

## 1 **The Phenome-wide Consequences of Anorexia Nervosa Genes**

2 Jessica S. Johnson, MPH<sup>1,2</sup>, Alanna C. Cote<sup>1,2</sup>, Amanda Dobbyn, PhD<sup>1-3</sup>, Laura G. Sloofman, MS<sup>4</sup>, Jiayi  
3 Xu, PhD<sup>1,2</sup>, Liam Cotter<sup>1,2</sup>, Alexander W. Charney, MD, PhD<sup>1-3,5,6</sup>, Eating Disorders Working Group of the  
4 Psychiatric Genomics Consortium<sup>\*</sup>, Jennifer Jordan, PhD<sup>7</sup>, Martin Kennedy, PhD<sup>7</sup>, Mikael Landén, MD,  
5 PhD<sup>8,9</sup>, Nicholas G Martin, PhD<sup>10</sup>, Preben Bo Mortensen, MD, DMSc<sup>11</sup>, Cynthia M. Bulik, PhD<sup>9,12,13</sup>, Laura  
6 M. Huckins, PhD<sup>1-6†</sup>

7

8 <sup>1</sup>Pamela Sklar Division of Psychiatric Genomics, Icahn School of Medicine at Mount Sinai, New York, NY  
9 10029, USA

10 <sup>2</sup>Department of Genetics and Genomics, Icahn School of Medicine at Mount Sinai, New York, NY 10029,  
11 USA

12 <sup>3</sup>Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York,  
13 NY 10029, USA

14 <sup>4</sup>Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New  
15 York, NY 10029, USA

16 <sup>5</sup>Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

17 <sup>6</sup>Mental Illness Research, Education and Clinical Centers, James J. Peters Department of Veterans  
18 Affairs Medical Center, Bronx, NY 14068, USA

19 <sup>7</sup>University of Otago, Department of Psychological Medicine, Christchurch School of Medicine & Health  
20 Sciences, 2 Riccarton Avenue, PO Box 4345, 8140 Christchurch, New Zealand

21 <sup>8</sup>Gothenburg University, Box 100, SE-405 30 Gothenburg, Sweden

22 <sup>9</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden

23 <sup>10</sup>QIMR Berghofer Medical Research Institute, Locked Bag 2000, Royal Brisbane Hospital QLD 4029,  
24 Australia

25 <sup>11</sup>Aarhus University, Norde Ringgade 1, 8000 Aarhus, Denmark

26 <sup>12</sup>Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27517, USA

27 <sup>13</sup>Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC 27517, USA

28

---

\* Full list of PGC-ED working group authors is available in the supplementary materials.

† Corresponding author: Dr. Laura M. Huckins, Icahn School of Medicine at Mount Sinai, 1425 Madison Ave., 3<sup>rd</sup> Floor, Room L3-70i, New York, NY 10029; [laura.huckins@mssm.edu](mailto:laura.huckins@mssm.edu), (212) 659-1613.

29 **ABSTRACT**

30

31 Anorexia nervosa (AN) is a psychiatric disorder with complex etiology, with a significant portion of disease  
32 risk imparted by genetics. Traditional GWAS studies produce principal evidence for the association of  
33 genetic variants with disease, and provide a jumping-off point for downstream functional analyses.

34 Transcriptomic imputation (TI) allows for the translation of SNPs into regulatory mechanisms, which can  
35 then be used to assess the functional outcome of genetically regulated gene expression (GReX) in a  
36 more broad setting through the use of phenome-wide association studies (PheWAS) in large and diverse  
37 clinical biobank populations with electronic health record (EHR) phenotypes. Here, we applied TI using S-  
38 PrediXcan to translate the most recent PGC-ED AN GWAS findings into AN-GReX. For significant genes,  
39 we imputed AN-GReX in the Mount Sinai BioMe™ Biobank and performed PheWAS on over 2000 clinical  
40 outcomes to test the clinical consequences of aberrant expression of these genes. We performed a  
41 secondary analysis to assess the impact of BMI on AN-GReX clinical associations.

42

43 Our S-PrediXcan analysis identified 47 genes associated with AN, including what is, to our knowledge,  
44 the first genetic association of AN with the Major Histocompatibility Complex (MHC). AN-GReX was  
45 associated with autoimmune, anthropometric, metabolic, psychiatric and gastrointestinal diagnoses in our  
46 biobank cohort, as well as measures of anthropometry, substance use, and pain score. Our analyses  
47 reveal that AN-GReX associations with measures of weight and substance use are modified by BMI, and  
48 indicate potential avenues of functional mechanism to investigate further.

49

50

## 51 INTRODUCTION

52

53 Anorexia nervosa (AN) is a severe eating disorder characterized by extreme low body weight and fear of  
54 gaining weight, and is often accompanied by compensatory behaviors to lose weight such as dietary  
55 restriction, purging, and higher than average physical activity(1). With a lifetime prevalence of 0.9-4%, AN  
56 has high mortality, with few effective treatments and an increased risk of suicide(2–4). Clinical and  
57 epidemiological studies have shown higher risk of other psychiatric disorders such as OCD, anxiety, and  
58 major depression in individuals with AN(3,5), and higher risk of substance use disorders in AN subtypes  
59 (6), as well as higher risk of metabolic and autoimmune disorders such as type 1 diabetes(7,8). Twin  
60 studies have established the heritability of AN to be between 50 to 60%, indicating a considerable  
61 contribution of genetic factors to AN disease liability(9,10).

62

63 Both metabolic and psychiatric factors contribute to anorexia nervosa. Epidemiological studies indicate  
64 high rates of psychiatric comorbidity, particularly with depression, anxiety disorders, and OCD(5), and  
65 many of these disorders share substantial symptomatology and environmental risk factors(11). Evidence  
66 is mounting that AN may lie at the extreme of a BMI spectrum, with obesity lying at the opposite end, and  
67 that AN may be a disorder of “fundamental metabolic dysregulation”(2). In line with this hypothesis,  
68 individuals with AN are able to reach and maintain incredibly low body weight, and often revert to this  
69 weight even after re-nourishment treatment, possibly indicating a fundamentally low BMI biological  
70 settling point(12,13). Additionally, individuals with AN have been shown to have lower concentrations of  
71 leptin, a hormone involved in regulation of body weight and other metabolic processes(13–16).

72 Microbiome studies have revealed abnormal microflora in the gut of AN patients(17), and evidence from  
73 many studies have pointed to a possible brain-gut etiology of AN(9,18–20).

74

75 Genetic studies of AN have provided further evidence of both psychiatric and metabolic etiology. The  
76 largest AN GWAS to date ( $N_{\text{Cases}}=16992$ ) uncovered eight loci associated with AN risk(9), and determined  
77 SNP-based heritability ( $h^2_{\text{SNP}}$ ) of 11-17%, similar to other psychiatric disorders, indicating that common  
78 variants contribute to polygenic risk of AN. In addition, genetic correlations demonstrate significant

79 overlap with other psychiatric disorders, including schizophrenia(9,21), MDD(9), OCD(9), alcohol use  
80 disorder(22) and anxiety disorders(9), as well as physical activity(9,23), indicating shared genetic  
81 variation between these traits. Significant negative genetic correlation between AN and anthropometric  
82 and metabolic traits, such as BMI(9,21,23), fat mass(9,21,23), obesity(9), type 2 diabetes(9), insulin  
83 resistance(9,21), fasting insulin(9,21) and leptin(9) have also been observed, further indicating metabolic  
84 components to AN disease risk. One study of AN polygenic risk scores (PRS) has indicated additional  
85 genetic associations of AN risk variants with anthropometric, behavioral, and psychiatric traits(24).

86  
87 Traditional observational genetic studies provide substantial data for factors contributing to risk of a  
88 particular disease, but may not capture the genetic factors that are specifically contributing to  
89 subthreshold or prodromal disease states. Electronic Health Records (EHR) contain information about an  
90 individual's health history such as diagnoses, medications, laboratory tests, vital signs, and family medical  
91 history, and can be utilized for scalable disease research. The vast amounts of data from the EHR can be  
92 queried to provide an understanding of the clinical spectrum of disease and disease progression across  
93 the lifetime of the patient. The initiation of biobanks to collect biological specimens from large numbers of  
94 patients within a healthcare system allows us to connect massive amounts of phenotype data from the  
95 EHR to genetic data. A phenome-wide association study or PheWAS is an effective method of querying  
96 the EHR to look for the associations of a trait of interest with the clinical phenome(25–27). Previous  
97 studies have demonstrated the efficacy of PheWAS methods in determining clinical outcomes associated  
98 with genomic disease risk variants, and have replicated associations from GWASs of strictly defined  
99 disease traits, including psychiatric disorders(28–30). Exploring the phenotypic associations with AN  
100 genetic architecture has the potential to clarify how some of these GWAS variants functionally contribute  
101 to AN disease risk, symptomatology, and clinical presentation.

102  
103 Transcriptomic imputation (TI) approaches leverage expression quantitative trait loci (eQTLs) derived  
104 from large, well-curated gene expression datasets (for example, the Genotype-Tissue Expression project  
105 (GTEx)(31), CommonMind Consortium (CMC)(32), and Depression Genes and Networks (DGN)(33)) in  
106 order to derive gene expression predictor models. These models may be applied to predict genetically

107 regulated gene expression (GReX) in large genotyped cohorts, without the need to collect tissue samples  
108 or RNA(19,34–36). Here, we explored the association between AN-associated GReX and the clinical  
109 phenome. We performed transcriptomic imputation using S-PrediXcan on the most recent PGC-ED AN  
110 GWAS to first find GReX associated with AN, and then tested for clinical associations of these genes with  
111 structured EHR-encoded phenotypes using PheWAS. Given the relationship of AN with BMI, we further  
112 investigated the effect of BMI on the GReX-phenotype associations by stratifying biobank individuals  
113 based on sex- and ancestry-adjusted categories of low, mid, and high BMI. Understanding the clinical  
114 consequences of aberrant gene expression across the phenome may clarify the biological mechanisms of  
115 clinically relevant AN GWAS risk variants.

116

117

118

119

120

121

122

123

124

125 **METHODS**

126

127 **AN GENELIST and TRANSCRIPTOMIC IMPUTATION**

128 We performed transcriptomic imputation using S-PrediXcan(35) on the largest available summary  
129 statistics of Anorexia Nervosa (AN) (16992 AN cases and 55525 controls) downloaded from the  
130 Psychiatric Genomics Consortium (PGC) website(9). We tested for association of genetically-regulated  
131 gene expression (GRex) using available GTEX, CMC and DGN predictor models(34–36) for a total of 50  
132 different tissues with AN case-control status. We established two thresholds for significance; first,  
133 correcting for all genes tested within each tissue ( $p_{\text{tissue}}$ , values shown in **Table S1**), and second,  
134 correcting for all tissues and genes tested ( $p_{\text{bonferroni}}=3.75 \times 10^{-8}$ ).

135

136 **BioMe<sup>TM</sup>**

137 The BioMe<sup>TM</sup> Biobank at the Icahn School of Medicine at Mount Sinai includes genotype and EHR data  
138 from 31704 individuals recruited from the general patient population of the Mount Sinai healthcare  
139 system. Individuals were genotyped on the Illumina Global Screening Array (GSA), an array designed for  
140 multi-ethnic populations. The BioMe<sup>TM</sup> population is very diverse, with the proportion of major ancestral  
141 population groups representative of the diversity of New York City and Mount Sinai's patient population:  
142 Hispanic (HA, 35%), European (EA, 34%), African (AA, 25%), East Asian (ESA, 3%), South Asian (SAS,  
143 2%), and Native American (NA, 0.3%) (Full BioMe<sup>TM</sup> demographics are shown in **Table 1**). Quality control  
144 and imputation of the genotyping data for BioMe<sup>TM</sup> is described in elsewhere(28).

145

146 Ancestry for BioMe<sup>TM</sup> was initially assigned from self-reported values into six ancestral groupings: African  
147 (AA), European (EA), Hispanic (HA), East Asian (ESA), Native American (NA), and O (Other). We plotted  
148 the first two principal components from the genotyping to confirm ancestral grouping. For individuals not  
149 clearly falling in an ancestral group, and for those designated "Other", we merged genotype information  
150 with phase 3 of 1000Genomes and reran the principal components analysis (PCA) using PLINK(37,38) in  
151 order to assign individuals to 1000Genomes super-populations. From this we were able to assign  
152 individuals to one of six ancestral groups: African (AA), European (EA), East Asian (ESA), Hispanic (HA),

153 Native American (NA), and South Asian (SAS). A total of 119 samples that could not be grouped were  
154 removed, leaving 31585 individuals for analysis.

155

## 156 **PheWAS**

157 We calculated GReX for all  $p_{tissue}$  significant genes across GTEx(31), DGN(33), and CMC(32,39) tissues  
158 (N=50) in BioMe™ individuals using PrediXcan-2(34,35) software, and performed a phenome-wide  
159 association (PheWAS) analysis. Logistic regression between the calculated AN-GReX and phecodes was  
160 performed using the PheWAS R package(40), adjusting for age, biological sex, and the first five  
161 genotype-derived principal components. PheWAS were run per BioMe™ ancestry cohort of ( $N_{Ancestry}=6$ )  
162 and results meta-analyzed using an inverse-variance based approach in METAL(41). Due to the nature of  
163 EHR data, sample sizes can vary greatly depending on the phenotype, which can, in turn, affect our  
164 ability to detect associations. We believe our study is sufficiently powered given results from a previous  
165 study(42) that used simulations of MAF, penetrance and sample number parameters to estimate power  
166 for pheWAS studies.

