

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

2 Jessica S. Johnson, MPH^{1,2}, Alanna C. Cote^{1,2}, Amanda Dobbyn, PhD¹⁻³, Laura G. Sloofman, MS⁴, Jiayi
3 Xu, PhD^{1,2}, Liam Cotter^{1,2}, Alexander W. Charney, MD, PhD^{1-3,5,6}, Eating Disorders Working Group of the
4 Psychiatric Genomics Consortium^{*}, Jennifer Jordan, PhD⁷, Martin Kennedy, PhD⁷, Mikael Landén, MD,
5 PhD^{8,9}, Sarah L Maguire, PhD¹⁰, Nicholas G Martin, PhD¹¹, Preben Bo Mortensen, MD, DMSc¹², Cynthia
6 M. Bulik, PhD^{9,13,14}, Laura M. Huckins, PhD^{1-6†}

7

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

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

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

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

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

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

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

21 ⁸Gothenburg University, Box 100, SE-405 30 Gothenburg, Sweden

22 ⁹Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden

23 ¹⁰School of Medicine, InsideOut Institute, Sydney, New South Wales, 2006, Australia

24 ¹¹QIMR Berghofer Medical Research Institute, Locked Bag 2000, Royal Brisbane Hospital QLD 4029,
25 Australia

26 ¹²Aarhus University, Norde Ringgade 1, 8000 Aarhus, Denmark

27 ¹³Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27517, USA

28 ¹⁴Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC 27517, USA

* 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., 3rd Floor, Room L3-70i, New York, NY 10029; 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_{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 GTE_x, 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_{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 **BioMeTM**

137 The BioMeTM 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 BioMeTM 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 BioMeTM demographics are shown in **Table 1**). Quality control
144 and imputation of the genotyping data for BioMeTM is described in elsewhere(28).

145

146 Ancestry for BioMeTM 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 1st quartile as “Low” BMI; those above the 3rd quartile as
228 “High” BMI, and those falling between the 1st and 3rd 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 **Testing for hidden case contamination**

231 It is possible that some of the clinical associations within our study stem from undiagnosed ED- and AN-
232 cases, rather than a direct effect of gene expression on phenomic expression. We term this “diagnostic
233 contamination”. To test whether this effect may drive the associations we observe, we simulated the
234 effects of diagnostic contamination occurring at rate p within our biobank sample, with effect β on GReX.

235
236 Full derivations and simulations are shown in our **Supplementary Material**; briefly, we derive the
237 expression among controls, and cases contaminated at rate p ; the difference in expression between the
238 two groups; the expected variance among cases, and pooled across all samples, and the expected
239 statistical significance (T-score, and p-value). Next, we simulated gene expression distributions for (i)
240 1000 cases and 1000 controls; (ii) 1000 cases and 10000 controls; (iii) 1000 cases and 30000 controls;
241 each for a range of p (0.1%, 0.5%, 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%) and β (1/10, 1/5, 1/4, 1/3, 1/2,
242 1, 2,3,4,5,10) values. We repeated our simulations 10000 times at each case-control proportion- p - β
243 combination, and demonstrated that our formulae accurately estimate the desired values (**Figure S3**).

244
245 Next, in order to test whether diagnostic contamination may account for the associations observed within
246 our PheWAS, we calculated the expected impact of diagnostic contamination for two genes (*NCKIPSD*-
247 Aorta; *SEMA3F*-Spinal Cord) with the largest effect sizes β observed in our S-PrediXcan analysis, under
248 two scenarios:

- 249 1. Diagnostic contamination occurs among our cases only, assuming the normal population rate for
250 AN: 0.9% among women; 0.3% among men.
- 251 2. Contamination occurs at significantly higher levels than might be expected in the population
252 (1.2%, 3%, 6%), and all of these samples fall into our case category.

253
254 For each scenario, we repeated our calculation using 500 and 1000 cases, matched with 1000, 5000, and
255 30000 controls.

256

257 **RESULTS**

258

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

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

276

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

278 We tested whether GReX of AN-associated genes was associated with anthropometric phenotypes in our
279 BioMe™ biobank (**Table 3**). In the overall cohort, upregulation of MHC-gene *CLIC1* was associated with
280 lowest recorded weight (Spleen-*CLIC1*, $p = 6.63 \times 10^{-4}$; Breast Mammary tissue-*CLIC1*, $p = 4.18 \times 10^{-4}$) and
281 mean height (Subcutaneous Adipose-*CLIC1*, $p = 1.39 \times 10^{-5}$; **Table S6**). Downregulation of multiple genes,
282 *ARIH2*, *NCKIPSD*, *NDUFAF3*, *NICN1*, *P4HTM*, *RNF123*, *SLC26A10*, was associated with decreased
283 measures of weight change over time in the overall cohort ($p < 9.86 \times 10^{-4}$)(**Table 3, Figure S4**).
284 Upregulation of *WDR6* was associated with increased measures of weight change over time ($p < 9.81 \times 10^{-4}$)

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

287

288 ***Patient BMI mediates AN-GREX-PheWAS associations***

289 Given the relationship of AN with BMI, we performed a secondary analysis to assess the effect of BMI on
290 the association between AN-GReX and anthropometric phenotypes. Specifically we wanted to look at
291 whether our AN-GReX-phenotype associations in individuals of High and Low BMI were still present
292 when BMI was within a normal range. We split the BioMe™ cohort into three groups: High, Mid, and Low
293 BMI (adjusted for ancestry and sex; see **Methods** and **Supplementary Information** for more details),
294 and re-ran our PheWAS within these groups. Indeed, our AN-GReX anthropometry associations (with
295 highest recorded weight, lowest recorded weight, and weight change over time) varied according to the
296 BMI group tested (**Figure 2, Table 3**). The association between *CLIC1*-GReX and highest and lowest
297 recorded weight measures was strongest in individuals of High BMI compared with the other BMI groups
298 (**Figure 2, Table 3**). Likewise, for multiple genes, associations between AN-GReX and weight change
299 over time differed markedly in the High BMI group (**Table 3, Figure S5**).

300

301 ***AN-GREX is associated with autoimmune and autoinflammatory diseases***

302 At a false discovery rate (FDR) $p < 0.05$, we found 17 FDR-significant gene-tissue associations with four
303 Phecode- and Encounter Diagnosis phenotypes: type I diabetes, celiac disease, peptic ulcer, and
304 unspecified immunodeficiency (**Figure 3**). Upregulation of *CLIC1* in spleen was associated with two
305 autoimmune/autoinflammatory phenotypes in the overall BioMe™ cohort: celiac disease and type I
306 diabetes ($p = 1.30 \times 10^{-11}$, $p = 5.08 \times 10^{-20}$, respectively) (**Tables S7 and S8**). Downregulation of *SLC26A6*
307 was associated with the phenotype “peptic ulcer, site unspecified” in the overall cohort (Liver-*SLC26A6*,
308 $p = 8.38 \times 10^{-8}$; Skeletal muscle-*SLC26A6*, $p = 1.71 \times 10^{-7}$; Thyroid-*SLC26A6*, $p = 7.16 \times 10^{-8}$).
309 Downregulation of *PFKFB4* was associated with ICD10 code D84.9, “immunodeficiency, unspecified”
310 (Skin not sun exposed-*PFKFB4*, $p = 1.34 \times 10^{-7}$). We stratified our cohort by BMI (see **Methods**) to assess
311 whether BMI was modifying the association of AN-GReX with Encounter Diagnosis and phecodes. We
312 saw significant associations of AN-GReX phecodes within our Mid BMI group; however, when plotted

313 across the BMI groups, there was no notable difference in the magnitude and direction of effect of AN-
314 GReX on either Encounter Diagnosis or phecodes across the BMI groups (**Figures S6 and S7**). Within
315 individuals with Mid BMI, upregulation of *CLIC1* was similarly associated with type 1 diabetes (Mid-
316 Subcutaneous adipose, Mid-Breast Mammary tissue, Mid-Spleen-*CLIC1*, $p < 1.96 \times 10^{-8}$).

317

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

319 Multiple AN genes were associated with both categorical and continuous measures of tobacco use (**Table**
320 **S9, Figure S8**). *APEH*, *CCDC36*, *DALRD3*, *GPX1*, *LINC00324*, *NICN1*, and *P4HTM* were associated
321 with years used/smoked of tobacco in the overall BioMe™ cohort ($p < 4.97 \times 10^{-4}$, **Table S9**) and
322 upregulation of *MGMT* was associated with being a smoker (Heart Atrial Appendage-*MGMT*, $p = 8.87 \times 10^{-5}$;
323 Putamen basal ganglia-*MGMT*, $p = 1.09 \times 10^{-4}$). Again, we note mediating effects of patient BMI category
324 on the association between AN-GReX and smoking phenotypes. Among the Mid and High BMI groups,
325 different genes were associated with both continuous and categorical measures of tobacco use (**Table**
326 **S9, Figure S8**), including years of tobacco use, tobacco cigarette packs per day, and smoking every day.
327 In addition to tobacco use measures, we see an association between AN genes and tobacco use related
328 ICD codes. Downregulation of *LINC00324* in Low BMI individuals was associated with ICD10 diagnosis
329 code F17.210 “Nicotine dependence, cigarettes, uncomplicated” (Low-Cells EBV-transformed-*LINC00324*,
330 $p = 3.03 \times 10^{-5}$).

331

332 Downregulation of *MGMT* was associated with decreased alcohol use (ounces per week) in individuals in
333 the Mid-BMI group ($p < 9.75 \times 10^{-4}$; **Figure S8**). In individuals of Low BMI, downregulation of *GPR75* and
334 *GPX1*, and upregulation of *NICN1* were associated with decreased and increased continuous measures
335 of illicit drug use frequency respectively (Liver-*GPR75*, $p = 3.34 \times 10^{-4}$; Nucleus accumbens-*GPX1*, $p = 2.04$
336 $\times 10^{-4}$; Testis-*NICN1*, $p = 3.34 \times 10^{-4}$; **Figure S8**).

