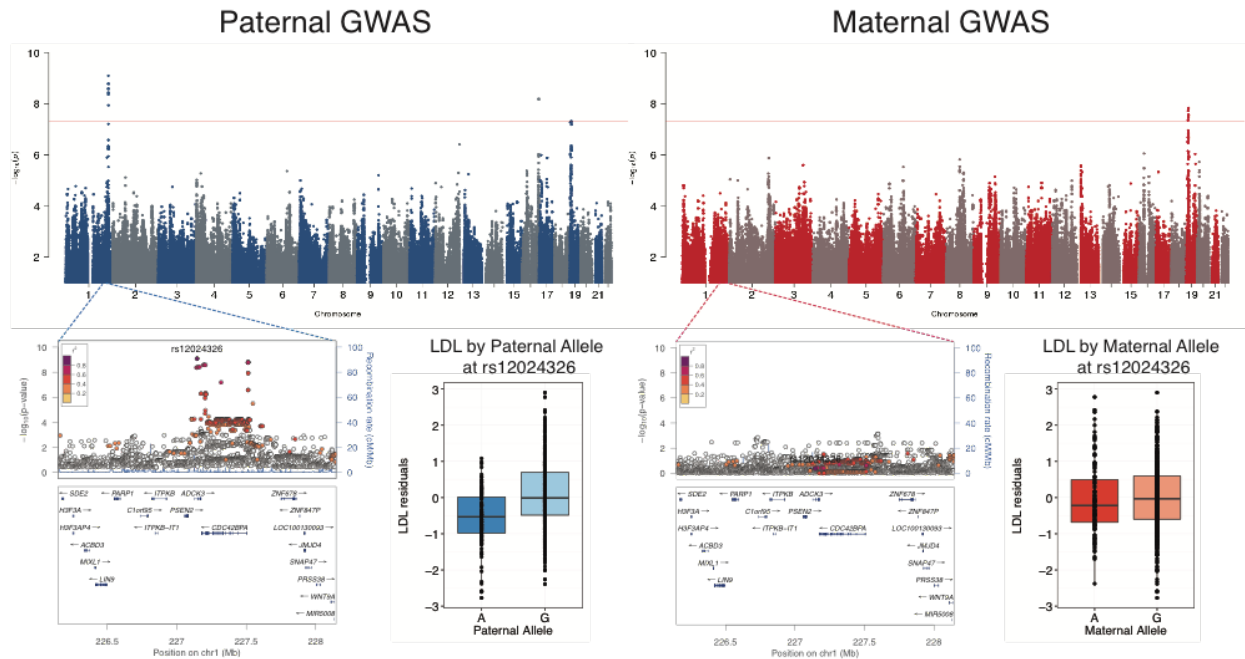
543

544

545 **FIGURES**  
546



547  
548 **Figure 1. Maternal and Paternal GWAS results for Age of Menarche.** The top panel  
549 shows the Manhattan plots from the maternal (left) and maternal (right) GWAS.  
550 LocusZoom plots for both GWAS are shown in the lower panel for the associated region  
551 in the GWAS. Boxplots show the distribution of age of menarche residuals (y-axes) by  
552 the corresponding maternal and paternal alleles at this SNP (x-axes). The horizontal bar  
553 of the boxplot shows the median, the box delineates the first and third quartile, and the  
554 whiskers show  $\pm 1.5 \times \text{IQR}$ .



555

556 **Figure 2. Maternal and Paternal GWAS results for LDL Cholesterol.** The top panel

557 shows the Manhattan plots from the maternal (left) and maternal (right) GWAS.