167

### 168 *Encounter Diagnosis*

169 Current diagnoses for each individual were recorded at each visit using the International Classification of  
170 Disease (ICD) coding system and recorded as “Encounter Diagnosis”. Phecodes were assigned from  
171 Encounter Diagnosis information by grouping ICD-9 and ICD-10 diagnostic codes(43). Individuals with at  
172 least two counts of a particular diagnostic code were considered “cases”, those with zero counts were  
173 marked as “controls”, and those with only one count were set to missing. After QC, there were a total of  
174 2178 unique Encounter Diagnosis codes and 1093 unique phecodes in the dataset.

175

176 For both encounter diagnoses and phecodes, we required at least 10 occurrences of each phenotype  
177 within each population group for inclusion in the analysis, and an overall effective sample size  $N_{eff} >$   
178 100 (**Equation 1**).

$$N_{eff} = \frac{4}{(1/N_{Cases}) + (1/N_{Controls})}$$

179 **Equation 1: Effective sample size ( $N_{eff}$ )**

180

181 Due to the high correlation between Encounter Diagnosis and phecode files, we combined all results from  
182 both PheWAS and performed an FDR correction in  $R$  to determine overall significant associations with  
183 diagnostic phenotypes. Significant associations were set at an FDR-corrected  $p$  threshold of 0.05.

184

### 185 **BioMe-R Structured Phenotype Database (BRSPD) PheWAS**

186 In addition to diagnosis code data, we also had access to other types of EHR-derived data that included  
187 information on allergy status, vital signs (weight, height, blood pressure, pulse, pulse oximetry,  
188 respirations, and temperature), family history, personal history, medication use, obstetrics and  
189 gynecology outcomes (OB/GYN), and other social and behavioral traits. This information derived from the  
190 EHR was collected into files under the heading BRSPD or BioMe-R Structured Phenotype Database.  
191 Summaries and descriptive statistics of all of the BRSPD phenotypes can be found in **Supplementary**  
192 **Methods**. For all BRSPD traits, we required at least ten cases in each ancestry group, and a total  
193  $N_{\text{eff}} > 100$  for inclusion in our analysis.

194

### 195 *Phenotype QC and Analysis for Continuous Traits*

196 We plotted the distribution of continuous traits across the cohort, and removed implausible or impossible  
197 values (likely due to errors in entry; for example, height listed as 1745 inches). Continuous traits were  
198 then analyzed as follows:

- 199 1. Height (cm): mean measure per individual.
- 200 2. Weight (kg): (i) highest recorded weight, (ii) lowest recorded weight, (iii) weight change over time.  
201 Weight change over time was calculated by taking the difference between the highest and lowest  
202 recorded weights, divided by the number of years between those two measures.
- 203 3. Pain scores were self-reported on a 1 to 10 scale and assessed for 25 different body locations.  
204 First, we tested pain score regardless of location, using three measures: (i) highest ever pain  
205 score, (ii) mean pain score, (iii) total pain score (sum), to encompass both acute and chronic pain  
206 conditions. We then repeated this analysis to each reported pain location, with the same three  
207 measures.

- 208 4. Measures of alcohol, tobacco, and illicit drug use were categorized as follows: alcohol- ounces  
209 (oz) per day; tobacco- (i) packs per day and (ii) years used; illicit drug use- frequency of use.  
210 5. Vital signs (blood pressure, pulse, pulse oximetry, respiration, temperature (Celsius)) were  
211 analyzed using (i) highest ever value, (ii) lowest ever value, (iii) mean, (iv) variance per individual.  
212

### 213 *Categorical Traits*

214 Categorical traits were converted to binary TRUE/FALSE values (phenotypes included are shown in  
215 **Supplemental Information**). For allergy and medication phenotypes, we collapsed and combined data  
216 on equivalent root components and/or active ingredients (e.g., allergies for the antibiotic Vancomycin  
217 were coded as “VANCOMYCIN”, “VANCOMYCIN HCL”, “VANCOMYCIN IN D5W”, “VANCOMYCIN IN  
218 0.9% SODIUM CL”, “VANCOMYCIN IN DEXTROSE ISO-OSM”, and “VANCOMYCIN (BULK)”), removed  
219 extraneous symbols and corrected typing errors. A full account of QC steps is documented in the  
220 **Supplementary Methods**.

221

### 222 **STRATIFICATION BY BMI**

223 Given the nature of AN symptomology and association with weight phenotypes, we tested whether AN-  
224 GReX associations varied based on BMI. We stratified BioMe™ individuals into three main BMI  
225 categories, defined based on ancestry- and sex-specific BMI distributions within our data (**Figures S1  
226 and S2; Tables S2 and S3**). BMI values used for these analyses were measured at intake into BioMe™.  
227 We categorized individuals with BMI within the 1<sup>st</sup> quartile as “Low” BMI; those above the 3<sup>rd</sup> quartile as  
228 “High” BMI, and those falling between the 1<sup>st</sup> and 3<sup>rd</sup> quartiles as “Mid” BMI (**Figure S1**). We then  
229 repeated our analysis in these three groups for all significant associations from our full analysis.

230 **RESULTS**

231

232 ***S-PrediXcan identifies gene-tissue associations with AN***

233 We applied S-PrediXcan to PGC-AN GWAS summary statistics, and identified 47 genes across 12 loci,  
234 two loci on chromosomes 2 and 3 that reached experiment-wide significance ( $p < 3.75 \times 10^{-8}$ ; **Figure 1**,  
235 **Table 2, Table S4**), and another 10 loci that reached tissue-specific significance (Each tissue type has its  
236 own p value threshold of significance; **Figure 1, Table 2, Table S4**). These 12 loci encompass 218 gene-  
237 tissue associations (which may include the same gene associated across multiple tissues). The most  
238 significant peak on chromosome 3 includes 29 genes with significant associations with AN; this region of  
239 the genome overlaps with a GWAS peak for inflammatory bowel disease (IBD)(44), Crohn's disease, and  
240 ulcerative colitis. Many of the AN-associated genes within this peak have been found to be associated  
241 with IBD. Genes included in our AN GReX gene set have been previously associated not only with IBD,  
242 but with other psychiatric traits such as schizophrenia (*ARNTL*, *CLIC1*), depression and/or depressive  
243 symptoms (*C3orf62*, *CCDC36*, *CLIC1*, *DAG1*, *GPX1*, *MST1R*), autism (*CLIC1*), and substance use  
244 (*RBM5*, *RBM6*), as well as anthropometric measures such as BMI, waist-hip ratio, and lean body mass  
245 (**Table S5**). Of these genes, *CLIC1* lies within the major histocompatibility complex (MHC), a large region  
246 on chromosome 6 that includes genes that code for components of the adaptive immune system (Tibial  
247 nerve-*CLIC1*,  $p = 1.47 \times 10^{-6}$ ). To our knowledge, this is the first genetic association for AN within the MHC  
248 region.

249

250 ***AN-GREX is associated with anthropometric measures***

251 We tested whether GReX of AN-associated genes was associated with anthropometric phenotypes in our  
252 BioMe™ biobank (**Table 3**). In the overall cohort, upregulation of MHC-gene *CLIC1* was associated with  
253 lowest recorded weight (Spleen-*CLIC1*,  $p = 6.63 \times 10^{-4}$ ; Breast Mammary tissue-*CLIC1*,  $p = 4.18 \times 10^{-4}$ ) and  
254 mean height (Subcutaneous Adipose-*CLIC1*,  $p = 1.39 \times 10^{-5}$ ; **Table S6**). Downregulation of multiple genes,  
255 *ARIH2*, *NCKIPSD*, *NDUFAF3*, *NICN1*, *P4HTM*, *RNF123*, *SLC26A10*, was associated with decreased  
256 measures of weight change over time in the overall cohort ( $p < 9.86 \times 10^{-4}$ )(**Table 3, Figure S3**).  
257 Upregulation of *WDR6* was associated with increased measures of weight change over time ( $p < 9.81 \times 10^{-4}$ )

258 <sup>4</sup>)(**Table 3, Figure S4**). *CTNNB1*, *GPR75* and *LINC00324* GReX was associated with mean height in  
259 BioMe™ ( $p < 6.51 \times 10^{-4}$ )(**Table S6**).

260

### 261 ***Patient BMI mediates AN-GREX-PheWAS associations***

262 Given the relationship of AN with BMI, we performed a secondary analysis to assess the effect of BMI on  
263 the association between AN-GReX and anthropometric phenotypes. Specifically we wanted to look at  
264 whether our AN-GReX-phenotype associations in individuals of High and Low BMI were still present  
265 when BMI was within a normal range. We split the BioMe™ cohort into three groups: High, Mid, and Low  
266 BMI (adjusted for ancestry and sex; see **Methods** and **Supplementary Information** for more details),  
267 and re-ran our PheWAS within these groups. Indeed, our AN-GReX anthropometry associations (with  
268 highest recorded weight, lowest recorded weight, and weight change over time) varied according to the  
269 BMI group tested (**Figure 2, Table 3**). The association between *CLIC1*-GReX and highest and lowest  
270 recorded weight measures was strongest in individuals of High BMI compared with the other BMI groups  
271 (**Figure 2, Table 3**). Likewise, for multiple genes, associations between AN-GReX and weight change  
272 over time differed markedly in the High BMI group (**Table 3, Figure S4**).

273

### 274 ***AN-GREX is associated with autoimmune and autoinflammatory diseases***

275 At a false discovery rate (FDR)  $p < 0.05$ , we found 17 FDR-significant gene-tissue associations with four  
276 Phecode- and Encounter Diagnosis phenotypes: type I diabetes, celiac disease, peptic ulcer, and  
277 unspecified immunodeficiency (**Figure 3**). Upregulation of *CLIC1* in spleen was associated with two  
278 autoimmune/autoinflammatory phenotypes in the overall BioMe™ cohort: celiac disease and type I  
279 diabetes ( $p = 1.30 \times 10^{-11}$ ,  $p = 5.08 \times 10^{-20}$ , respectively) (**Tables S7 and S8**). Downregulation of *SLC26A6*  
280 was associated with the phenotype “peptic ulcer, site unspecified” in the overall cohort (Liver-*SLC26A6*,  
281  $p = 8.38 \times 10^{-8}$ ; Skeletal muscle-*SLC26A6*,  $p = 1.71 \times 10^{-7}$ ; Thyroid-*SLC26A6*,  $p = 7.16 \times 10^{-8}$ ).  
282 Downregulation of *PFKFB4* was associated with ICD10 code D84.9, “immunodeficiency, unspecified”  
283 (Skin not sun exposed-*PFKFB4*,  $p = 1.34 \times 10^{-7}$ ). We stratified our cohort by BMI (see **Methods**) to assess  
284 whether BMI was modifying the association of AN-GReX with Encounter Diagnosis and phecodes. We  
285 saw significant associations of AN-GReX phecodes within our Mid BMI group; however, when plotted

286 across the BMI groups, there was no notable difference in the magnitude and direction of effect of AN-  
287 GReX on either Encounter Diagnosis or phecodes across the BMI groups (**Figures S5 and S6**). Within  
288 individuals with Mid BMI, upregulation of *CLIC1* was similarly associated with type 1 diabetes (Mid-  
289 Subcutaneous adipose, Mid-Breast Mammary tissue, Mid-Spleen-*CLIC1*,  $p < 1.96 \times 10^{-8}$ ).

290

### 291 ***AN-GREX is associated with tobacco, alcohol, and illicit drug use***

292 Multiple AN genes were associated with both categorical and continuous measures of tobacco use (**Table**  
293 **S9, Figure S7**). *APEH*, *CCDC36*, *DALRD3*, *GPX1*, *LINC00324*, *NICN1*, and *P4HTM* were associated  
294 with years used/smoked of tobacco in the overall BioMe™ cohort ( $p < 4.97 \times 10^{-4}$ , **Table S9**) and  
295 upregulation of *MGMT* was associated with being a smoker (Heart Atrial Appendage-*MGMT*,  $p = 8.87 \times 10^{-5}$ ;  
296 Putamen basal ganglia-*MGMT*,  $p = 1.09 \times 10^{-4}$ ). Again, we note mediating effects of patient BMI category  
297 on the association between AN-GReX and smoking phenotypes. Among the Mid and High BMI groups,  
298 different genes were associated with both continuous and categorical measures of tobacco use (**Table**  
299 **S9, Figure S7**), including years of tobacco use, tobacco cigarette packs per day, and smoking every day.  
300 In addition to tobacco use measures, we see an association between AN genes and tobacco use related  
301 ICD codes. Downregulation of *LINC00324* in Low BMI individuals was associated with ICD10 diagnosis  
302 code F17.210 “Nicotine dependence, cigarettes, uncomplicated” (Low-Cells EBV-transformed-*LINC00324*,  
303  $p = 3.03 \times 10^{-5}$ ).

304

305 Downregulation of *MGMT* was associated with decreased alcohol use (ounces per week) in individuals in  
306 the Mid-BMI group ( $p < 9.75 \times 10^{-4}$ ; **Figure S7**). In individuals of Low BMI, downregulation of *GPR75* and  
307 *GPX1*, and upregulation of *NICN1* were associated with decreased and increased continuous measures  
308 of illicit drug use frequency respectively (Liver-*GPR75*,  $p = 3.34 \times 10^{-4}$ ; Nucleus accumbens-*GPX1*,  $p = 2.04$   
309  $\times 10^{-4}$ ; Testis-*NICN1*,  $p = 3.34 \times 10^{-4}$ ; **Figure S7**).