337

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

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

341 pain score measurement and location across all BMI groups (**Table S10**). Most significantly, upregulation
342 of *TNFSF12* was associated with total knee pain score (sum) in individuals with High BMI (High-Brain
343 Substantia nigra-*TNFSF12*, $p=1.50 \times 10^{-3211}$). Downregulation of *TNFSF12* was associated with total
344 abdomen pain score (sum) in individuals with Low BMI (Low-Brain Substantia nigra-*TNFSF12*, $p=3.93 \times$
345 10^{-58}) (**Table S10**).

346
347 When looking at pain location irrespective of pain score, upregulation of *TUSC2* in the overall cohort is
348 associated with presence of knee pain (Pancreas-*TUSC2*, $p=7.62 \times 10^{-6}$) and downregulation of *CTNNB1*
349 in Mid BMI individuals is associated with the presence of left leg pain (Mid-Stomach-*CTNNB1*, $p=1.93 \times$
350 10^{-5}) (**Figure 4, Table S10**). Repeating this analysis according to BMI category reveals a very large
351 number of associations specific to the Low BMI category (**Figure 4, Table S10**); in particular, we note 129
352 gene-tissue associations with presence of foot pain, nine of which pass phenotype-wide significance,
353 which may indicate a propensity to excess exercise, or exercise-related injuries among these individuals.

354

355 ***Upregulation of CLIC1 is associated with glucagon medication***

356 AN-GReX PheWAS with medication phenotypes was associated with glucagon, a hormone used to treat
357 severe hypoglycemia, in the overall BioMe™ cohort. Upregulation of *CLIC1* was associated with glucagon
358 in the overall cohort (Spleen-*CLIC1*, $p=4.20 \times 10^{-9}$) and also in individuals with Mid BMI (Mid-Spleen-
359 *CLIC1*, $p=2.09 \times 10^{-6}$) (**Figure S9, Table S11**).

360

361 **Identifying potential case contamination effects**

362 Our BioMe™ patients have not all been explicitly assessed for eating disorders (ED), and information
363 regarding ED diagnoses and assessments earlier in life may be omitted from the records due to
364 incomplete clinical history assessments. Therefore, it is possible that diagnostic contamination within
365 some of our sample is responsible for the associations observed within our data. In order to test this, we
366 performed a simple experiment.

367

368 For this simple experiment, we assume that diagnostic contamination occurs among true cases only, at
369 prevalence p . In this case, we expect gene expression among those individuals to be up- or down-
370 regulated by effect size β ; consequently, we expect a gene expression difference between cases and
371 controls proportional to the prevalence of the contamination within the sample (p) and the effect size
372 (β); (**Supplementary information**). We tested whether such contamination may drive the associations
373 observed in our study, assuming the following different contamination scenarios (**Table S12**), for two
374 genes with large S-PrediXcan effect sizes: (i) that contamination occurs among cases, assuming the
375 highest common estimate for AN prevalence (0.9% among women, 0.3% among men); (ii) that
376 contamination occurs within our biobank, assuming the highest common estimate for AN prevalence
377 (0.6%, 190 cases), and that all of these samples fall into our 'case' category.

378
379 Our model indicates that diagnostic contamination among cases is unlikely to result in significant
380 PheWAS associations in our data. Contamination occurring at population ED prevalence (0.6%) did not
381 result in significant associations for any of the case-control scenarios in our model (**Table S12**); nor did
382 contamination at 1.2% or 3%. Assuming 6% contamination resulted in potentially significant associations
383 (estimated $p < 2.8 \times 10^{-8}$; see **Methods** and **Supplementary Methods** for more detail).

384

385 **DISCUSSION**

386

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

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

411

412 *CLIC1* (Chloride Intracellular 1) encodes a chloride ion channel protein involved in many necessary
413 cellular functions including the regulation of cell membrane potential, and the proliferation and
414 differentiation of cells(57), including a role in axonal outgrowth of neurons⁵¹. In addition to neurons,
415 *CLIC1* has been found to be expressed in many different tissue types(59), and in particular, *CLIC1*
416 protein has been found to accumulate in peripheral blood mononuclear cells (PBMCs) when the central
417 nervous system is in a state of chronic inflammation(58), and increased *CLIC1* expression has been
418 shown to occur in microglia stimulated by the presence of β -amyloid *in vitro*(60,61). Therefore, *CLIC1*
419 may potentially act as a biomarker for neurodegenerative disorders such as Alzheimer's disease, and
420 downregulation of *CLIC1* has been shown to have a neuroprotective effect in the presence of β -amyloid
421 proteins(58,60–62). Variants in *CLIC1* may contribute to risk of multiple psychiatric disorders and
422 behaviors, such as schizophrenia, MDD and PTSD, as well as many autoimmune and immune-related
423 phenotypes such as asthma and systemic lupus erythematosus(50–53)(63) (**Table S5**). Many studies
424 have posed the role of inflammation in the development of psychiatric disorders, and in the etiology of
425 AN(7,8,45,46,64).

426
427 Multiple genes (8/47 or 17%) in our results were associated with regular gym attendance (**Table S5**), or
428 measures of physical activity. Excessive and compulsive exercise is a behavior often seen in individuals
429 with AN(65,66), and evidence is suggestive that hyperactivity increases risk of chronic AN(67,68).
430 Likewise, genetic studies have shown a strong genetic correlation between AN and physical activity
431 ($r_g=0.17$)(9,23,69). These genes may reflect a genetic liability to compulsive behaviors of physical activity.
432 Our pain location results point to possible symptoms of excessive activity, with the strong association of
433 AN-GReX with foot pain in Low BMI individuals. Given the known association of AN disease with low
434 bone mineral density (BMD) and propensity for bone fracture(70), our results may reflect the result of
435 genes associated with compulsive exercise and BMD contributing to increased osteoarticular pain. We
436 also see multiple genes that have been previously associated with decreased sleep and insomnia
437 phenotypes (including a core circadian clock gene *ARNTL*), which have previously been implicated in a
438 range of psychiatric disorders(71–82) , including eating disorders(83), as well as in satiety and
439 hunger(84,85).

440

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

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

464

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

468 of AN with metabolic traits and abnormalities has been fairly well established. The Anorexia Nervosa
469 Genetics Initiative reported significant negative genetic correlation between AN and a number of
470 metabolic traits including insulin resistance ($r_g=-0.29$), fasting insulin ($r_g=-0.24$), leptin ($r_g=-0.26$) and type
471 2 diabetes ($r_g=-0.22$), and a strong positive genetic correlation with HDL cholesterol ($r_g=0.21$)(9).
472 Additional studies have shown associations of AN with fasting insulin(87) and insulin sensitivity(88).
473 Notably, our results point to a similar role of aberrant glycemc regulation in the etiology of AN. Future
474 analyses including EHR-derived lab results (LabWAS) studies may further elucidate AN genes associated
475 with abnormal metabolic regulation and clinical features.

476

477 ***BMI modification of AN-GReX on clinical outcomes***

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

492

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

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

502
503 Although the hypothesis-free, phenome-wide design of our study allows for powerful detection of clinical
504 and biological consequences of AN risk genes, the same design also bears some notable limitations and
505 caveats. One key caveat to our results is the lack of diagnostic detail and insights available for the
506 patients in our study. EHR analyses leverage large, existing data sets to rapidly amass cohorts for
507 analysis, and to yield insights into whole phenome consequences of genotype and GReX associations;
508 however, the scale and scope of these studies precludes deep phenotyping or performing clinical
509 interviews. This lack of diagnostic precision may arise from a number of factors.

510
511 First, we make use of ICD codes within the medical record in order to infer diagnoses; since these are
512 primarily used by clinicians for billing purposes, they likely provide an imperfect proxy for true disease
513 state. In order to mitigate spurious results stemming from imperfectly assigned codes, we established
514 criteria for inclusion of diagnoses in our analyses: First, diagnosis codes with fewer than 10 counts in the
515 biobank were removed. Second, we only included diagnoses in our pheWAS with at least 10 counts of a
516 code in the ancestral population (i.e., ancestral groups with fewer than 10 case counts were not included
517 in the analysis for that diagnosis), and with a total effective sample size after meta-analysis that was
518 greater than 100. Lastly, within each pheWAS, individuals had to have two or more counts of a diagnosis
519 code to count as a “case”, whereas individuals with only one count of a diagnosis code were set to
520 “missing” and not included in the analysis.

521
522 Second, it is possible that our patients regularly receive treatment at multiple different hospital systems;
523 as such, we may be capturing only partial data for each of our patients. In order to mitigate this, we

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

527
528 Importantly, our patients have not all been explicitly assessed for EDs, and information regarding ED
529 diagnoses and assessments earlier in life may be omitted from the records due to incomplete clinical
530 history assessments. Therefore, it is possible that case-contamination within some of our sample is
531 responsible for the associations observed within our data. In order to address this, we performed a simple
532 experiment to simulate the effect sizes expected if undiagnosed AN contamination drives our result. We
533 estimated the expected association statistics observed due to two possible levels of contamination among
534 PheWAS cases; i, that 0.06% of PheWAS-trait cases are missing a true AN diagnosis- i.e., the proportion
535 of individuals we expect to receive an AN diagnosis based on standard prevalence estimates; ii, that
536 1.2%, 3% and 6% of individuals in BioMe are missing an AN diagnosis (i.e., double, five times, and ten
537 times the expected population rate), and that all of these individuals fall within our PheWAS trait case
538 group, and applied this model for two genes with large S-PrediXcan effect sizes (*NCKIPSD-Aorta*;
539 *SEMA3F-Spinal Cord*). Our model indicates that diagnostic contamination at 0.06%, 1.2% or 3% is
540 insufficient to account for the effect sizes observed within our PheWAS analysis. Moreover, since these
541 genes are selected due to their large S-PrediXcan effect sizes, we expect that the contamination effects
542 observed for these two genes will be greater than any others in our study; as such, we do not believe that
543 our findings are attributable to hidden case contamination. Contamination at 6% resulted in significant
544 differences; however, since this estimate represents a >five-fold increase in expected numbers of AN
545 within a population, we do not expect that this level of contamination occurs in our sample.