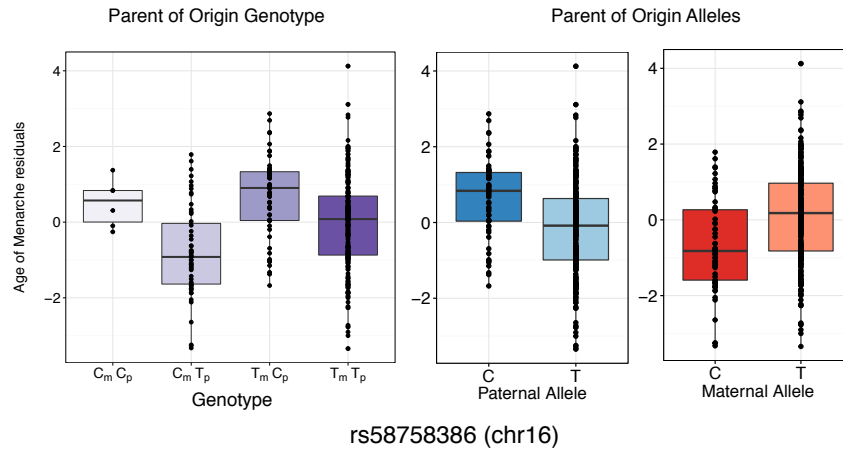
558 LocusZoom plots for both GWAS are shown in the lower panel for the associated region

559 in the GWAS. Boxplots show the distribution of LDL residuals (y-axes) by the

560 corresponding maternal and paternal alleles at this SNP (x-axes). The horizontal bar of

561 the boxplot shows the median, the box delineates the first and third quartile, and the

562 whiskers show  $\pm 1.5 \times$  IQR.



563

564 **Figure 3. Opposite Effect Parent of Origin GWAS Result for Age of Menarche.** Box

565 plots of age of menarche residuals (y-axes) are shown for each of the four genotypes

566 (left panel; x-axis), and for paternal (center panel; x-axis) and maternal (right panel; x-

567 axis) alleles. The maternal C allele is associated with decreased and maternal T allele

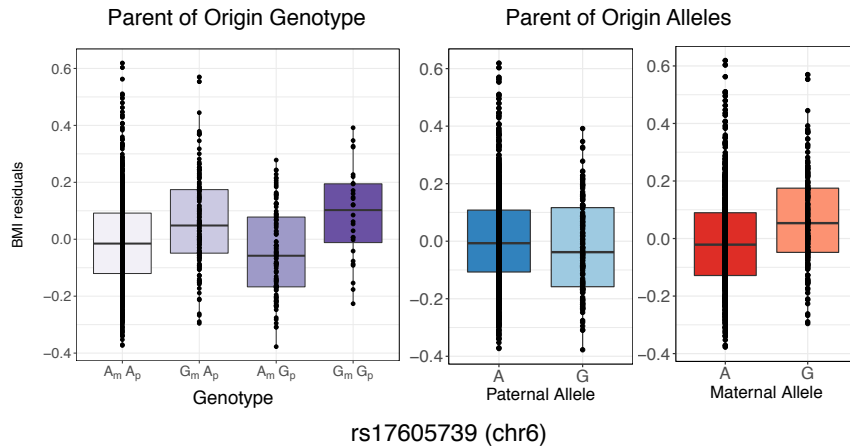
568 with increased age of menarche. The paternal C allele is associated with increased and

569 the paternal T allele with decreased age of menarche. The horizontal bar of the boxplot

570 shows the median, the box delineates the first and third quartile, and the whiskers show

571  $\pm 1.5 \times$  IQR.

572



573

574 **Figure 4. Opposite Effect Parent of Origin GWAS Result for BMI.** Box plots of two  
575 significant loci plot BMI residuals (y-axes) for each of the four genotypes (left panel; x-  
576 axis), and for paternal (center panel; x-axis) and maternal (right panel; x-axis) alleles.  
577 For the **a.** SNP on chromosome 5 the maternal A allele is associated with decreased  
578 and maternal G allele with increased BMI. The paternal A allele is associated with  
579 increased and the paternal G allele with decreased BMI. For the **b.** SNP on  
580 chromosome 6 the maternal A allele is associated with decreased and maternal G allele  
581 with increased BMI. The paternal A allele is associated with increased and the paternal  
582 G allele with decreased BMI. The horizontal bar of the boxplot shows the median, the  
583 box delineates the first and third quartile, and the whiskers show  $\pm 1.5 \times$  IQR.