310

### 311 ***AN-GREX is associated with measures of pain score and pain location***

312 We tested for association between AN-GReX and three sets of pain score measures: (1) pain location, (2)  
313 pain score overall, (3) pain score by body location. AN-GReX was associated with multiple measures of

314 pain score measurement and location across all BMI groups (**Table S10**). Most significantly, upregulation  
315 of *TNFSF12* was associated with total knee pain score (sum) in individuals with High BMI (High-Brain  
316 Substantia nigra-*TNFSF12*,  $p=1.50 \times 10^{-3211}$ ). Downregulation of *TNFSF12* was associated with total  
317 abdomen pain score (sum) in individuals with Low BMI (Low-Brain Substantia nigra-*TNFSF12*,  $p=3.93 \times$   
318  $10^{-58}$ ) (**Table S10**).

319  
320 When looking at pain location irrespective of pain score, upregulation of *TUSC2* in the overall cohort is  
321 associated with presence of knee pain (Pancreas-*TUSC2*,  $p=7.62 \times 10^{-6}$ ) and downregulation of *CTNNB1*  
322 in Mid BMI individuals is associated with the presence of left leg pain (Mid-Stomach-*CTNNB1*,  $p=1.93 \times$   
323  $10^{-5}$ ) (**Figure 4, Table S10**). Repeating this analysis according to BMI category reveals a very large  
324 number of associations specific to the Low BMI category (**Figure 4, Table S10**); in particular, we note 129  
325 gene-tissue associations with presence of foot pain, nine of which pass phenotype-wide significance,  
326 which may indicate a propensity to excess exercise, or exercise-related injuries among these individuals.

327

### 328 ***Upregulation of CLIC1 is associated with glucagon medication***

329 AN-GReX PheWAS with medication phenotypes was associated with glucagon, a hormone used to treat  
330 severe hypoglycemia, in the overall BioMe™ cohort. Upregulation of *CLIC1* was associated with glucagon  
331 in the overall cohort (Spleen-*CLIC1*,  $p=4.20 \times 10^{-9}$ ) and also in individuals with Mid BMI (Mid-Spleen-  
332 *CLIC1*,  $p=2.09 \times 10^{-6}$ ) (**Figure S8, Table S11**).

333

334

335

336 **DISCUSSION**

337

338 ***AN-associated genes are enriched for associations with psychiatric, autoimmune, and***  
339 ***anthropometric phenotypes.***

340 Our S-PrediXcan gene results illustrate the contribution of genes associated with metabolic,  
341 anthropometric, autoimmune, and psychiatric phenotypes to AN, and 53% of our genes (25/47) overlap  
342 with those from a previous prediXcan analysis of AN cohorts(9). 24/47 AN genes (51%) had variants that  
343 were previously associated with some autoimmune disorder in GWAS studies (**Table S5**). The  
344 experiment-wide significant loci on chromosome 3 overlaps with a known GWAS peak for inflammatory  
345 bowel disease(44) (IBD), and many of the genes within that peak that we found associated with AN were  
346 also associated with disorders under the IBD umbrella such as Crohn's and ulcerative colitis (**Table S5**),  
347 demonstrating further shared genetic architecture of AN with autoimmune and autoinflammatory  
348 disease(7,8,45,46). Earlier AN GWAS studies have found AN risk loci that overlap with loci implicated in  
349 autoimmune disease(47), but the association of *CLIC1* gene expression in tibial nerve that we have found  
350 is the first known association of AN with the MHC locus, a region that has been associated with many  
351 other psychiatric disorders(48,49) as well as autoimmune disorders. In particular, *CLIC1* has previously  
352 been identified as associated with schizophrenia(50), autism(50), major depressive disorder(51), post-  
353 traumatic stress disorder(52), neuroticism(53) and depressive phenotypes(53). Importantly, *CLIC1*  
354 variants have also been associated with complement component C4 and C3 protein levels in the  
355 blood(54). *C4* is a gene with two isotypes, *C4A* and *C4B*, that, through various structural allelic  
356 combinations, give rise to complement component 4, which, along with complement component 3 (C3), is  
357 part of the complement cascade of proteins(55). Previous research has shown that the complement  
358 cascade is involved in not only immunological functions of pathogen clearance (e.g., tagging pathogens  
359 for destruction), but also in synaptic pruning and neuronal connectivity(56). The *C4A* isotype has been  
360 associated with schizophrenia risk(55), and both *C4A* and *C4B* genes had subthreshold associations with  
361 AN in our dataset.

362

363 *CLIC1* (Chloride Intracellular 1) encodes a chloride ion channel protein involved in many necessary  
364 cellular functions including the regulation of cell membrane potential, and the proliferation and  
365 differentiation of cells(57), including a role in axonal outgrowth of neurons<sup>51</sup>. In addition to neurons,  
366 *CLIC1* has been found to be expressed in many different tissue types(59), and in particular, *CLIC1*  
367 protein has been found to accumulate in peripheral blood mononuclear cells (PBMCs) when the central  
368 nervous system is in a state of chronic inflammation(58), and increased *CLIC1* expression has been  
369 shown to occur in microglia stimulated by the presence of  $\beta$ -amyloid *in vitro*(60,61). Therefore, *CLIC1*  
370 may potentially act as a biomarker for neurodegenerative disorders such as Alzheimer's disease, and  
371 downregulation of *CLIC1* has been shown to have a neuroprotective effect in the presence of  $\beta$ -amyloid  
372 proteins(58,60–62). Variants in *CLIC1* may contribute to risk of multiple psychiatric disorders and  
373 behaviors, such as schizophrenia, MDD and PTSD, as well as many autoimmune and immune-related  
374 phenotypes such as asthma and systemic lupus erythematosus(50–53)(63) (**Table S5**). Many studies  
375 have posed the role of inflammation in the development of psychiatric disorders, and in the etiology of  
376 AN(7,8,45,46,64).

377  
378 Multiple genes (8/47 or 17%) in our results were associated with regular gym attendance (**Table S5**), or  
379 measures of physical activity. Excessive and compulsive exercise is a behavior often seen in individuals  
380 with AN(65,66), and evidence is suggestive that hyperactivity increases risk of chronic AN(67,68).  
381 Likewise, genetic studies have shown a strong genetic correlation between AN and physical activity  
382 ( $r_g=0.17$ )(9,23,69). These genes may reflect a genetic liability to compulsive behaviors of physical activity.  
383 Our pain location results point to possible symptoms of excessive activity, with the strong association of  
384 AN-GReX with foot pain in Low BMI individuals. Given the known association of AN disease with low  
385 bone mineral density (BMD) and propensity for bone fracture(70), our results may reflect the result of  
386 genes associated with compulsive exercise and BMD contributing to increased osteoarticular pain. We  
387 also see multiple genes that have been previously associated with decreased sleep and insomnia  
388 phenotypes (including a core circadian clock gene *ARNTL*), which have previously been implicated in a  
389 range of psychiatric disorders(71–82), including eating disorders(83), as well as in satiety and  
390 hunger(84,85).

391

392 ***In a patient population, expression of AN genes is associated with psychiatric, metabolic, and***  
393 ***autoimmune phenotypes***

394 Unlike GWAS, which include carefully constructed case-control cohorts, PheWAS encompass all  
395 individuals within a given healthcare system, including patients with subthreshold or partial presentations  
396 of a disorder, or individuals with commonly comorbid or co-diagnosed conditions. Importantly, because  
397 individuals enrolled in the biobank are not ascertained for any specific disorder, they likely represent a  
398 comprehensive clinical picture of the comorbidities and symptomatology associated with AN gene  
399 expression. Examining the consequences of aberrant predicted gene expression among these patients  
400 (i.e., testing for GReX-associations) may reveal clinical and biological consequences of these genes; for  
401 example, studying whether AN-associated genes have anthropometric and metabolic consequences  
402 among adults with no evidence for previous AN- or ED-diagnoses may disentangle whether certain  
403 endophenotypes present as a cause, or consequence of AN. For example, it has been postulated that  
404 behaviors of food avoidance and restriction may arise due to gastrointestinal complaints and distress that  
405 provoke these behaviors and precede development of AN(86). Likewise, autoimmune disorders of the  
406 gastrointestinal tract, such as celiac disease and Crohn's disease, show a bidirectional relationship with  
407 AN, with previous diagnosis of a GI-associated autoimmune disorders increasing the risk of AN and vice  
408 versa(8). Our pheWAS results of AN-GReX associations with gastrointestinal symptoms such as  
409 abdominal pain, ascites, and peptic ulcer, as well as GI-related autoimmune disorders like celiac disease,  
410 suggest AN-GReX may contribute directly to these diseases and phenotypes, and that food aversive  
411 behaviors and gastric distress may be genetically regulated in these individuals, rather than occurring as  
412 a consequence of AN. While these associations could be the consequence of undiagnosed AN  
413 individuals in our biobank, they more likely reflect real biological associations of expression of those  
414 particular genes with the phenotype.

415

416 Our results further confirm the contribution of metabolic factors to AN etiology as we see a very robust  
417 association of AN-GReX with type 1 diabetes, and our top findings for AN-GReX with medication  
418 prescriptions include the hyperglycemic hormone glucagon and various forms of insulin. The association

419 of AN with metabolic traits and abnormalities has been fairly well established. The Anorexia Nervosa  
420 Genetics Initiative reported significant negative genetic correlation between AN and a number of  
421 metabolic traits including insulin resistance ( $r_g=-0.29$ ), fasting insulin ( $r_g=-0.24$ ), leptin ( $r_g=-0.26$ ) and type  
422 2 diabetes ( $r_g=-0.22$ ), and a strong positive genetic correlation with HDL cholesterol ( $r_g=0.21$ )(9).  
423 Additional studies have shown associations of AN with fasting insulin(87) and insulin sensitivity(88).  
424 Notably, our results point to a similar role of aberrant glycemc regulation in the etiology of AN. Future  
425 analyses including EHR-derived lab results (LabWAS) studies may further elucidate AN genes associated  
426 with abnormal metabolic regulation and clinical features.

427

#### 428 ***BMI modification of AN-GReX on clinical outcomes***

429 We stratified our pheWAS analyses by BMI in order to observe whether the effects of predicted AN gene  
430 expression on clinical outcomes were modified by BMI status, and whether this modification had a  
431 significant impact on clinical phenotypes. Specifically, are the AN-GReX-phenotype associations that we  
432 see in individuals of High and Low BMI still present when BMI is within the normal “healthy” range?  
433 Understanding how BMI is influencing the association of AN-GReX with the clinical phenome might give  
434 us further insight into the biological pathways leading to those outcomes and the conferral of AN disease  
435 risk. Additionally, we want to understand how BMI could be interacting with AN genetic risk to influence  
436 clinical outcomes. We know from our S-PrediXcan analyses and previous GWAS that many genes found  
437 to be associated with AN have variants that have been associated with BMI and other anthropometric  
438 traits (**Table S5**). 20/47 AN-genes (42.6%) had evidence from previous GWAS of association with  
439 anthropometric measures such as BMI, body fat distribution, waist-to-hip ratio, fat-free mass, obesity, and  
440 hip circumference. BMI polygenic risk scores (PRS) have been shown to be associated with disordered  
441 eating, with high BMI PRS negatively associated with thin ideal internalization and external eating, where  
442 individuals tend to eat in response to external cues of the sight or smell of food(89,90).

443

444 We postulate that the modification of AN-GReX-weight associations by BMI could be due to actual  
445 biological differences in the regulation of AN gene expression and downstream processes in the presence  
446 of different measures of BMI. CLIC1 is involved in pathways of inflammation, and individuals with a higher

447 BMI have been shown to have increased systemic inflammation(91), especially in adipose tissues, so it is  
448 possible that the modification effect by BMI on the effect of *CLIC1*-GReX on weight measures is  
449 happening via a tissue-specific inflammatory process. If we view AN and obesity as metabolic disorders  
450 on opposite ends of the spectrum, then perhaps the interaction of BMI with risk for either disorder may  
451 lead to differences in metabolic regulation resulting in higher weight phenotypes for those with increased  
452 obesity risk, and lower weight phenotypes in those with higher AN risk.

453  
454 Although the hypothesis-free, phenome-wide design of our study allows for powerful detection of clinical  
455 and biological consequences of AN risk genes, the same design also bears some notable limitations and  
456 caveats. One key caveat to our results is the lack of diagnostic detail and insights available for the  
457 patients in our study. EHR analyses leverage large, existing data sets to rapidly amass cohorts for  
458 analysis, and to yield insights into whole phenome consequences of genotype and GReX associations;  
459 however, the scale and scope of these studies precludes deep phenotyping or performing clinical  
460 interviews. This lack of diagnostic precision may arise from a number of factors.

461  
462 First, we make use of ICD codes within the medical record in order to infer diagnoses; since these are  
463 primarily used by clinicians for billing purposes, they likely provide an imperfect proxy for true disease  
464 state. In order to mitigate spurious results stemming from imperfectly assigned codes, we established  
465 criteria for inclusion of diagnoses in our analyses: First, diagnosis codes with fewer than 10 counts in the  
466 biobank were removed. Second, we only included diagnoses in our pheWAS with at least 10 counts of a  
467 code in the ancestral population (i.e., ancestral groups with fewer than 10 case counts were not included  
468 in the analysis for that diagnosis), and with a total effective sample size after meta-analysis that was  
469 greater than 100. Lastly, within each pheWAS, individuals had to have two or more counts of a diagnosis  
470 code to count as a “case”, whereas individuals with only one count of a diagnosis code were set to  
471 “missing” and not included in the analysis.