546
547 Our results demonstrate that there are real clinical consequences to differences in AN gene expression.
548 Characterization of the phenotypic consequences of AN gene expression in a clinical setting can give us
549 more insight into the biological mechanisms underlying AN and, consequently, how to diagnose and treat
550 the disorder. By understanding the associations of AN gene expression with symptomatology, prodromal
551 or subthreshold disease states, we may gain insights into the biology of the disease, and perhaps identify

552 therapeutic targets and opportunities for clinical interventions. For example, if gastrointestinal complaints
553 are truly the consequence of aberrant AN gene expression, and contribute to disordered eating due to
554 gastrointestinal distress, treatment of those symptoms may help alleviate other AN symptoms or prevent
555 development of later AN(92,93). An understanding of the clinical consequences of AN gene expression
556 can further augment the definition of AN, and could allow clinicians to more broadly identify individuals at
557 greater risk of AN, or those who present with symptom constellations that do not yet meet the established
558 diagnostic threshold.

559 **Acknowledgements**

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

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

562

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

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

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

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

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

568

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

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

571 Medicine at Mount Sinai.

572

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

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

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

576 of Health.

577

578 **Competing Interests**

579 CMB reports: Shire (grant recipient, Scientific Advisory Board member); Idorsia (consultant);

580 Pearson (author, royalty recipient). ML declares that over the past 36 months, he has received lecture

581 honoraria from Lundbeck pharmaceutical (No other equity ownership, profit-sharing agreements,

582 royalties, or patent). The remaining authors declare no competing interests.

583

584

585 **REFERENCES:**

586

- 587 1. Zipfel S, Giel KE, Bulik CM, Hay P, Schmidt U. Anorexia nervosa: aetiology, assessment, and
588 treatment. *Lancet Psychiatry*. 2015 Dec;2(12):1099–111.
- 589 2. Bulik CM, Flatt R, Abbaspour A, Carroll I. Reconceptualizing anorexia nervosa. *Psychiatry Clin
590 Neurosci*. 2019;73(9):518–25.
- 591 3. Mitchell JE, Peterson CB. Anorexia Nervosa. *N Engl J Med*. 2020 02;382(14):1343–51.
- 592 4. Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a
593 meta-review. *World Psychiatry Off J World Psychiatr Assoc WPA*. 2014 Jun;13(2):153–60.
- 594 5. Ulfvebrand S, Birgegård A, Norring C, Högdahl L, von Hausswolff-Juhlin Y. Psychiatric comorbidity in
595 women and men with eating disorders results from a large clinical database. *Psychiatry Res*. 2015
596 Dec 15;230(2):294–9.
- 597 6. Root TL, Pinheiro AP, Thornton L, Strober M, Fernandez-Aranda F, Brandt H, et al. Substance use
598 disorders in women with anorexia nervosa. *Int J Eat Disord*. 2010 Jan;43(1):14–21.
- 599 7. Raevuori A, Haukka J, Vaarala O, Suvisaari JM, Gissler M, Grainger M, et al. The increased risk for
600 autoimmune diseases in patients with eating disorders. *PLoS One*. 2014;9(8):e104845.
- 601 8. Hedman A, Breithaupt L, Hübel C, Thornton LM, Tillander A, Norring C, et al. Bidirectional relationship
602 between eating disorders and autoimmune diseases. *J Child Psychol Psychiatry*. 2019;60(7):803–
603 12.
- 604 9. Anorexia Nervosa Genetics Initiative, Eating Disorders Working Group of the Psychiatric Genomics
605 Consortium, Watson HJ, Yilmaz Z, Thornton LM, Hübel C, et al. Genome-wide association study
606 identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nat Genet*.
607 2019 Aug;51(8):1207–14.
- 608 10. Yilmaz Z, Hardaway JA, Bulik CM. Genetics and Epigenetics of Eating Disorders. *Adv Genomics
609 Genet*. 2015;5:131–50.
- 610 11. Levinson CA, Zerwas SC, Brosf LC, Thornton LM, Strober M, Pivarunas B, et al. Associations
611 between dimensions of anorexia nervosa and obsessive-compulsive disorder: An examination of
612 personality and psychological factors in patients with anorexia nervosa. *Eur Eat Disord Rev J Eat
613 Disord Assoc*. 2019;27(2):161–72.
- 614 12. Kaye WH, Wierenga CE, Bailer UF, Simmons AN, Bischoff-Grethe A. Nothing tastes as good as
615 skinny feels: the neurobiology of anorexia nervosa. *Trends Neurosci*. 2013 Feb 1;36(2):110–20.
- 616 13. Speakman JR, Levitsky DA, Allison DB, Bray MS, Castro JM de, Clegg DJ, et al. Set points, settling
617 points and some alternative models: theoretical options to understand how genes and environments
618 combine to regulate body adiposity. *Dis Model Mech*. 2011 Nov 1;4(6):733–45.
- 619 14. Müller MJ, Bösny-Westphal A, Heymsfield SB. Is there evidence for a set point that regulates human
620 body weight? *F1000 Med Rep [Internet]*. 2010 Aug 9 [cited 2020 Oct 20];2. Available from:
621 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2990627/>
- 622 15. Hebebrand J, Müller TD, Holtkamp K, Herpertz-Dahlmann B. The role of leptin in anorexia nervosa:
623 clinical implications. *Mol Psychiatry*. 2007 Jan;12(1):23–35.

- 624 16. Hübel C, Yilmaz Z, Schaumberg KE, Breithaupt L, Hunjan A, Horne E, et al. Body composition in
625 anorexia nervosa: Meta-analysis and meta-regression of cross-sectional and longitudinal studies.
626 *Int J Eat Disord.* 2019;52(11):1205–23.
- 627 17. Igudesman D, Sweeney M, Carroll IM, Mayer-Davis EJ, Bulik CM. Gut-Brain Interactions:
628 Implications for a Role of the Gut Microbiota in the Treatment and Prognosis of Anorexia Nervosa
629 and Comparison to Type I Diabetes. *Gastroenterol Clin North Am.* 2019;48(3):343–56.
- 630 18. Huckins LM, Hatzikotoulas K, Southam L, Thornton LM, Steinberg J, Aguilera-Mckay F, et al.
631 Investigation of common, low-frequency and rare genome-wide variation in anorexia nervosa. *Mol*
632 *Psychiatry.* 2018;23(5).
- 633 19. Huckins LM, Dobbyn A, McFadden W, Ruderfer D, Wang W, Gamazon E, et al. Identifying tissues
634 implicated in Anorexia Nervosa using Transcriptomic Imputation. *bioRxiv.* 2018;265017.
- 635 20. Hinney A, Kesselmeier M, Jall S, Volckmar A-L, Föcker M, Antel J, et al. Evidence for three genetic
636 loci involved in both anorexia nervosa risk and variation of body mass index. *Mol Psychiatry.*
637 2017;22(2).
- 638 21. Duncan L, Yilmaz Z, Gaspar H, Walters R, Goldstein J, Anttila V, et al. Significant Locus and
639 Metabolic Genetic Correlations Revealed in Genome-Wide Association Study of Anorexia Nervosa.
640 *Am J Psychiatry.* 2017 01;174(9):850–8.
- 641 22. Munn-Chernoff MA, Johnson EC, Chou Y-L, Coleman JRI, Thornton LM, Walters RK, et al. Shared
642 genetic risk between eating disorder- and substance-use-related phenotypes: Evidence from
643 genome-wide association studies. *Addict Biol.* 2020 Feb 16;e12880.
- 644 23. Hübel C, Gaspar HA, Coleman JRI, Finucane H, Purves KL, Hanscombe KB, et al. Genomics of
645 body fat percentage may contribute to sex bias in anorexia nervosa. *Am J Med Genet Part B*
646 *Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet.* 2019;180(6):428–38.
- 647 24. C H, M A, M H, Rj L, G B, Cm B, et al. Binge-eating disorder, anorexia nervosa, and constitutional
648 thinness differ in their associations with anthropometric and psychiatric polygenic scores. 2020 Mar
649 26 [cited 2020 Aug 20]; Available from: <https://europepmc.org/article/ppr/ppr126993>
- 650 25. Smoller JW. The use of electronic health records for psychiatric phenotyping and genomics. *Am J*
651 *Med Genet Part B Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet.* 2018;177(7):601–12.
- 652 26. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, et al. PheWAS:
653 demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations.
654 *Bioinforma Oxf Engl.* 2010 May;26(9):1205–10.
- 655 27. Pendergrass SA, Brown-Gentry K, Dudek SM, Torstenson ES, Ambite JL, Avery CL, et al. The use
656 of phenome-wide association studies (PheWAS) for exploration of novel genotype-phenotype
657 relationships and pleiotropy discovery. *Genet Epidemiol.* 2011 Jul;35(5):410–22.
- 658 28. Zheutlin AB, Dennis J, Karlsson Linnér R, Moscati A, Restrepo N, Straub P, et al. Penetrance and
659 Pleiotropy of Polygenic Risk Scores for Schizophrenia in 106,160 Patients Across Four Health Care
660 Systems. *Am J Psychiatry.* 2019 Aug 16;176(10):846–55.
- 661 29. Leppert B, Millard LAC, Riglin L, Davey Smith G, Thapar A, Tilling K, et al. A cross-disorder PRS-
662 pheWAS of 5 major psychiatric disorders in UK Biobank. *PLoS Genet.* 2020 May;16(5):e1008185.