472  
473 Second, it is possible that our patients regularly receive treatment at multiple different hospital systems;  
474 as such, we may be capturing only partial data for each of our patients. In order to mitigate this, we

475 restricted our analysis to patients with multiple data points within our EHR. Future analyses that seek to  
476 harmonize or meta-analyze patient data across EHR (e.g., NYC-CDRN, PsycheMERGE) are ongoing,  
477 and will be vital to disentangling this effect further.

478  
479 Our results demonstrate that there are real clinical consequences to differences in AN gene expression.  
480 Characterization of the phenotypic consequences of AN gene expression in a clinical setting can give us  
481 more insight into the biological mechanisms underlying AN and, consequently, how to diagnose and treat  
482 the disorder. By understanding the associations of AN gene expression with symptomatology, prodromal  
483 or subthreshold disease states, we may gain insights into the biology of the disease, and perhaps identify  
484 therapeutic targets and opportunities for clinical interventions. For example, if gastrointestinal complaints  
485 are truly the consequence of aberrant AN gene expression, and contribute to disordered eating due to  
486 gastrointestinal distress, treatment of those symptoms may help alleviate other AN symptoms or prevent  
487 development of later AN(92,93). An understanding of the clinical consequences of AN gene expression  
488 can further augment the definition of AN, and could allow clinicians to more broadly identify individuals at  
489 greater risk of AN, or those who present with symptom constellations that do not yet meet the established  
490 diagnostic threshold.

491 **Acknowledgements**

492 We are deeply grateful for the mentorship of Pamela Sklar, whose guidance and wisdom we miss daily.

493 We strive to continue her legacy of thoughtful, innovative, and collaborative science.

494

495 JJ and LMH were supported by funding from the Klarman Family Foundation (2019 Eating Disorders

496 Research Grants Program) and the NIMH (R01MH118278). CMB is supported by NIMH (R01MH120170;

497 R01MH119084; R01MH118278; U01 MH109528); Brain and Behavior Research Foundation

498 Distinguished Investigator Grant; Swedish Research Council (Vetenskapsrådet, award: 538-2013-8864);

499 Lundbeck Foundation (Grant no. R276-2018-4581).

500

501 This work was supported in part through the resources and staff expertise provided by the Charles

502 Bronfman Institute for Personalized Medicine and The BioMe™ Biobank Program at the Icahn School of

503 Medicine at Mount Sinai.

504

505 Research reported in this paper was supported by the Office of Research Infrastructure of the National

506 Institutes of Health under award numbers S10OD018522 and S10OD026880. The content is solely the

507 responsibility of the authors and does not necessarily represent the official views of the National Institutes

508 of Health.

509

510 **Competing Interests**

511 CM Bulik reports: Shire (grant recipient, Scientific Advisory Board member); Idorsia (consultant);

512 Pearson (author, royalty recipient).

513

514

515 **REFERENCES:**

516

- 517 1. Zipfel S, Giel KE, Bulik CM, Hay P, Schmidt U. Anorexia nervosa: aetiology, assessment, and  
518 treatment. *Lancet Psychiatry*. 2015 Dec;2(12):1099–111.
- 519 2. Bulik CM, Flatt R, Abbaspour A, Carroll I. Reconceptualizing anorexia nervosa. *Psychiatry Clin  
520 Neurosci*. 2019;73(9):518–25.
- 521 3. Mitchell JE, Peterson CB. Anorexia Nervosa. *N Engl J Med*. 2020 02;382(14):1343–51.
- 522 4. Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a  
523 meta-review. *World Psychiatry Off J World Psychiatr Assoc WPA*. 2014 Jun;13(2):153–60.
- 524 5. Ulfvebrand S, Birgegård A, Norring C, Högdahl L, von Hausswolff-Juhlin Y. Psychiatric comorbidity in  
525 women and men with eating disorders results from a large clinical database. *Psychiatry Res*. 2015  
526 Dec 15;230(2):294–9.
- 527 6. Root TL, Pinheiro AP, Thornton L, Strober M, Fernandez-Aranda F, Brandt H, et al. Substance use  
528 disorders in women with anorexia nervosa. *Int J Eat Disord*. 2010 Jan;43(1):14–21.
- 529 7. Raevuori A, Haukka J, Vaarala O, Suvisaari JM, Gissler M, Grainger M, et al. The increased risk for  
530 autoimmune diseases in patients with eating disorders. *PLoS One*. 2014;9(8):e104845.
- 531 8. Hedman A, Breithaupt L, Hübel C, Thornton LM, Tillander A, Norring C, et al. Bidirectional relationship  
532 between eating disorders and autoimmune diseases. *J Child Psychol Psychiatry*. 2019;60(7):803–  
533 12.
- 534 9. Anorexia Nervosa Genetics Initiative, Eating Disorders Working Group of the Psychiatric Genomics  
535 Consortium, Watson HJ, Yilmaz Z, Thornton LM, Hübel C, et al. Genome-wide association study  
536 identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nat Genet*.  
537 2019 Aug;51(8):1207–14.
- 538 10. Yilmaz Z, Hardaway JA, Bulik CM. Genetics and Epigenetics of Eating Disorders. *Adv Genomics  
539 Genet*. 2015;5:131–50.
- 540 11. Levinson CA, Zerwas SC, Brosf LC, Thornton LM, Strober M, Pivarunas B, et al. Associations  
541 between dimensions of anorexia nervosa and obsessive-compulsive disorder: An examination of  
542 personality and psychological factors in patients with anorexia nervosa. *Eur Eat Disord Rev J Eat  
543 Disord Assoc*. 2019;27(2):161–72.
- 544 12. Kaye WH, Wierenga CE, Bailer UF, Simmons AN, Bischoff-Grethe A. Nothing tastes as good as  
545 skinny feels: the neurobiology of anorexia nervosa. *Trends Neurosci*. 2013 Feb 1;36(2):110–20.
- 546 13. Speakman JR, Levitsky DA, Allison DB, Bray MS, Castro JM de, Clegg DJ, et al. Set points, settling  
547 points and some alternative models: theoretical options to understand how genes and environments  
548 combine to regulate body adiposity. *Dis Model Mech*. 2011 Nov 1;4(6):733–45.
- 549 14. Müller MJ, Bösny-Westphal A, Heymsfield SB. Is there evidence for a set point that regulates human  
550 body weight? *F1000 Med Rep [Internet]*. 2010 Aug 9 [cited 2020 Oct 20];2. Available from:  
551 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2990627/>
- 552 15. Hebebrand J, Müller TD, Holtkamp K, Herpertz-Dahlmann B. The role of leptin in anorexia nervosa:  
553 clinical implications. *Mol Psychiatry*. 2007 Jan;12(1):23–35.

- 554 16. Hübel C, Yilmaz Z, Schaumberg KE, Breithaupt L, Hunjan A, Horne E, et al. Body composition in  
555 anorexia nervosa: Meta-analysis and meta-regression of cross-sectional and longitudinal studies.  
556 *Int J Eat Disord.* 2019;52(11):1205–23.
- 557 17. Igudesman D, Sweeney M, Carroll IM, Mayer-Davis EJ, Bulik CM. Gut-Brain Interactions:  
558 Implications for a Role of the Gut Microbiota in the Treatment and Prognosis of Anorexia Nervosa  
559 and Comparison to Type I Diabetes. *Gastroenterol Clin North Am.* 2019;48(3):343–56.
- 560 18. Huckins LM, Hatzikotoulas K, Southam L, Thornton LM, Steinberg J, Aguilera-Mckay F, et al.  
561 Investigation of common, low-frequency and rare genome-wide variation in anorexia nervosa. *Mol*  
562 *Psychiatry.* 2018;23(5).
- 563 19. Huckins LM, Dobbyn A, McFadden W, Ruderfer D, Wang W, Gamazon E, et al. Identifying tissues  
564 implicated in Anorexia Nervosa using Transcriptomic Imputation. *bioRxiv.* 2018;265017.
- 565 20. Hinney A, Kesselmeier M, Jall S, Volckmar A-L, Föcker M, Antel J, et al. Evidence for three genetic  
566 loci involved in both anorexia nervosa risk and variation of body mass index. *Mol Psychiatry.*  
567 2017;22(2).
- 568 21. Duncan L, Yilmaz Z, Gaspar H, Walters R, Goldstein J, Anttila V, et al. Significant Locus and  
569 Metabolic Genetic Correlations Revealed in Genome-Wide Association Study of Anorexia Nervosa.  
570 *Am J Psychiatry.* 2017 01;174(9):850–8.
- 571 22. Munn-Chernoff MA, Johnson EC, Chou Y-L, Coleman JRI, Thornton LM, Walters RK, et al. Shared  
572 genetic risk between eating disorder- and substance-use-related phenotypes: Evidence from  
573 genome-wide association studies. *Addict Biol.* 2020 Feb 16;e12880.
- 574 23. Hübel C, Gaspar HA, Coleman JRI, Finucane H, Purves KL, Hanscombe KB, et al. Genomics of  
575 body fat percentage may contribute to sex bias in anorexia nervosa. *Am J Med Genet Part B*  
576 *Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet.* 2019;180(6):428–38.
- 577 24. C H, M A, M H, Rj L, G B, Cm B, et al. Binge-eating disorder, anorexia nervosa, and constitutional  
578 thinness differ in their associations with anthropometric and psychiatric polygenic scores. 2020 Mar  
579 26 [cited 2020 Aug 20]; Available from: <https://europepmc.org/article/ppr/ppr126993>
- 580 25. Smoller JW. The use of electronic health records for psychiatric phenotyping and genomics. *Am J*  
581 *Med Genet Part B Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet.* 2018;177(7):601–12.
- 582 26. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, et al. PheWAS:  
583 demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations.  
584 *Bioinforma Oxf Engl.* 2010 May;26(9):1205–10.
- 585 27. Pendergrass SA, Brown-Gentry K, Dudek SM, Torstenson ES, Ambite JL, Avery CL, et al. The use  
586 of phenome-wide association studies (PheWAS) for exploration of novel genotype-phenotype  
587 relationships and pleiotropy discovery. *Genet Epidemiol.* 2011 Jul;35(5):410–22.
- 588 28. Zheutlin AB, Dennis J, Karlsson Linnér R, Moscati A, Restrepo N, Straub P, et al. Penetrance and  
589 Pleiotropy of Polygenic Risk Scores for Schizophrenia in 106,160 Patients Across Four Health Care  
590 Systems. *Am J Psychiatry.* 2019 Aug 16;176(10):846–55.
- 591 29. Leppert B, Millard LAC, Riglin L, Davey Smith G, Thapar A, Tilling K, et al. A cross-disorder PRS-  
592 pheWAS of 5 major psychiatric disorders in UK Biobank. *PLoS Genet.* 2020 May;16(5):e1008185.

- 593 30. Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, Mosley JD, et al. Systematic comparison  
594 of phenome-wide association study of electronic medical record data and genome-wide  
595 association study data. *Nat Biotechnol.* 2013 Dec;31(12):1102–10.
- 596 31. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue  
597 Expression (GTEx) project. *Nat Genet.* 2013 Jun;45(6):580–5.
- 598 32. Hoffman GE, Bendl J, Voloudakis G, Montgomery KS, Sloofman L, Wang Y-C, et al. CommonMind  
599 Consortium provides transcriptomic and epigenomic data for Schizophrenia and Bipolar Disorder.  
600 *Sci Data.* 2019 Sep 24;6(1):180.
- 601 33. Battle A, Mostafavi S, Zhu X, Potash JB, Weissman MM, McCormick C, et al. Characterizing the  
602 genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. *Genome Res.*  
603 2014 Jan;24(1):14–24.
- 604 34. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-  
605 based association method for mapping traits using reference transcriptome data. *Nat Genet.* 2015  
606 Sep;47(9):1091–8.
- 607 35. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, et al. Exploring the  
608 phenotypic consequences of tissue specific gene expression variation inferred from GWAS  
609 summary statistics. *Nat Commun.* 2018 May 8;9(1):1825.
- 610 36. Huckins LM, Dobbyn A, Ruderfer DM, Hoffman G, Wang W, Pardiñas AF, et al. Gene expression  
611 imputation across multiple brain regions provides insights into schizophrenia risk. *Nat Genet.*  
612 2019;51(4):659–74.
- 613 37. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for  
614 whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007  
615 Sep;81(3):559–75.
- 616 38. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising  
617 to the challenge of larger and richer datasets. *GigaScience.* 2015;4:7.
- 618 39. Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, et al. Gene  
619 expression elucidates functional impact of polygenic risk for schizophrenia. *Nat Neurosci.*  
620 2016;19(11):1442–53.
- 621 40. PheWAS/PheWAS [Internet]. PheWAS; 2019 [cited 2019 Nov 26]. Available from:  
622 <https://github.com/PheWAS/PheWAS>
- 623 41. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association  
624 scans. *Bioinformatics.* 2010 Sep 1;26(17):2190–1.
- 625 42. Verma A, Bradford Y, Dudek S, Lucas AM, Verma SS, Pendergrass SA, et al. A simulation study  
626 investigating power estimates in phenome-wide association studies. *BMC Bioinformatics.* 2018 Apr  
627 4;19(1):120.
- 628 43. Wu P, Gifford A, Meng X, Li X, Campbell H, Varley T, et al. Developing and Evaluating Mappings of  
629 ICD-10 and ICD-10-CM Codes to PheCodes. *bioRxiv.* 2019 Jul 3;462077.
- 630 44. de Lange KM, Moutsianas L, Lee JC, Lamb CA, Luo Y, Kennedy NA, et al. Genome-wide  
631 association study implicates immune activation of multiple integrin genes in inflammatory bowel  
632 disease. *Nat Genet.* 2017 Feb;49(2):256–61.

- 633 45. Mårild K, Størdal K, Bulik CM, Rewers M, Ekblom A, Liu E, et al. Celiac Disease and Anorexia  
634 Nervosa: A Nationwide Study. *Pediatrics*. 2017 May;139(5).
- 635 46. S Z, Jt L, L P, Lm T, M Q, Sv K, et al. Eating Disorders, Autoimmune, and Autoinflammatory  
636 Disease. *Pediatrics* [Internet]. 2017 Nov 9 [cited 2020 Nov 16];140(6). Available from:  
637 <https://europepmc.org/article/med/29122972>
- 638 47. Gibson D, Mehler PS. Anorexia Nervosa and the Immune System—A Narrative Review. *J Clin Med*.  
639 2019 Nov;8(11):1915.
- 640 48. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetsky V, et al. Genome-wide  
641 association study identifies 30 loci associated with bipolar disorder. *Nat Genet*. 2019 May;51(5):793.
- 642 49. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108  
643 schizophrenia-associated genetic loci. *Nature*. 2014 Jul;511(7510):421–7.
- 644 50. The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium. Meta-  
645 analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus  
646 at 10q24.32 and a significant overlap with schizophrenia. *Mol Autism*. 2017 Dec;8(1):21.
- 647 51. Zhu Z, Zhu X, Liu C-L, Shi H, Shen S, Yang Y, et al. Shared genetics of asthma and mental health  
648 disorders: a large-scale genome-wide cross-trait analysis. *Eur Respir J* [Internet]. 2019 Dec 1 [cited  
649 2020 May 21];54(6). Available from: <https://erj.ersjournals.com/content/54/6/1901507-0>
- 650 52. Marchese S, Hoffman G, Pietrzak RH, Lebovitch D, Daskalakis NP, Yehuda R, et al. Altered gene  
651 expression and PTSD symptom dimensions in World Trade Center responders. *In Prep*. 2020;
- 652 53. Baselmans BML, Jansen R, Ip HF, van Dongen J, Abdellaoui A, van de Weijer MP, et al.  
653 Multivariate Genome-wide and integrated transcriptome and epigenome-wide analyses of the Well-  
654 being spectrum [Internet]. *Genetics*; 2017 Mar [cited 2020 May 21]. Available from:  
655 <http://biorxiv.org/lookup/doi/10.1101/115915>
- 656 54. Yang X, Sun J, Gao Y, Tan A, Zhang H, Hu Y, et al. Genome-Wide Association Study for Serum  
657 Complement C3 and C4 Levels in Healthy Chinese Subjects. *PLOS Genet*. 2012 Sep  
658 13;8(9):e1002916.
- 659 55. Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from  
660 complex variation of complement component 4. *Nature*. 2016;530(7589):177–83.
- 661 56. Stephan AH, Barres BA, Stevens B. The complement system: an unexpected role in synaptic  
662 pruning during development and disease. *Annu Rev Neurosci*. 2012;35:369–89.
- 663 57. Li B, Mao Y, Wang Z, Chen Y, Wang Y, Zhai C, et al. CLIC1 Promotes the Progression of Gastric  
664 Cancer by Regulating the MAPK/AKT Pathways. *Cell Physiol Biochem*. 2018;46(3):907–24.
- 665 58. Carlini V, Verduci I, Cianci F, Cannavale G, Fenoglio C, Galimberti D, et al. CLIC1 Protein  
666 Accumulates in Circulating Monocyte Membrane during Neurodegeneration. *Int J Mol Sci*. 2020  
667 Jan;21(4):1484.
- 668 59. Facal F, Costas J. Evidence of association of the DISC1 interactome gene set with schizophrenia  
669 from GWAS. *Prog Neuropsychopharmacol Biol Psychiatry*. 2019 20;95:109729.
- 670 60. Milton RH, Abeti R, Averaimo S, DeBiasi S, Vitellaro L, Jiang L, et al. CLIC1 function is required for  
671 beta-amyloid-induced generation of reactive oxygen species by microglia. *J Neurosci Off J Soc*  
672 *Neurosci*. 2008 Nov 5;28(45):11488–99.

- 673 61. Novarino G, Fabrizi C, Tonini R, Denti MA, Malchiodi-Albedi F, Lauro GM, et al. Involvement of the  
674 intracellular ion channel CLIC1 in microglia-mediated beta-amyloid-induced neurotoxicity. *J*  
675 *Neurosci Off J Soc Neurosci*. 2004 Jun 9;24(23):5322–30.
- 676 62. Swarup V, Chang TS, Duong DM, Dammer EB, Dai J, Lah JJ, et al. Identification of Conserved  
677 Proteomic Networks in Neurodegenerative Dementia. *Cell Rep*. 2020 Jun 23;31(12):107807.
- 678 63. Lee YH, Bae S-C, Choi SJ, Ji JD, Song GG. Genome-wide pathway analysis of genome-wide  
679 association studies on systemic lupus erythematosus and rheumatoid arthritis. *Mol Biol Rep*. 2012  
680 Dec 1;39(12):10627–35.
- 681 64. Nilsson IAK, Millischer V, Göteson A, Hübel C, Thornton LM, Bulik CM, et al. Aberrant inflammatory  
682 profile in acute but not recovered anorexia nervosa. *Brain Behav Immun*. 2020 Aug 1;88:718–24.
- 683 65. Bulik CM. Genetic and Biological Risk Factors. In: Thompson JK, editor. *Handbook of Eating*  
684 *Disorders and Obesity* [Internet]. Hoboken, NJ: Wiley; 2004. p. 3. Available from:  
685 [https://books.google.com/books/about/Handbook%7B\\_%7Dof%7B\\_%7DEating%7B\\_%7DDisorders%7B\\_%7Dand%7B\\_%7DObesity.html?id=KZhHAAAAMAAJ%7B%7Dpgis=1](https://books.google.com/books/about/Handbook%7B_%7Dof%7B_%7DEating%7B_%7DDisorders%7B_%7Dand%7B_%7DObesity.html?id=KZhHAAAAMAAJ%7B%7Dpgis=1)  
686
- 687 66. Shroff H, Reba L, Thornton LM, Tozzi F, Klump KL, Berrettini WH, et al. Features associated with  
688 excessive exercise in women with eating disorders. *Int J Eat Disord*. 2006 Sep;39(6):454–61.
- 689 67. Achamrah N, Coëffier M, Déchelotte P. Physical activity in patients with anorexia nervosa. *Nutr Rev*.  
690 2016 May 1;74(5):301–11.
- 691 68. Strober M, Freeman R, Morrell W. The long-term course of severe anorexia nervosa in adolescents:  
692 Survival analysis of recovery, relapse, and outcome predictors over 10–15 years in a prospective  
693 study. *Int J Eat Disord*. 1997;22(4):339–60.
- 694 69. Hübel C, Gaspar HA, Coleman JRI, Hanscombe KB, Purves K, Prokopenko I, et al. Genetic  
695 correlations of psychiatric traits with body composition and glycemic traits are sex- and age-  
696 dependent. *Nat Commun*. 2019 18;10(1):5765.
- 697 70. Solmi M, Veronese N, Correll CU, Favaro A, Santonastaso P, Caregato L, et al. Bone mineral  
698 density, osteoporosis, and fractures among people with eating disorders: a systematic review and  
699 meta-analysis. *Acta Psychiatr Scand*. 2016 May;133(5):341–51.
- 700 71. Etain B, Jamain S, Milhiet V, Lajnef M, Boudebese C, Dumaine A, et al. Association between  
701 circadian genes, bipolar disorders and chronotypes. *Chronobiol Int*. 2014 Aug;31(7):807–14.
- 702 72. Melo MCA, Abreu RLC, Linhares Neto VB, de Bruin PFC, de Bruin VMS. Chronotype and circadian  
703 rhythm in bipolar disorder: A systematic review. *Sleep Med Rev*. 2017 Aug;34:46–58.
- 704 73. Huckins L, Dobbyn A, McFadden W, Wang W, Ruderfer D, Hoffman G, et al. Transcriptomic  
705 Imputation of Bipolar Disorder and Bipolar subtypes reveals 29 novel associated genes. *bioRxiv*.  
706 2017 Nov;222786.
- 707 74. Meyrer R, Demling J, Kornhuber J, Nowak M. Effects of night shifts in bipolar disorders and extreme  
708 morningness. *Bipolar Disord*. 2009 Dec;11(8):897–9.
- 709 75. Alloy LB, Ng TH, Titone MK, Boland EM. Circadian Rhythm Dysregulation in Bipolar Spectrum  
710 Disorders. *Curr Psychiatry Rep*. 2017 Apr;19(4):21.
- 711 76. Boivin DB. Influence of sleep-wake and circadian rhythm disturbances in psychiatric disorders. *J*  
712 *Psychiatry Neurosci JPN*. 2000 Nov;25(5):446–58.

- 713 77. Bunney BG, Bunney WE. Mechanisms of Rapid Antidepressant Effects of Sleep Deprivation  
714 Therapy: Clock Genes and Circadian Rhythms. *Biol Psychiatry*. 2013 Jun;73(12):1164–71.
- 715 78. Bunney BG, Li JZ, Walsh DM, Stein R, Vawter MP, Cartagena P, et al. Circadian dysregulation of  
716 clock genes: clues to rapid treatments in major depressive disorder. *Mol Psychiatry*. 2015  
717 Feb;20(1):48–55.
- 718 79. Karatsoreos IN. Links between Circadian Rhythms and Psychiatric Disease. *Front Behav Neurosci*.  
719 2014;8:162.
- 720 80. Kripke DF, Mullaney DJ, Atkinson M, Wolf S. Circadian rhythm disorders in manic-depressives. *Biol*  
721 *Psychiatry*. 1978 Jun;13(3):335–51.
- 722 81. Mansour HA, Talkowski ME, Wood J, Chowdari KV, McClain L, Prasad K, et al. Association study of  
723 21 circadian genes with bipolar I disorder, schizoaffective disorder, and schizophrenia. *Bipolar*  
724 *Disord*. 2009 Nov;11(7):701–10.
- 725 82. Murray G, Allen NB, Trinder J, Burgess H. Is weakened circadian rhythmicity a characteristic of  
726 neuroticism? *J Affect Disord*. 2002 Dec;72(3):281–9.
- 727 83. Allison KC, Spaeth A, Hopkins CM. Sleep and Eating Disorders. *Curr Psychiatry Rep*.  
728 2016;18(10):92.
- 729 84. Arble DM, Bass J, Laposky AD, Vitaterna MH, Turek FW. Circadian timing of food intake contributes  
730 to weight gain. *Obesity*. 2009;17(11):2100–2.
- 731 85. Herpertz S, Albers N, Wagner R, Pelz B, Köpp W, Mann K, et al. Longitudinal changes of circadian  
732 leptin, insulin and cortisol plasma levels and their correlation during refeeding in patients with  
733 anorexia nervosa. *Eur J Endocrinol*. 2000 Apr;142(4):373–9.
- 734 86. Zucker NL, Bulik CM. On bells, saliva, and abdominal pain or discomfort: Early aversive visceral  
735 conditioning and vulnerability for anorexia nervosa. *Int J Eat Disord*. 2020;53(4):508–12.
- 736 87. Adams DM, Reay WR, Geaghan MP, Cairns MJ. Investigation of glycaemic traits in psychiatric  
737 disorders using Mendelian randomisation revealed a causal relationship with anorexia nervosa.  
738 *Neuropsychopharmacology*. 2020 Sep 13;1–10.
- 739 88. Ilyas A, Hübel C, Stahl D, Stadler M, Ismail K, Breen G, et al. The metabolic underpinning of eating  
740 disorders: A systematic review and meta-analysis of insulin sensitivity. *Mol Cell Endocrinol*. 2019  
741 Nov 1;497:110307.
- 742 89. Abdulkadir M, Herle M, De Stavola BL, Hübel C, Santos Ferreira DL, Loos RJF, et al. Polygenic  
743 Score for Body Mass Index Is Associated with Disordered Eating in a General Population Cohort. *J*  
744 *Clin Med* [Internet]. 2020 Apr 21 [cited 2020 Nov 3];9(4). Available from:  
745 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7231239/>
- 746 90. Nagata JM, Braudt DB, Domingue BW, Bibbins-Domingo K, Garber AK, Griffiths S, et al. Genetic  
747 risk, body mass index, and weight control behaviors: Unlocking the triad. *Int J Eat Disord*.  
748 2019;52(7):825–33.
- 749 91. Kolb R, Sutterwala FS, Zhang W. Obesity and cancer: inflammation bridges the two. *Curr Opin*  
750 *Pharmacol*. 2016 Aug 1;29:77–89.

751 92. Wiklund CA, Kuja-Halkola R, Thornton LM, Hübel C, Leppä V, Bulik CM. Prolonged constipation  
752 and diarrhea in childhood and disordered eating in adolescence. *J Psychosom Res.*  
753 2019;126:109797.