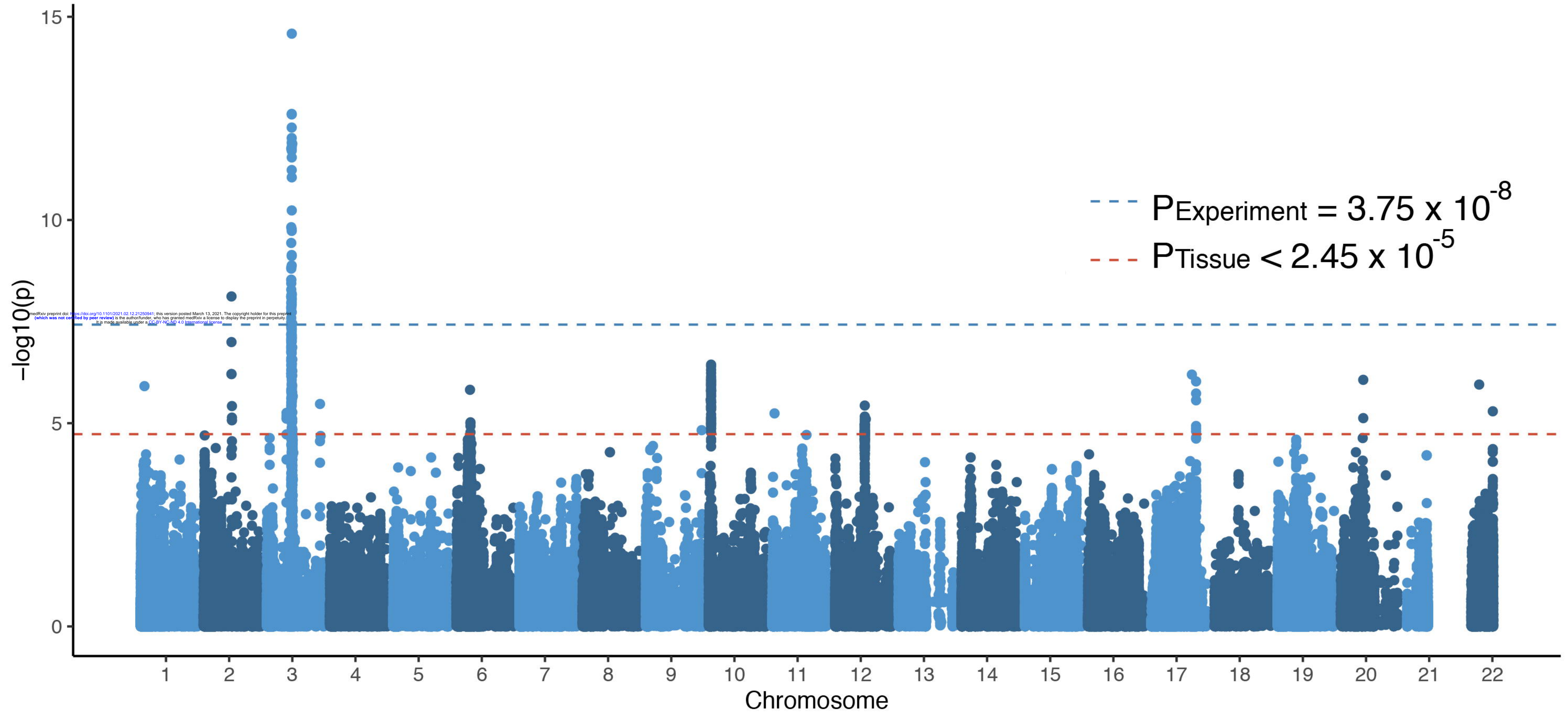
- 663 30. Denny JC, Bastarache L, Ritchie MD, Carroll RJ, Zink R, Mosley JD, et al. Systematic comparison
664 of phenome-wide association study of electronic medical record data and genome-wide
665 association study data. *Nat Biotechnol.* 2013 Dec;31(12):1102–10.
- 666 31. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue
667 Expression (GTEx) project. *Nat Genet.* 2013 Jun;45(6):580–5.
- 668 32. Hoffman GE, Bendl J, Voloudakis G, Montgomery KS, Sloofman L, Wang Y-C, et al. CommonMind
669 Consortium provides transcriptomic and epigenomic data for Schizophrenia and Bipolar Disorder.
670 *Sci Data.* 2019 Sep 24;6(1):180.
- 671 33. Battle A, Mostafavi S, Zhu X, Potash JB, Weissman MM, McCormick C, et al. Characterizing the
672 genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. *Genome Res.*
673 2014 Jan;24(1):14–24.
- 674 34. Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ, et al. A gene-
675 based association method for mapping traits using reference transcriptome data. *Nat Genet.* 2015
676 Sep;47(9):1091–8.
- 677 35. Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM, et al. Exploring the
678 phenotypic consequences of tissue specific gene expression variation inferred from GWAS
679 summary statistics. *Nat Commun.* 2018 May 8;9(1):1825.
- 680 36. Huckins LM, Dobbyn A, Ruderfer DM, Hoffman G, Wang W, Pardiñas AF, et al. Gene expression
681 imputation across multiple brain regions provides insights into schizophrenia risk. *Nat Genet.*
682 2019;51(4):659–74.
- 683 37. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for
684 whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007
685 Sep;81(3):559–75.
- 686 38. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising
687 to the challenge of larger and richer datasets. *GigaScience.* 2015;4:7.
- 688 39. Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, et al. Gene
689 expression elucidates functional impact of polygenic risk for schizophrenia. *Nat Neurosci.*
690 2016;19(11):1442–53.
- 691 40. PheWAS/PheWAS [Internet]. PheWAS; 2019 [cited 2019 Nov 26]. Available from:
692 <https://github.com/PheWAS/PheWAS>
- 693 41. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association
694 scans. *Bioinformatics.* 2010 Sep 1;26(17):2190–1.
- 695 42. Verma A, Bradford Y, Dudek S, Lucas AM, Verma SS, Pendergrass SA, et al. A simulation study
696 investigating power estimates in phenome-wide association studies. *BMC Bioinformatics.* 2018 Apr
697 4;19(1):120.
- 698 43. Wu P, Gifford A, Meng X, Li X, Campbell H, Varley T, et al. Developing and Evaluating Mappings of
699 ICD-10 and ICD-10-CM Codes to PheCodes. *bioRxiv.* 2019 Jul 3;462077.
- 700 44. de Lange KM, Moutsianas L, Lee JC, Lamb CA, Luo Y, Kennedy NA, et al. Genome-wide
701 association study implicates immune activation of multiple integrin genes in inflammatory bowel
702 disease. *Nat Genet.* 2017 Feb;49(2):256–61.

- 703 45. Mårild K, Størdal K, Bulik CM, Rewers M, Ekblom A, Liu E, et al. Celiac Disease and Anorexia
704 Nervosa: A Nationwide Study. *Pediatrics*. 2017 May;139(5).
- 705 46. S Z, Jt L, L P, Lm T, M Q, Sv K, et al. Eating Disorders, Autoimmune, and Autoinflammatory
706 Disease. *Pediatrics* [Internet]. 2017 Nov 9 [cited 2020 Nov 16];140(6). Available from:
707 <https://europepmc.org/article/med/29122972>
- 708 47. Gibson D, Mehler PS. Anorexia Nervosa and the Immune System—A Narrative Review. *J Clin Med*.
709 2019 Nov;8(11):1915.
- 710 48. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetsky V, et al. Genome-wide
711 association study identifies 30 loci associated with bipolar disorder. *Nat Genet*. 2019 May;51(5):793.
- 712 49. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108
713 schizophrenia-associated genetic loci. *Nature*. 2014 Jul;511(7510):421–7.
- 714 50. The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium. Meta-
715 analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus
716 at 10q24.32 and a significant overlap with schizophrenia. *Mol Autism*. 2017 Dec;8(1):21.
- 717 51. Zhu Z, Zhu X, Liu C-L, Shi H, Shen S, Yang Y, et al. Shared genetics of asthma and mental health
718 disorders: a large-scale genome-wide cross-trait analysis. *Eur Respir J* [Internet]. 2019 Dec 1 [cited
719 2020 May 21];54(6). Available from: <https://erj.ersjournals.com/content/54/6/1901507-0>
- 720 52. Marchese S, Hoffman G, Pietrzak RH, Lebovitch D, Daskalakis NP, Yehuda R, et al. Altered gene
721 expression and PTSD symptom dimensions in World Trade Center responders. In Prep. 2020;
- 722 53. Baselmans BML, Jansen R, Ip HF, van Dongen J, Abdellaoui A, van de Weijer MP, et al.
723 Multivariate Genome-wide and integrated transcriptome and epigenome-wide analyses of the Well-
724 being spectrum [Internet]. *Genetics*; 2017 Mar [cited 2020 May 21]. Available from:
725 <http://biorxiv.org/lookup/doi/10.1101/115915>
- 726 54. Yang X, Sun J, Gao Y, Tan A, Zhang H, Hu Y, et al. Genome-Wide Association Study for Serum
727 Complement C3 and C4 Levels in Healthy Chinese Subjects. *PLOS Genet*. 2012 Sep
728 13;8(9):e1002916.
- 729 55. Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from
730 complex variation of complement component 4. *Nature*. 2016;530(7589):177–83.
- 731 56. Stephan AH, Barres BA, Stevens B. The complement system: an unexpected role in synaptic
732 pruning during development and disease. *Annu Rev Neurosci*. 2012;35:369–89.
- 733 57. Li B, Mao Y, Wang Z, Chen Y, Wang Y, Zhai C, et al. CLIC1 Promotes the Progression of Gastric
734 Cancer by Regulating the MAPK/AKT Pathways. *Cell Physiol Biochem*. 2018;46(3):907–24.
- 735 58. Carlini V, Verduci I, Cianci F, Cannavale G, Fenoglio C, Galimberti D, et al. CLIC1 Protein
736 Accumulates in Circulating Monocyte Membrane during Neurodegeneration. *Int J Mol Sci*. 2020
737 Jan;21(4):1484.
- 738 59. Facal F, Costas J. Evidence of association of the DISC1 interactome gene set with schizophrenia
739 from GWAS. *Prog Neuropsychopharmacol Biol Psychiatry*. 2019 20;95:109729.
- 740 60. Milton RH, Abeti R, Averaimo S, DeBiasi S, Vitellaro L, Jiang L, et al. CLIC1 function is required for
741 beta-amyloid-induced generation of reactive oxygen species by microglia. *J Neurosci Off J Soc*
742 *Neurosci*. 2008 Nov 5;28(45):11488–99.