754 93. Riedlinger C, Schmidt G, Weiland A, Stengel A, Giel KE, Zipfel S, et al. Which Symptoms,  
755 Complaints and Complications of the Gastrointestinal Tract Occur in Patients With Eating  
756 Disorders? A Systematic Review and Quantitative Analysis. *Front Psychiatry.* 2020;11:195.

757

758

759 **MAIN FIGURES AND TABLES**

760 **Figure 1: S-PrediXcan results for PGC-ED AN GWAS ( $N_{\text{Cases}}=16,992$ ,  $N_{\text{Controls}}=55,525$ ).** S-PrediXcan  
761 transcriptomic imputation of the PGC-ED AN GWAS summary statistics to determine genetically-  
762 regulated gene expression (GReX), with tests for association of GReX with AN disease status. Manhattan  
763 plot of S-PrediXcan gene-tissue associations with anorexia nervosa for 50 tissues. Each point represents  
764 a different gene-tissue association result; i.e., the same gene may have multiple points within a peak.

765  
766 **Figure 2: Significant AN-GReX-tissue associations with weight phenotypes.** Effect estimates and  
767 95% CI for associations of AN-GReX with (i) highest recorded weight (kg) and (ii) lowest recorded weight  
768 (kg). Associations are shown for four BMI groups: Overall, High, Mid and Low BMI. Tissue-specific  
769 significance threshold set at  $p < 0.001$ .

770  
771 **Figure 3: AN-GReX associations with full BioMe™ Encounter Diagnosis ICD codes and Phecodes.**  
772 Diagnosis codes are plotted along the X-axis and grouped by category. (A) Manhattan PheWAS plot of  
773 AN-GReX gene associations with Encounter Diagnosis ICD codes ( $N=2,178$ ) (B) Manhattan PheWAS  
774 plot of AN-GReX gene associations with Phecodes ( $N=1,093$ ). Encounter Diagnosis ICD-10 code  
775 descriptions: D84.9 – Immunodeficiency, unspecified; E10.65 – Type 1 diabetes mellitus with  
776 hyperglycemia; E10.9 – Type 1 diabetes mellitus without complications; K90.0 – Celiac disease. Phecode  
777 descriptions: 250.1 – Type 1 diabetes; 531.4 – Peptic ulcer, site unspecified; 557.1 – Celiac disease.

778  
779 **Figure 4: AN-GReX associations with pain location.** Pain locations are shown along the x-axis ( $N=25$ ).  
780 Associations are shown for four BMI groups (A) Overall cohort (B) High BMI (C) Mid BMI and (D) Low  
781 BMI. Bonferroni significance threshold is demarcated by the dotted line ( $p=4.17 \times 10^{-5}$ ).

782  
783 **Table 1: BioMe™ Demographics ( $N=31,585$ ).** Ancestry designations: AA (African), EA (European), ESA  
784 (East Asian), HA (Hispanic), SAS (South Asian), NA (Native American).

785

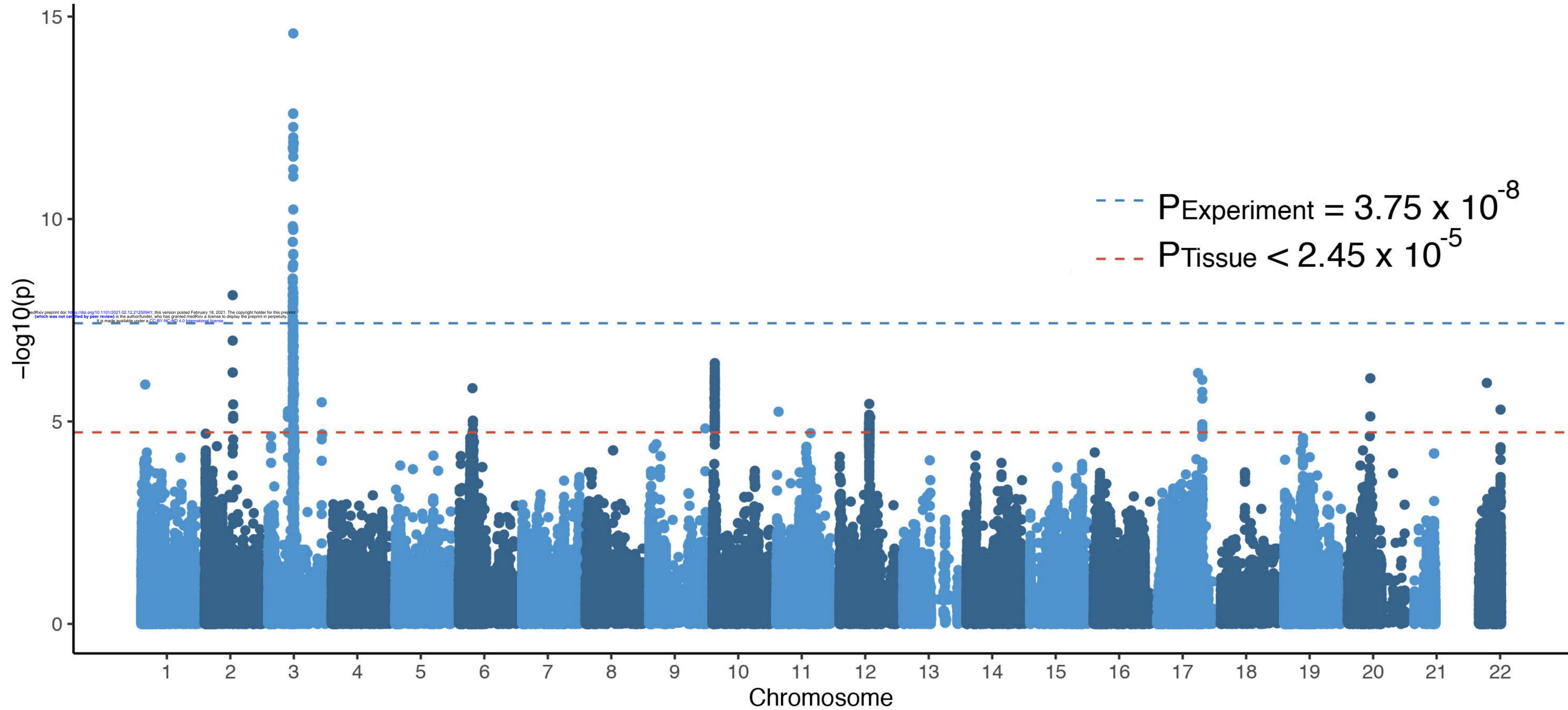
786 **Table 2: S-PrediXcan loci results.** Summary of S-PrediXcan of PGC-ED AN GWAS results. Genes  
787 associated with Anorexia nervosa from each locus are shown, as well as the top finding Zscore, p-value,  
788 gene and tissue for each locus. P-values shaded in bold indicate experiment-wide significant loci ( $p < 3.75$   
789  $\times 10^{-8}$ ).

790

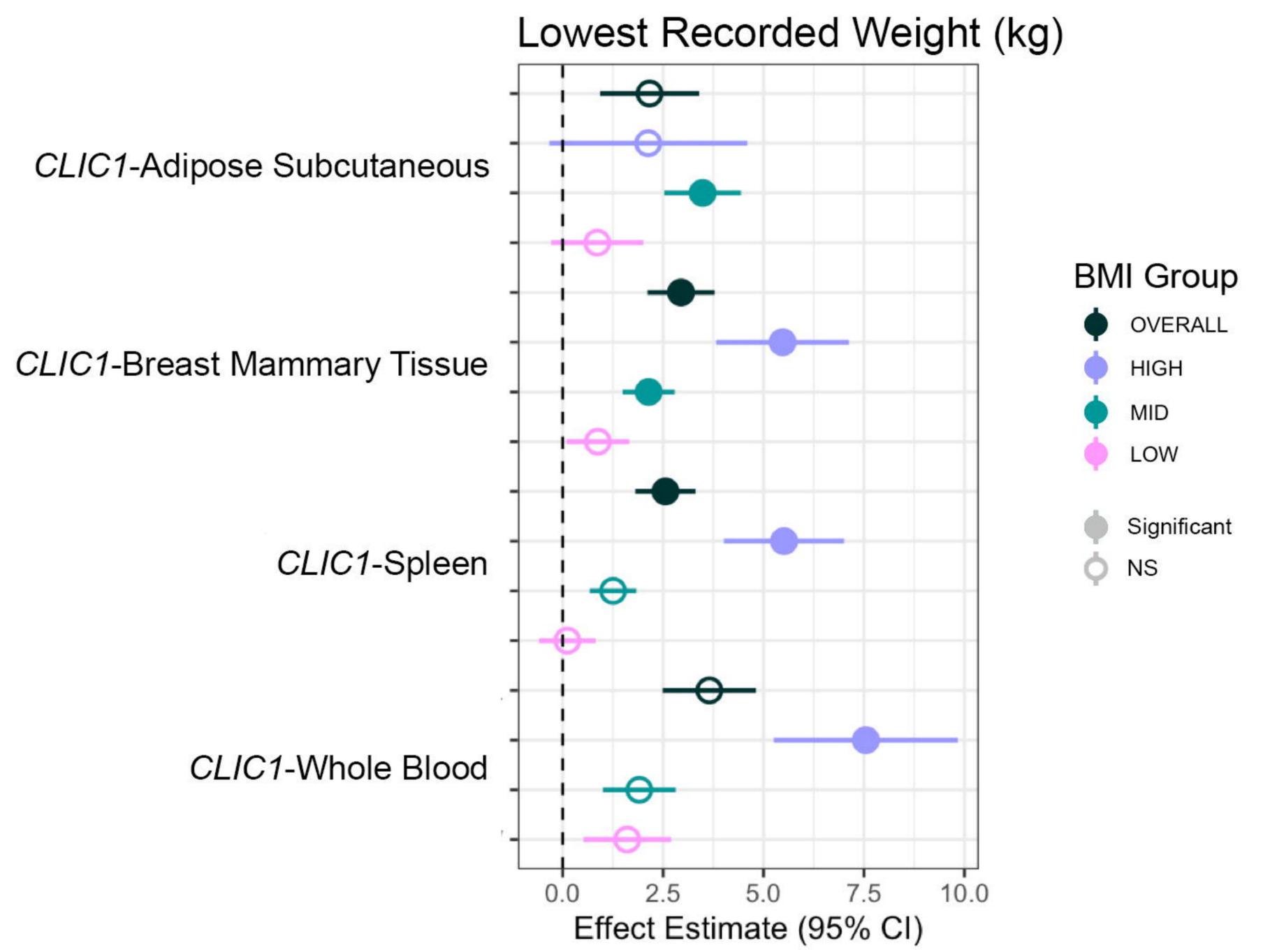
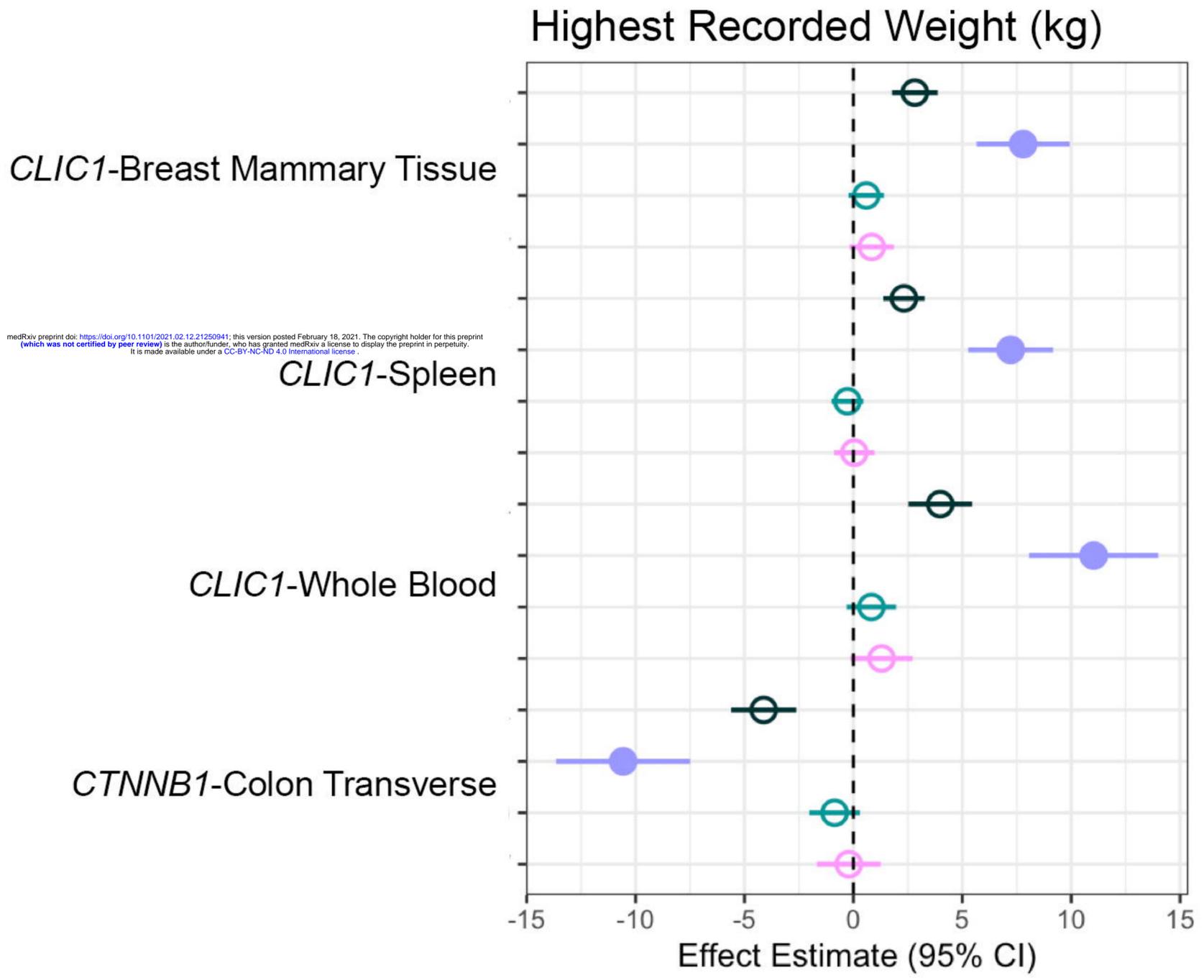
791 **Table 3: PheWAS AN-GReX tissue associations with weight phenotypes.** AN-GReX association with  
792 phenotypes of (i) Highest recorded weight (kg), (ii) Lowest recorded weight (kg), (iii) weight change over  
793 time (kg/yr) for four BMI groups: Overall, High, Mid and Low BMI. Tissue-specific significance threshold  
794 set at  $p < 0.001$ .

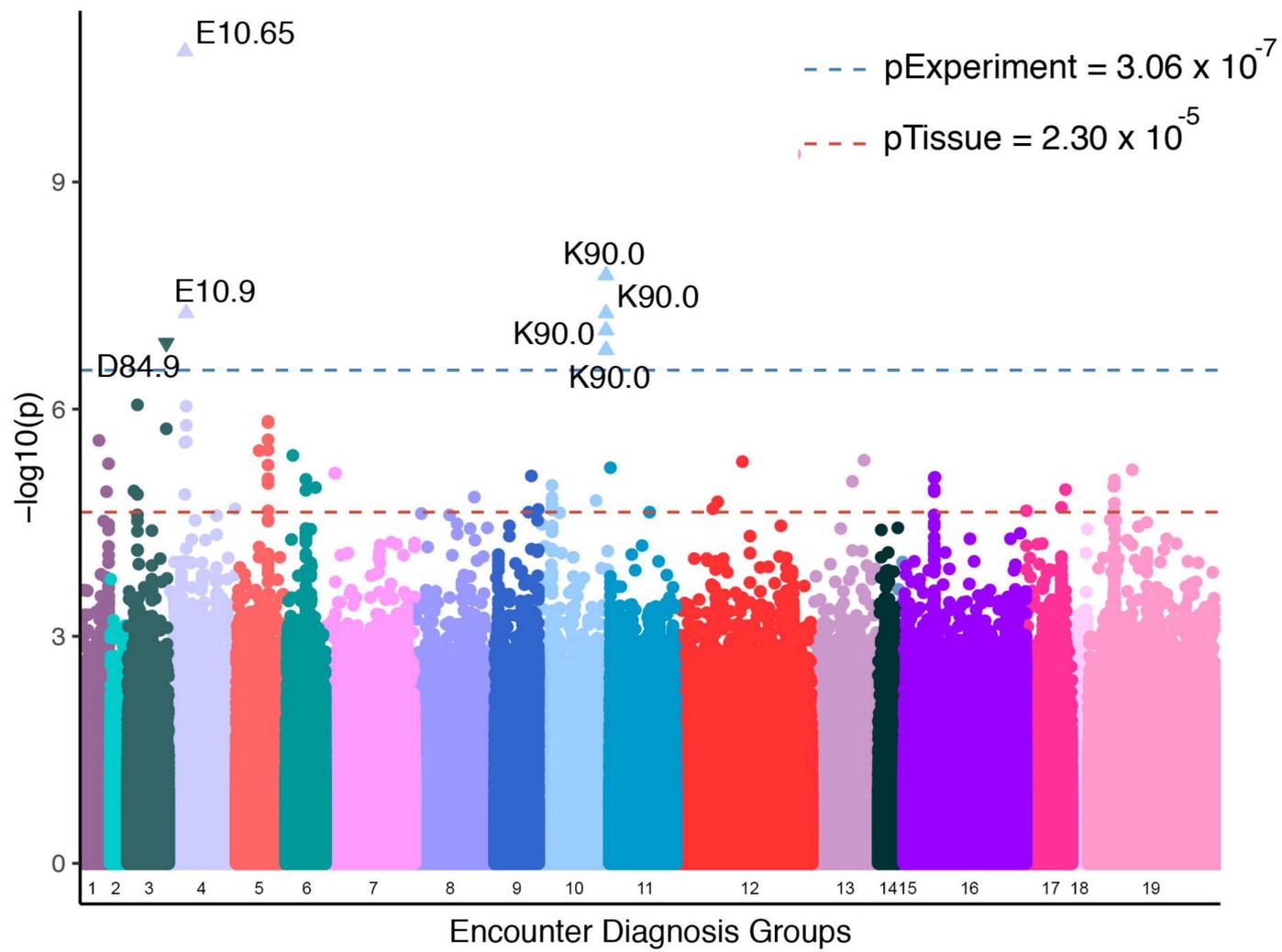
795

796



medRxiv preprint doi: <https://doi.org/10.1101/2021.02.12.21250941>; this version posted February 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



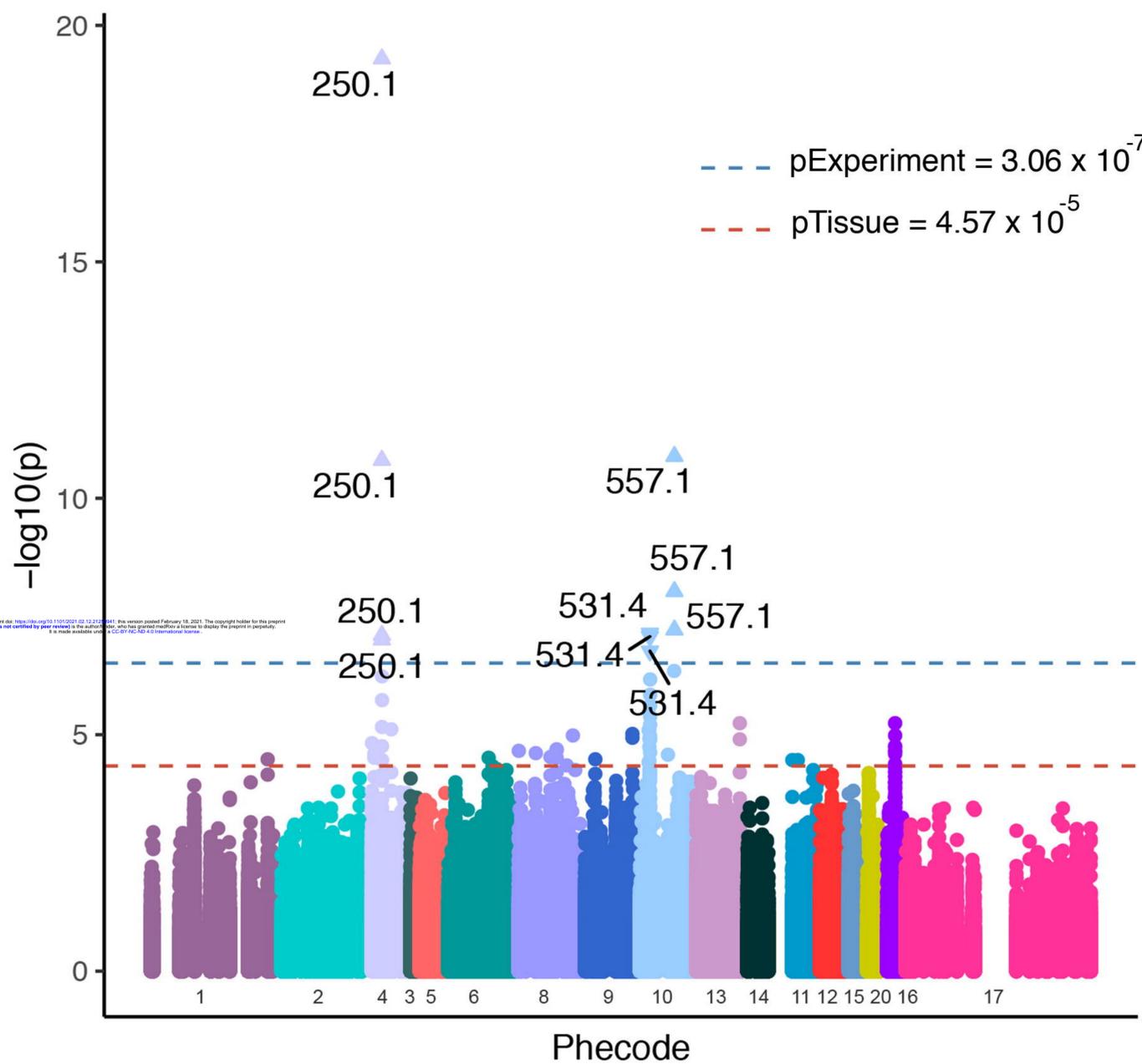


### Diagnosis Group

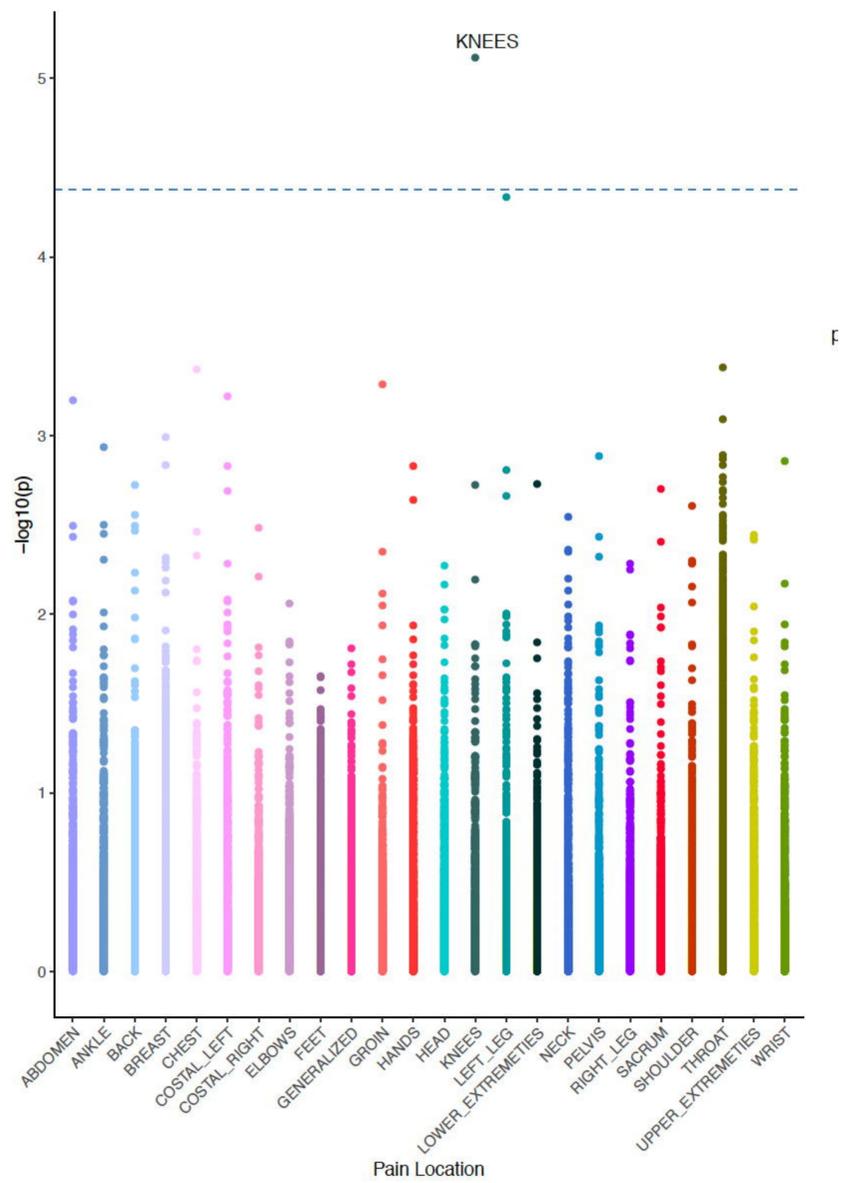
- 1 Infectious and parasitic
- 2 Neoplasms
- 3 Blood disorders
- 4 Endocrine/Nutritional/Metabolic
- 5 Mental/Behavioral/Neurodevelopmental
- 6 Nervous system
- 7 Eye/Ear
- 8 Circulatory system
- 9 Respiratory system
- 10 Digestive system
- 11 Skin/Subcutaneous tissue
- 12 Musculoskeletal system/Connective tissue
- 13 Genitourinary system
- 14 Pregnancy/Childbirth/Puerperium
- 15 Congenital abnormalities
- 16 Symptoms/signs
- 17 Injury/poisoning
- 18 External causes of morbidity
- 19 Factors influencing health status
- 20 Conditions originating in the perinatal period