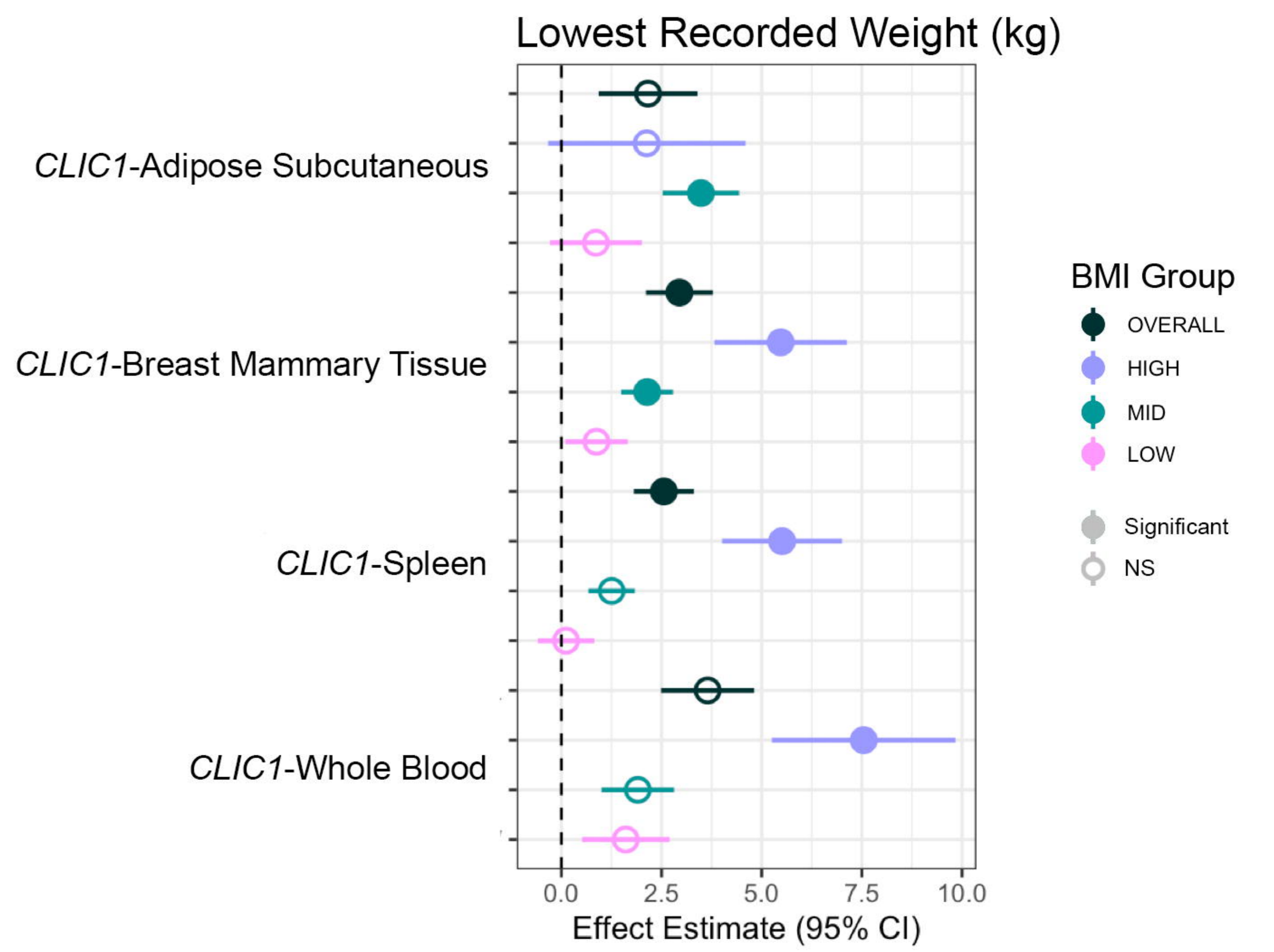
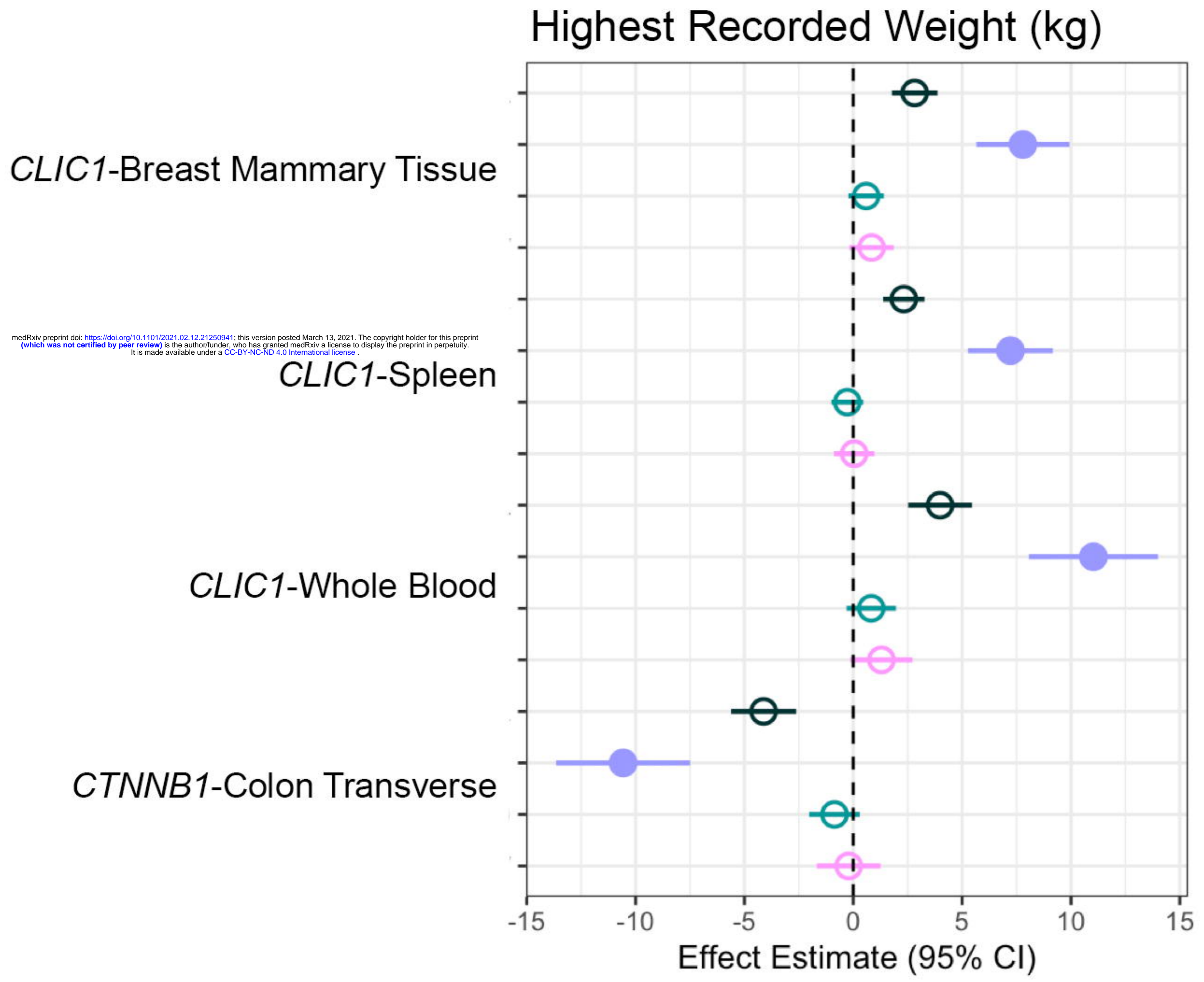
- 743 61. Novarino G, Fabrizi C, Tonini R, Denti MA, Malchiodi-Albedi F, Lauro GM, et al. Involvement of the
744 intracellular ion channel CLIC1 in microglia-mediated beta-amyloid-induced neurotoxicity. *J*
745 *Neurosci Off J Soc Neurosci*. 2004 Jun 9;24(23):5322–30.
- 746 62. Swarup V, Chang TS, Duong DM, Dammer EB, Dai J, Lah JJ, et al. Identification of Conserved
747 Proteomic Networks in Neurodegenerative Dementia. *Cell Rep*. 2020 Jun 23;31(12):107807.
- 748 63. Lee YH, Bae S-C, Choi SJ, Ji JD, Song GG. Genome-wide pathway analysis of genome-wide
749 association studies on systemic lupus erythematosus and rheumatoid arthritis. *Mol Biol Rep*. 2012
750 Dec 1;39(12):10627–35.
- 751 64. Nilsson IAK, Millischer V, Göteson A, Hübel C, Thornton LM, Bulik CM, et al. Aberrant inflammatory
752 profile in acute but not recovered anorexia nervosa. *Brain Behav Immun*. 2020 Aug 1;88:718–24.
- 753 65. Bulik CM. Genetic and Biological Risk Factors. In: Thompson JK, editor. *Handbook of Eating*
754 *Disorders and Obesity* [Internet]. Hoboken, NJ: Wiley; 2004. p. 3. Available from:
755 https://books.google.com/books/about/Handbook%7B_%7Dof%7B_%7DEating%7B_%7DDisorders%7B_%7Dand%7B_%7DObesity.html?id=KZhHAAAAMAAJ%7B%7Dpgis=1
756
- 757 66. Shroff H, Reba L, Thornton LM, Tozzi F, Klump KL, Berrettini WH, et al. Features associated with
758 excessive exercise in women with eating disorders. *Int J Eat Disord*. 2006 Sep;39(6):454–61.
- 759 67. Achamrah N, Coëffier M, Déchelotte P. Physical activity in patients with anorexia nervosa. *Nutr Rev*.
760 2016 May 1;74(5):301–11.
- 761 68. Strober M, Freeman R, Morrell W. The long-term course of severe anorexia nervosa in adolescents:
762 Survival analysis of recovery, relapse, and outcome predictors over 10–15 years in a prospective
763 study. *Int J Eat Disord*. 1997;22(4):339–60.
- 764 69. Hübel C, Gaspar HA, Coleman JRI, Hanscombe KB, Purves K, Prokopenko I, et al. Genetic
765 correlations of psychiatric traits with body composition and glycemic traits are sex- and age-
766 dependent. *Nat Commun*. 2019 18;10(1):5765.
- 767 70. Solmi M, Veronese N, Correll CU, Favaro A, Santonastaso P, Caregari L, et al. Bone mineral
768 density, osteoporosis, and fractures among people with eating disorders: a systematic review and
769 meta-analysis. *Acta Psychiatr Scand*. 2016 May;133(5):341–51.
- 770 71. Etain B, Jamain S, Milhiet V, Lajnef M, Boudebese C, Dumaine A, et al. Association between
771 circadian genes, bipolar disorders and chronotypes. *Chronobiol Int*. 2014 Aug;31(7):807–14.
- 772 72. Melo MCA, Abreu RLC, Linhares Neto VB, de Bruin PFC, de Bruin VMS. Chronotype and circadian
773 rhythm in bipolar disorder: A systematic review. *Sleep Med Rev*. 2017 Aug;34:46–58.
- 774 73. Huckins L, Dobbyn A, McFadden W, Wang W, Ruderfer D, Hoffman G, et al. Transcriptomic
775 Imputation of Bipolar Disorder and Bipolar subtypes reveals 29 novel associated genes. *bioRxiv*.
776 2017 Nov;222786.
- 777 74. Meyrer R, Demling J, Kornhuber J, Nowak M. Effects of night shifts in bipolar disorders and extreme
778 morningness. *Bipolar Disord*. 2009 Dec;11(8):897–9.
- 779 75. Alloy LB, Ng TH, Titone MK, Boland EM. Circadian Rhythm Dysregulation in Bipolar Spectrum
780 Disorders. *Curr Psychiatry Rep*. 2017 Apr;19(4):21.
- 781 76. Boivin DB. Influence of sleep-wake and circadian rhythm disturbances in psychiatric disorders. *J*
782 *Psychiatry Neurosci JPN*. 2000 Nov;25(5):446–58.

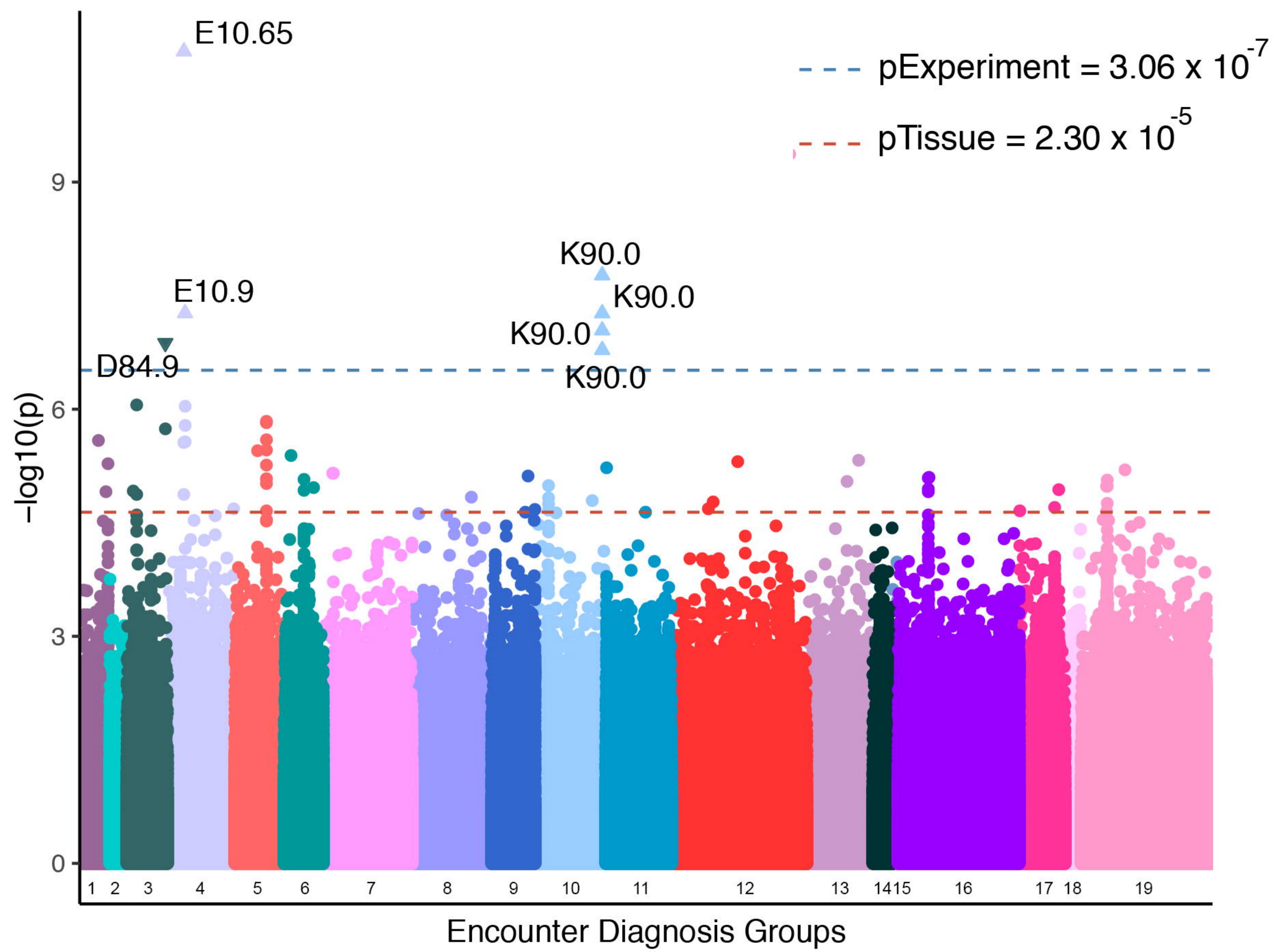
- 783 77. Bunney BG, Bunney WE. Mechanisms of Rapid Antidepressant Effects of Sleep Deprivation
784 Therapy: Clock Genes and Circadian Rhythms. *Biol Psychiatry*. 2013 Jun;73(12):1164–71.
- 785 78. Bunney BG, Li JZ, Walsh DM, Stein R, Vawter MP, Cartagena P, et al. Circadian dysregulation of
786 clock genes: clues to rapid treatments in major depressive disorder. *Mol Psychiatry*. 2015
787 Feb;20(1):48–55.
- 788 79. Karatsoreos IN. Links between Circadian Rhythms and Psychiatric Disease. *Front Behav Neurosci*.
789 2014;8:162.
- 790 80. Kripke DF, Mullaney DJ, Atkinson M, Wolf S. Circadian rhythm disorders in manic-depressives. *Biol*
791 *Psychiatry*. 1978 Jun;13(3):335–51.
- 792 81. Mansour HA, Talkowski ME, Wood J, Chowdari KV, McClain L, Prasad K, et al. Association study of
793 21 circadian genes with bipolar I disorder, schizoaffective disorder, and schizophrenia. *Bipolar*
794 *Disord*. 2009 Nov;11(7):701–10.
- 795 82. Murray G, Allen NB, Trinder J, Burgess H. Is weakened circadian rhythmicity a characteristic of
796 neuroticism? *J Affect Disord*. 2002 Dec;72(3):281–9.
- 797 83. Allison KC, Spaeth A, Hopkins CM. Sleep and Eating Disorders. *Curr Psychiatry Rep*.
798 2016;18(10):92.
- 799 84. Arble DM, Bass J, Laposky AD, Vitaterna MH, Turek FW. Circadian timing of food intake contributes
800 to weight gain. *Obesity*. 2009;17(11):2100–2.
- 801 85. Herpertz S, Albers N, Wagner R, Pelz B, Köpp W, Mann K, et al. Longitudinal changes of circadian
802 leptin, insulin and cortisol plasma levels and their correlation during refeeding in patients with
803 anorexia nervosa. *Eur J Endocrinol*. 2000 Apr;142(4):373–9.
- 804 86. Zucker NL, Bulik CM. On bells, saliva, and abdominal pain or discomfort: Early aversive visceral
805 conditioning and vulnerability for anorexia nervosa. *Int J Eat Disord*. 2020;53(4):508–12.
- 806 87. Adams DM, Reay WR, Geaghan MP, Cairns MJ. Investigation of glycaemic traits in psychiatric
807 disorders using Mendelian randomisation revealed a causal relationship with anorexia nervosa.
808 *Neuropsychopharmacology*. 2020 Sep 13;1–10.
- 809 88. Ilyas A, Hübel C, Stahl D, Stadler M, Ismail K, Breen G, et al. The metabolic underpinning of eating
810 disorders: A systematic review and meta-analysis of insulin sensitivity. *Mol Cell Endocrinol*. 2019
811 Nov 1;497:110307.
- 812 89. Abdulkadir M, Herle M, De Stavola BL, Hübel C, Santos Ferreira DL, Loos RJF, et al. Polygenic
813 Score for Body Mass Index Is Associated with Disordered Eating in a General Population Cohort. *J*
814 *Clin Med* [Internet]. 2020 Apr 21 [cited 2020 Nov 3];9(4). Available from:
815 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7231239/>
- 816 90. Nagata JM, Braudt DB, Domingue BW, Bibbins-Domingo K, Garber AK, Griffiths S, et al. Genetic
817 risk, body mass index, and weight control behaviors: Unlocking the triad. *Int J Eat Disord*.
818 2019;52(7):825–33.
- 819 91. Kolb R, Sutterwala FS, Zhang W. Obesity and cancer: inflammation bridges the two. *Curr Opin*
820 *Pharmacol*. 2016 Aug 1;29:77–89.

- 821 92. Wiklund CA, Kuja-Halkola R, Thornton LM, Hübel C, Leppä V, Bulik CM. Prolonged constipation
822 and diarrhea in childhood and disordered eating in adolescence. *J Psychosom Res.*
823 2019;126:109797.
- 824 93. Riedlinger C, Schmidt G, Weiland A, Stengel A, Giel KE, Zipfel S, et al. Which Symptoms,
825 Complaints and Complications of the Gastrointestinal Tract Occur in Patients With Eating
826 Disorders? A Systematic Review and Quantitative Analysis. *Front Psychiatry.* 2020;11:195.
- 827



medRxiv preprint doi: <https://doi.org/10.1101/2021.02.12.21250941>; this version posted March 13, 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/).