### Direction of effect

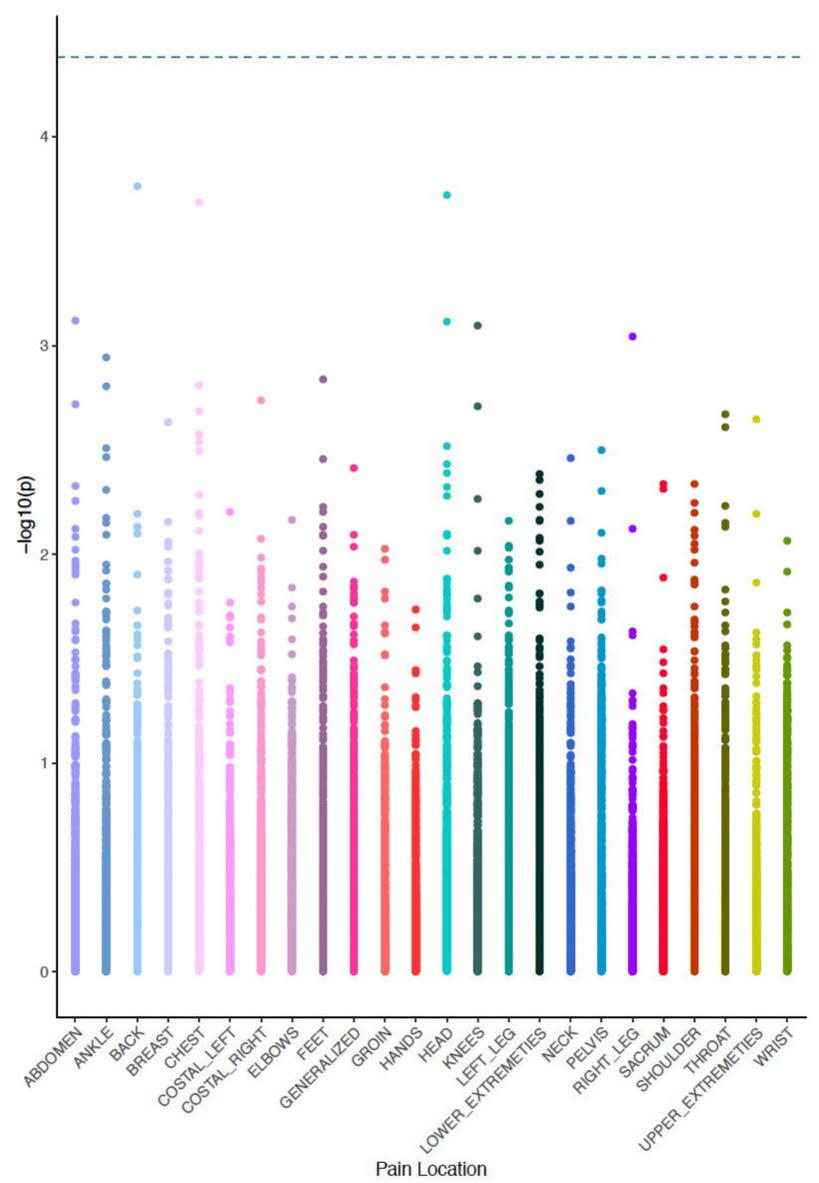
- ▼ DOWN
- ▲ UP



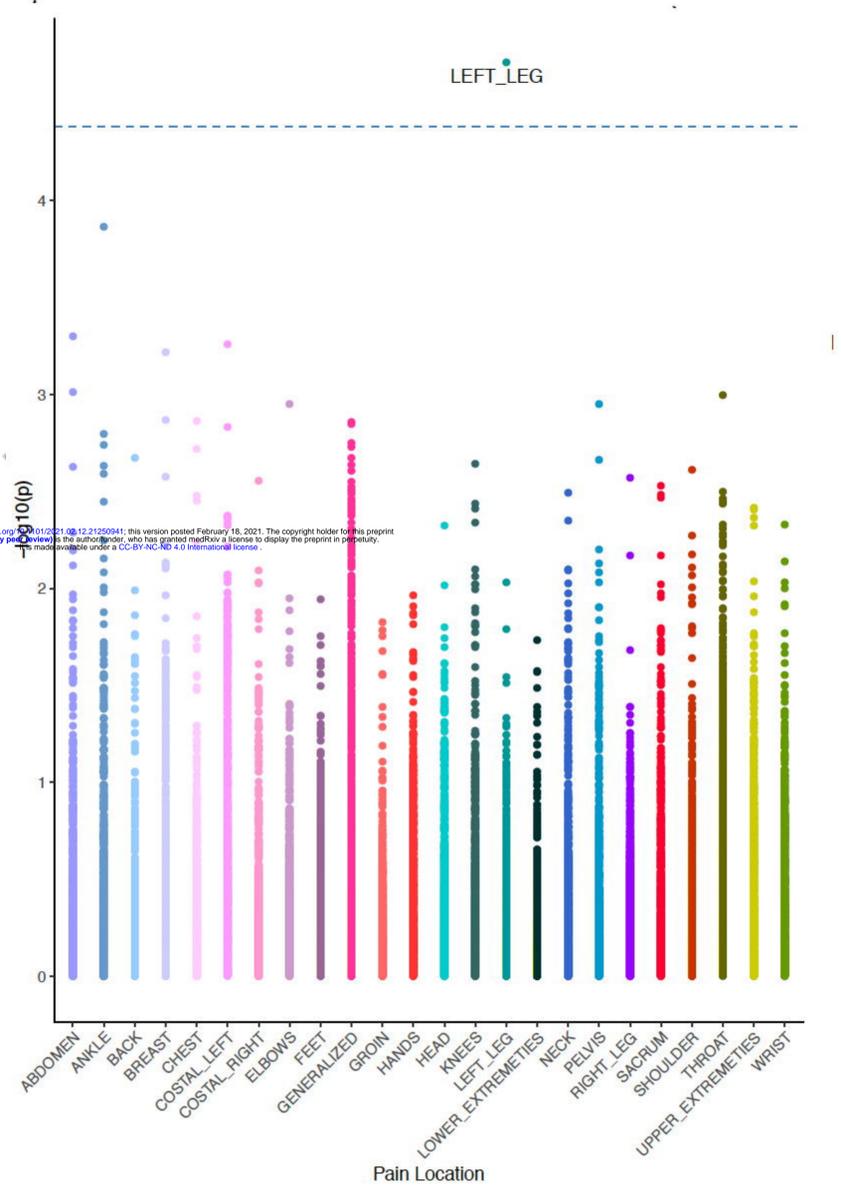
i. Overall



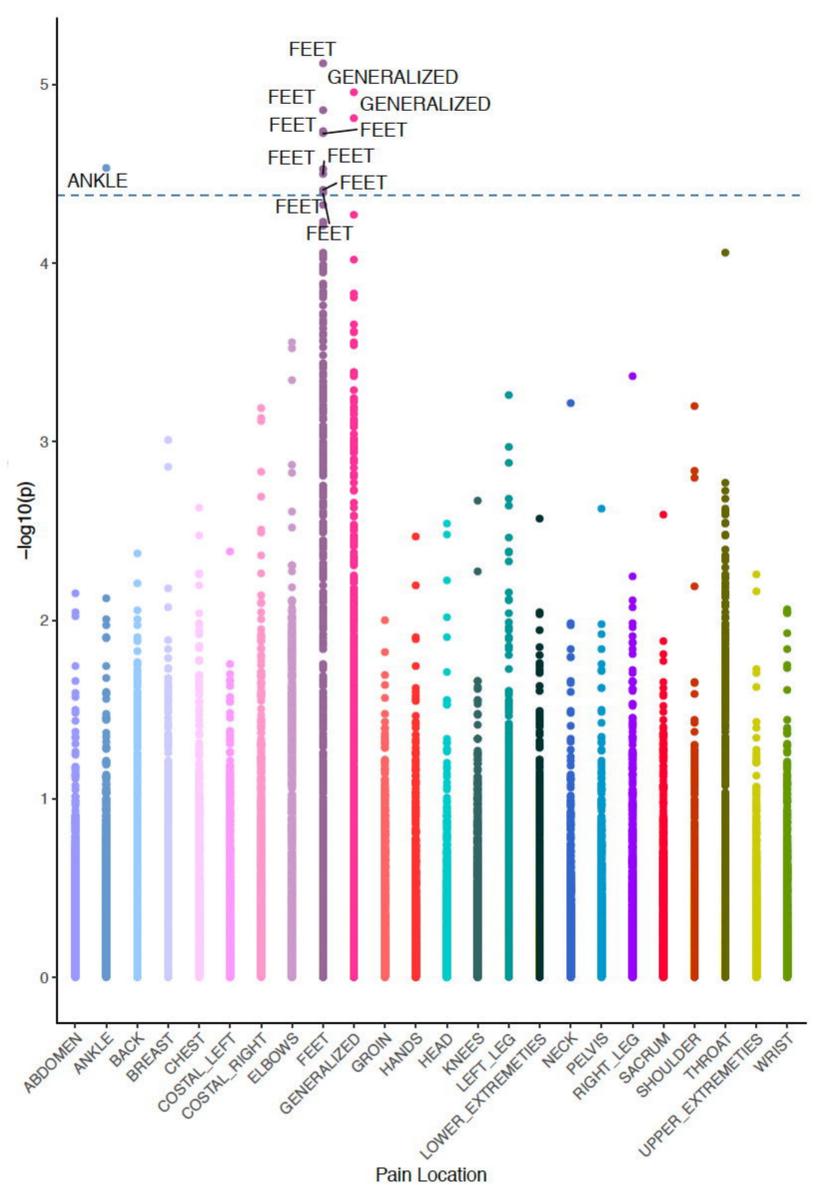
ii. High BMI



iii. Mid BMI



iv. Low BMI



**BioMe BioBank Demographic Information**

|              | <b>Total N</b> | <b>N Self-report Female (%)</b> | <b>N Self-report Male (%)</b> | <b>Mean Age (1stQ-3rdQ)</b> | <b>Mean BMI (1stQ-3rdQ)</b> |                     |
|--------------|----------------|---------------------------------|-------------------------------|-----------------------------|-----------------------------|---------------------|
| Total Cohort | 31585          | 17968 (57)                      | 12417 (39)                    | 57.12 (45-69)               | 28.69 (27.46-32.09)         |                     |
| Ancestry     | AA             | 8044                            | 4945 (61)                     | 2937 (37)                   | 55.76 (45-66)               | 30.63 (25.06-34.81) |
|              | EA             | 10788                           | 5273 (49)                     | 4766 (44)                   | 58.79 (45-72)               | 26.58 (22.64-29.39) |
|              | ESA            | 957                             | 567 (59)                      | 332 (35)                    | 50 (38-60)                  | 24.26 (21.22-26.21) |
|              | HA             | 11047                           | 6863 (62)                     | 3980 (36)                   | 57.42 (45-70)               | 29.64 (25.17-32.95) |
|              | SAS            | 661                             | 275 (42)                      | 360 (54)                    | 52.93 (39-66)               | 26.14 (23.04-28.52) |
|              | NA             | 88                              | 45 (51)                       | 42 (48)                     | 57.03 (46.5-69)             | 28.35 (24.24-31.43) |

| Chromosome | Genes   | Zscore    | P-value         | Top Finding (Gene) | Top Finding (Tissue)       |
|------------|---|-----------|-----------------|--------------------|----------------------------|
| 1          | <i>FBLIM1</i>   | -4.851060 | 1.23e-06        | <i>FBLIM1</i>      | Liver                      |
| 2          | <i>ASB3, CHAC2, GPR75</i>   | -5.780228 | <b>7.46e-09</b> | <i>GPR75</i>       | Whole_Blood                |
| 3          | <i>SPINK8, PFKFB4, TMEM89, SLC26A6, CELSR3, NCKIPSD, ARIH2OS, ARIH2, P4HTM, WDR6, DALRD3, NDUFAF3, USP19, LAMB2, CCDC71, CCDC36, C3orf62, GPX1, NICN1, DAG1, APEH, MST1, MST1R, RNF123, RBM6, RBM5, SEMA3F, TUSC2</i> | 7.321127  | <b>2.46e-13</b> | <i>WDR6</i>        | Skin_Sun_Exposed_Lower_leg |
| 3          | <i>PROS1, DHFRL1</i>  | 4.687382  | 2.77e-06        | <i>DHFRL1</i>      | DGN_Whole_Blood            |
| 6          | <i>CLIC1</i>  | 4.814629  | 1.47e-06        | <i>CLIC1</i>       | Nerve_Tibial               |
| 10         | <i>MGMT, EBF3</i>   | -5.086391 | 3.65e-07        | <i>MGMT</i>        | Thyroid                    |
| 11         | <i>ARNTL</i>  | 4.534957  | 5.76e-06        | <i>ARNTL</i>       | Whole_Blood                |
| 12         | <i>SUOX, RPS26, SLC26A10</i>  | 4.628307  | 3.69e-06        | <i>RPS26</i>       | Uterus                     |
| 17         | <i>TNFSF12, LINC00324</i>   | 4.982641  | 6.27e-07        | <i>TNFSF12</i>     | Breast_Mammary_Tissue      |
| 20         | <i>SLC2A10</i>  | 4.921176  | 8.60e-07        | <i>SLC2A10</i>     | Nerve_Tibial               |
| 22         | <i>KREMEN1</i>  | 4.872670  | 1.10e-06        | <i>KREMEN1</i>     | Nerve_Tibial               |
| 22         | <i>SCO2</i>   | -4.560015 | 5.11e-06        | <i>SCO2</i>        | Brain_Cerebellum           |

1 medRxiv preprint doi: <https://doi.org/10.1101/2021.02.17.21250911>; this version posted February 18, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).