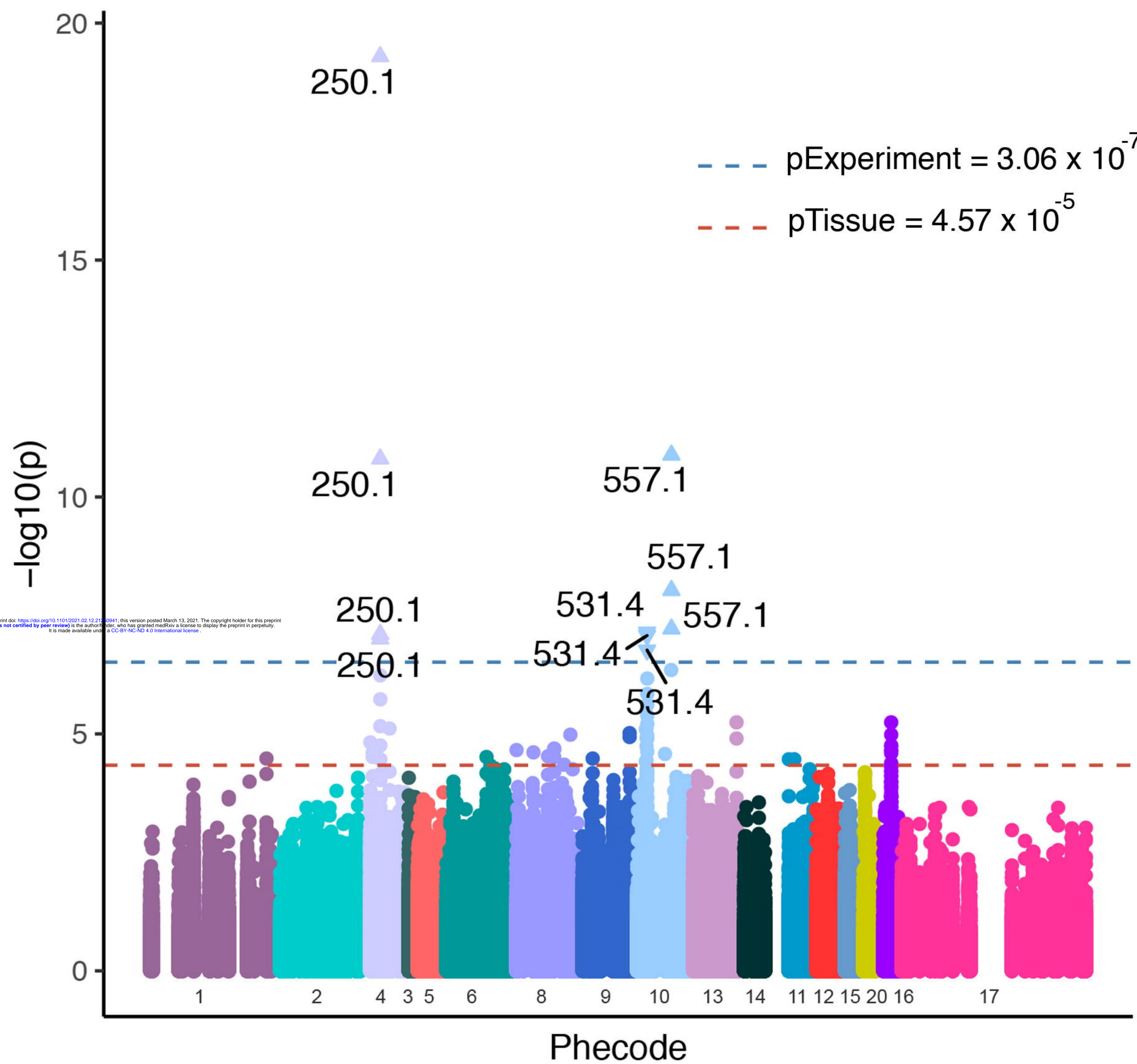


Diagnosis Group

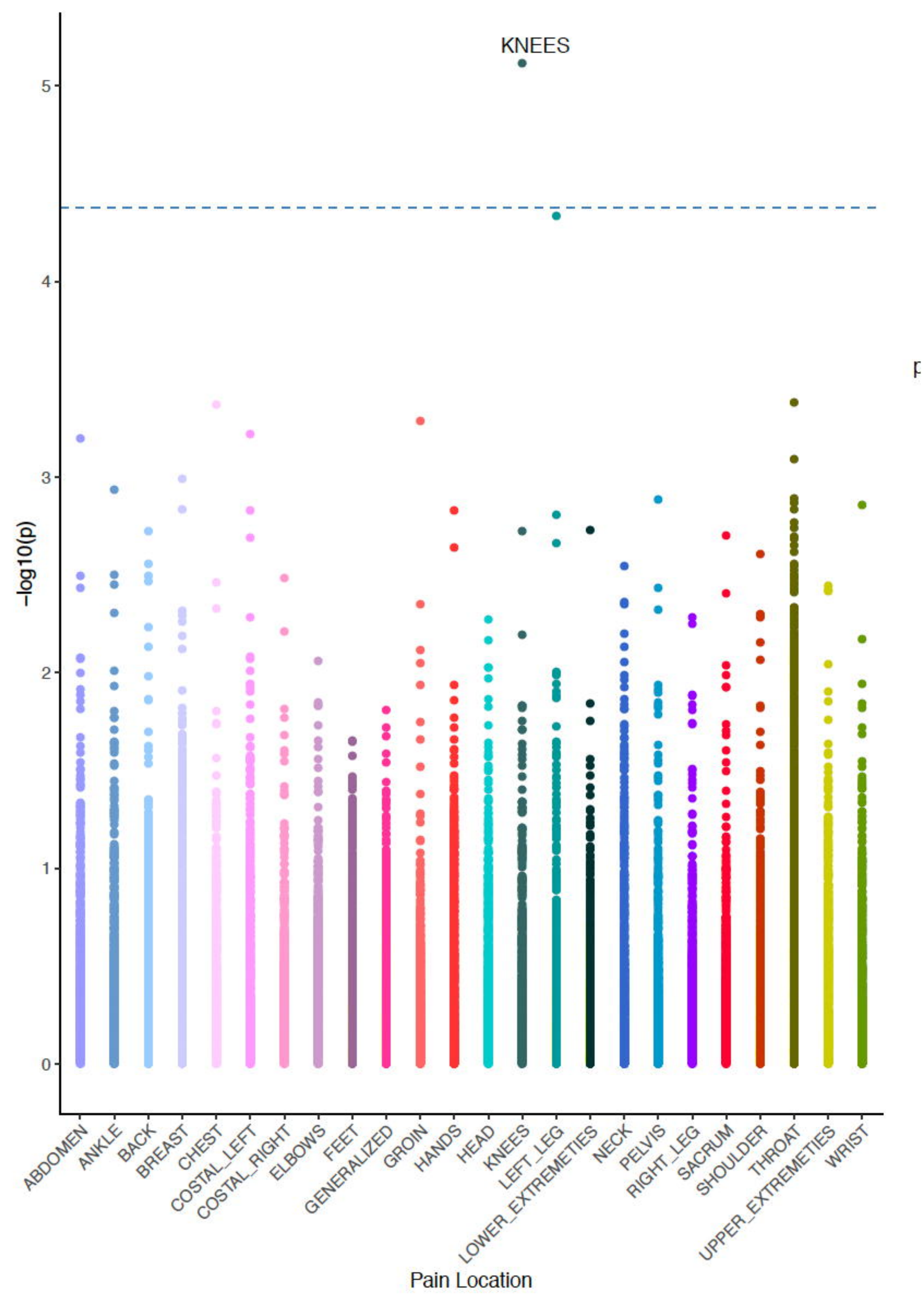
- 1 Infectious and parasitic
- 2 Neoplasms
- 3 Blood disorders
- 4 Endocrine/Nutritional/Metabolic
- 5 Mental/Behavioral/Neurodevelopmental
- 6 Nervous system
- 7 Eye/Ear
- 8 Circulatory system
- 9 Respiratory system
- 10 Digestive system
- 11 Skin/Subcutaneous tissue
- 12 Musculoskeletal system/Connective 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

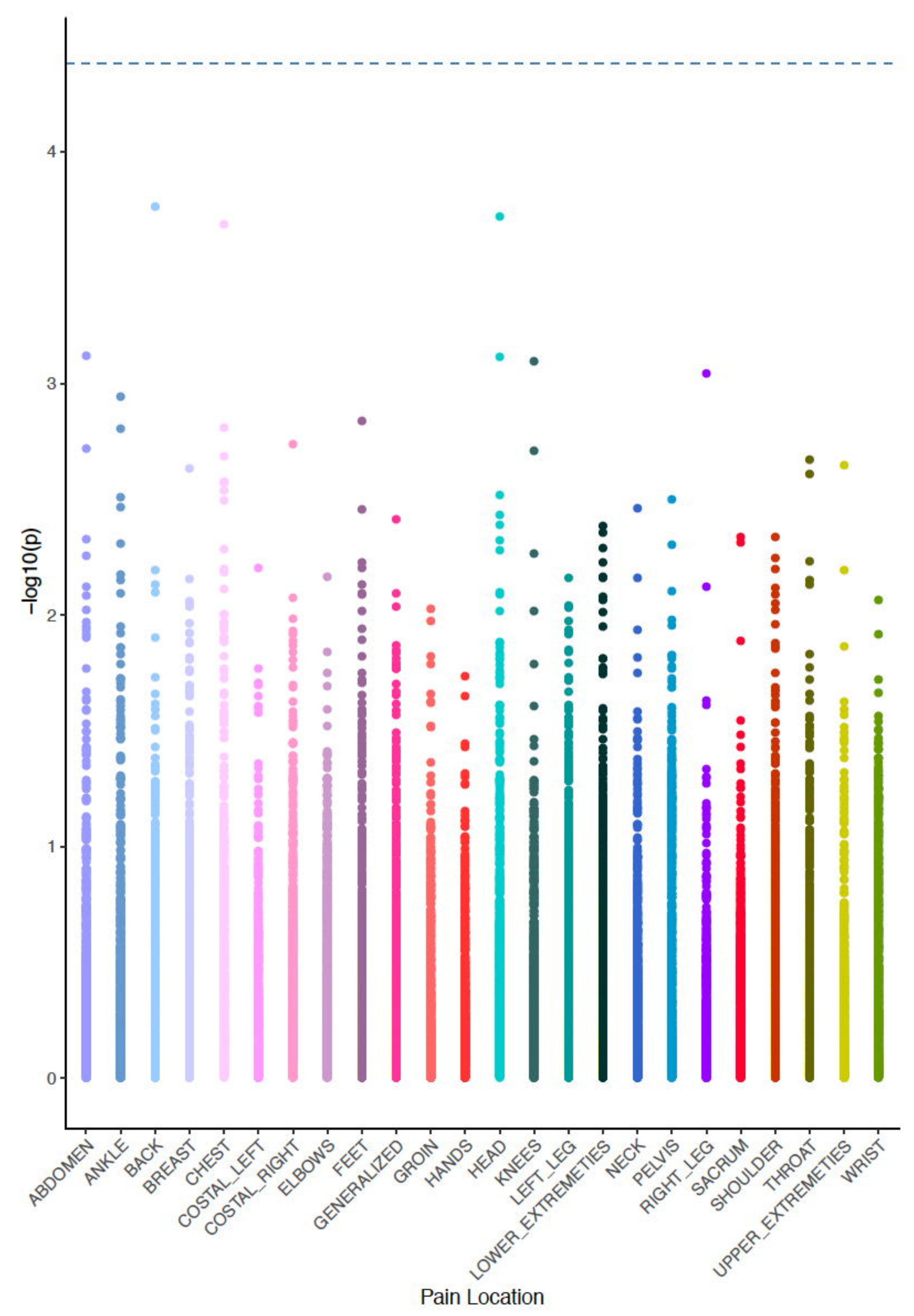
- ▼ DOWN
- ▲ UP



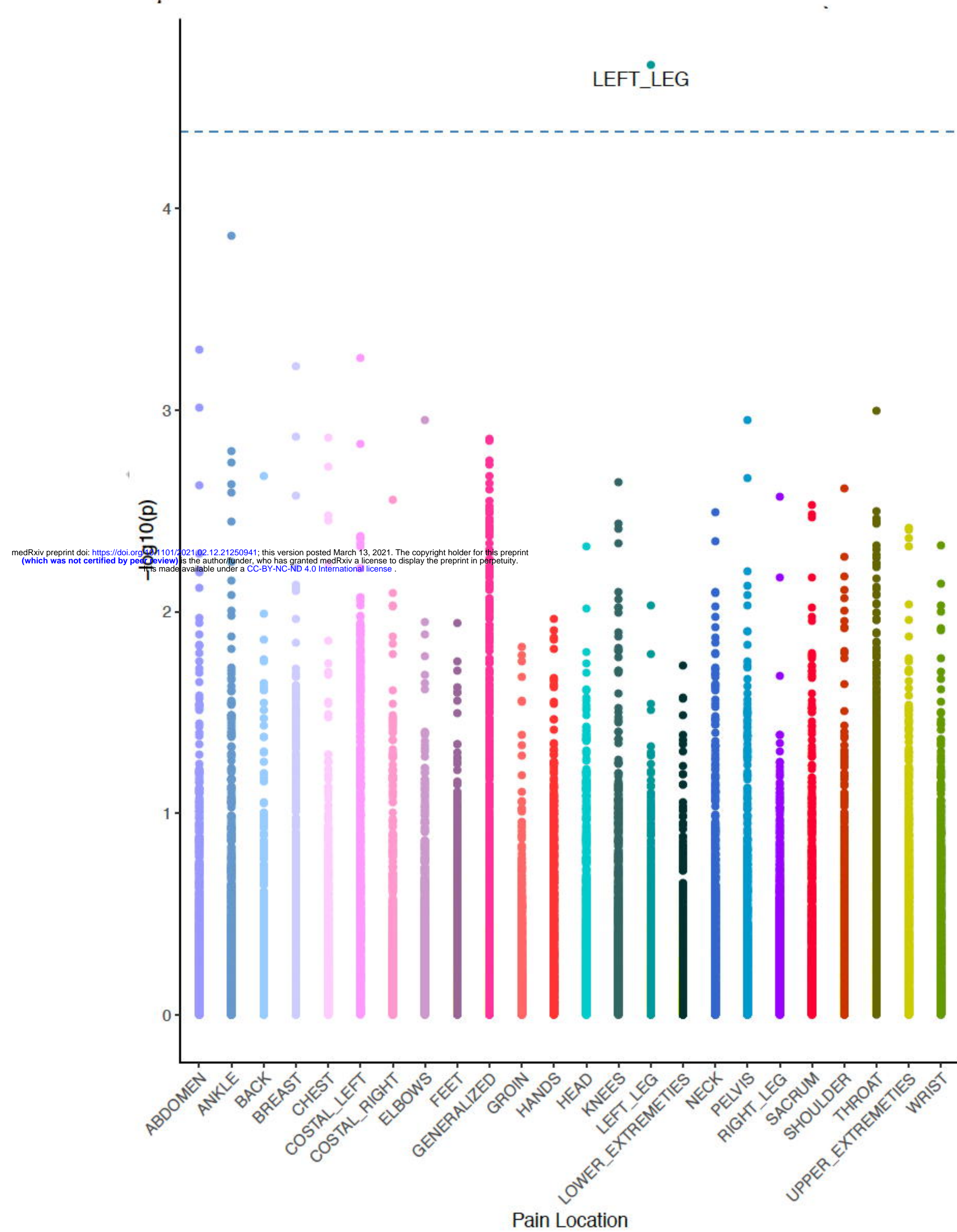
i. Overall



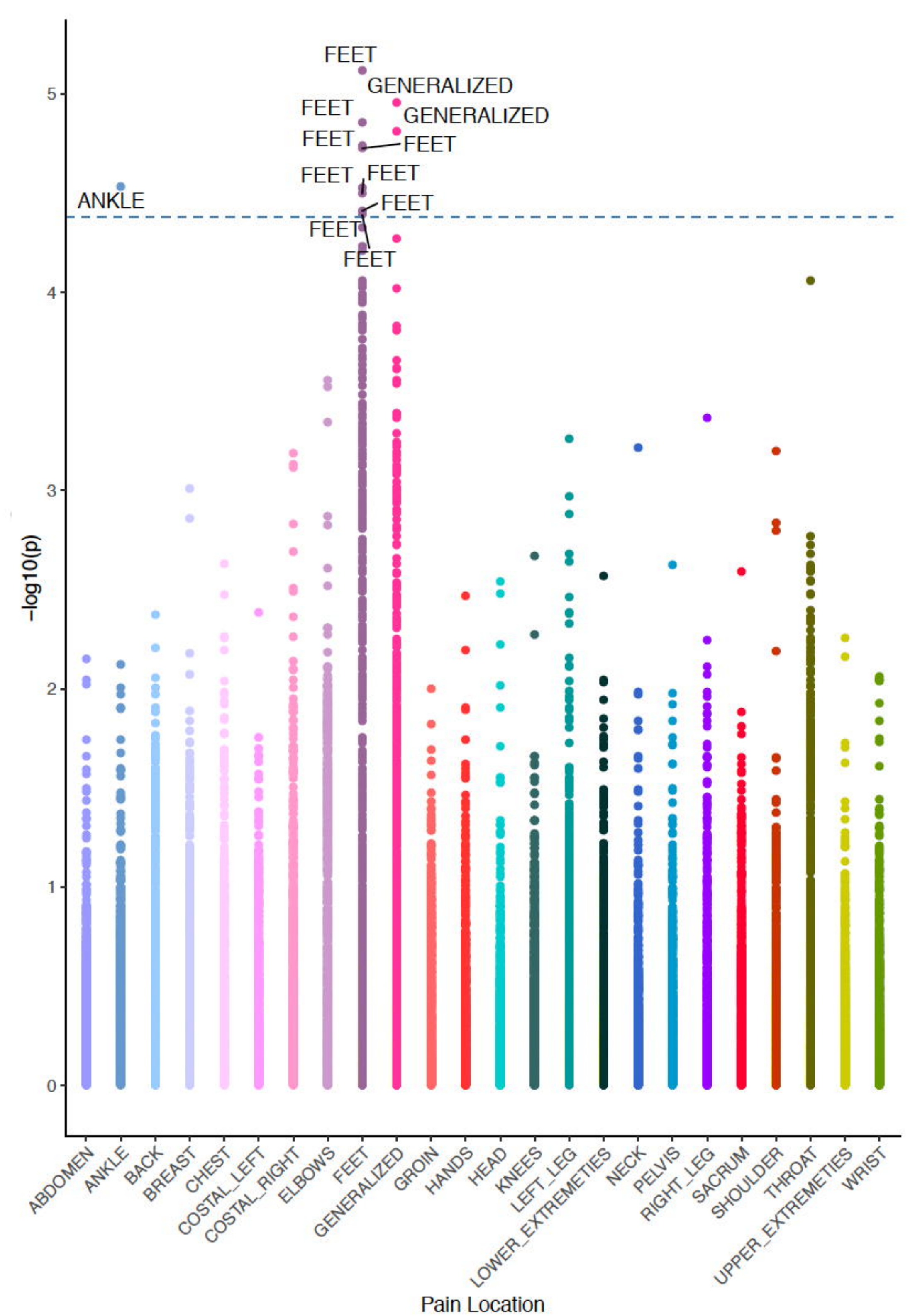
ii. High BMI



iii. Mid BMI



iv. Low BMI



BioMe BioBank Demographic Information

	Total N	N Self-report Female (%)	N Self-report Male (%)	Mean Age (1stQ-3rdQ)	Mean BMI (1stQ-3rdQ)	
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	7.46e-09	<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	2.46e-13	<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/10.1101/2021.02.12.21250441>; this version posted March 13, 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/).