Mapping anorexia nervosa genes to clinical phenotypes

Jessica S. Johnson[1,2] (iD), Alanna C. Cote[1,2] (iD), Amanda Dobbyn[1,2,3] (iD),
Laura G. Sloofman[4] (iD), Jiayi Xu[1,2] (iD), Liam Cotter[1,2] (iD),
Alexander W. Charney[1,2,3,5,6] (iD), Eating Disorders Working Group of the
Psychiatric Genomics Consortium[1,*], Andreas Birgegård[7] (iD), Jennifer Jordan[8] (iD),
Martin Kennedy[8] (iD), Mikaél Landén[7,9] (iD), Sarah L. Maguire[10] (iD), Nicholas
G. Martin[11] (iD), Preben Bo Mortensen[12,13] (iD), Laura M. Thornton[14] (iD),
Cynthia M. Bulik[7,14,15] (iD) and Laura M. Huckins[1,2,3,4,5,6] (iD)

[1]Pamela Sklar Division of Psychiatric Genomics, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [2]Department of Genetics and Genomics, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [3]Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [4]Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [5]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [6]James J. Peters Department of Veterans Affairs Medical Center, Mental Illness Research, Education and Clinical Centers, Bronx, NY 14068, USA; [7]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; [8]Department of Psychological Medicine, Christchurch School of Medicine & Health Sciences, University of Otago, 2 Riccarton Avenue, PO Box 4345, 8140 Christchurch, New Zealand; [9]Institute of Neuroscience and Physiology, Sahlgrenska Academy at Gothenburg University, SE-413 45 Gothenburg, Sweden; [10]InsideOut Institute, University of Sydney, New South Wales 2006, Australia; [11]QIMR Berghofer Medical Research Institute, Locked Bag 2000, Royal Brisbane Hospital, Herston, QLD 4029, Australia; [12]The Lundbeck Foundation Initiative for Integrative Psychiatric Research, Aarhus, Denmark; [13]National Centre for Register-Based Research, Aarhus University, Aarhus, Denmark; [14]Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27517, USA and [15]Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC 27517, USA

Abstract

Background. Anorexia nervosa (AN) is a psychiatric disorder with complex etiology, with a significant portion of disease risk imparted by genetics. Traditional genome-wide association studies (GWAS) produce principal evidence for the association of genetic variants with disease. Transcriptomic imputation (TI) allows for the translation of those variants into regulatory mechanisms, which can then be used to assess the functional outcome of genetically regulated gene expression (GReX) in a broader setting through the use of phenome-wide association studies (pheWASs) in large and diverse clinical biobank populations with electronic health record phenotypes.

Methods. Here, we applied TI using S-PrediXcan to translate the most recent PGC-ED AN GWAS findings into AN-GReX. For significant genes, we imputed AN-GReX in the Mount Sinai BioMe™ Biobank and performed pheWASs on over 2000 outcomes to test the clinical consequences of aberrant expression of these genes. We performed a secondary analysis to assess the impact of body mass index (BMI) and sex on AN-GReX clinical associations.

Results. Our S-PrediXcan analysis identified 53 genes associated with AN, including what is, to our knowledge, the first-genetic association of AN with the major histocompatibility complex. AN-GReX was associated with autoimmune, metabolic, and gastrointestinal diagnoses in our biobank cohort, as well as measures of cholesterol, medications, substance use, and pain. Additionally, our analyses showed moderation of AN-GReX associations with measures of cholesterol and substance use by BMI, and moderation of AN-GReX associations with celiac disease by sex.

Conclusions. Our BMI-stratified results provide potential avenues of functional mechanism for AN-genes to investigate further.

CAMBRIDGE
UNIVERSITY PRESS

Introduction

Anorexia nervosa (AN) is a severe eating disorder with a lifetime prevalence of 0.9–4%, characterized by extreme low body weight, fear of gaining weight, and compensatory weight-loss behaviors such as dietary restrictions, purging, and excessive exercise (Zipfel, Giel, Bulik, Hay, & Schmidt, 2015). Despite having one of the highest mortality rate of any psychiatric disorder and an increased risk of suicide (Bulik, Flatt, Abbaspour, & Carroll, 2019; Chesney, Goodwin, & Fazel, 2014; Mitchell & Peterson, 2020), relatively little is known about the

biological mechanisms underlying AN, and effective, evidence-based treatments are scant, especially for adults (Watson & Bulik, 2013). Twin studies have established AN-heritability between 50% and 60%, indicating a considerable contribution of genetic factors to disease liability (Watson et al., 2019; Yilmaz, Hardaway, & Bulik, 2015).

Genetic studies of AN provide evidence of both psychiatric and metabolic etiology. The largest AN genome-wide association study (GWAS) to date (N_{Cases} = 16 992) uncovered eight AN-associated loci (Watson et al., 2019), and determined single-nucleotide polymorphism (SNP)-based heritability (h^2_{SNP}) of 11–17%, similar to other psychiatric disorders. In addition, genetic correlations demonstrate significant shared genetic variation between AN and other psychiatric disorders, including major depressive disorder (MDD), obsessive-compulsive disorder, anxiety disorders (Watson et al., 2019), schizophrenia (Duncan et al., 2017; Watson et al., 2019), alcohol use disorder (Munn-Chernoff et al., 2021), as well as with physical activity (Hübel et al., 2019a; Watson et al., 2019). Significant negative genetic correlation between AN and anthropometric and metabolic traits [including body mass index (BMI), fat mass, obesity, type-2 diabetes, leptin, and insulin-related traits; Duncan et al., 2017; Hübel et al., 2019a; Watson et al., 2019] have also been observed, further indicating metabolic components to AN disease risk. Studies of AN polygenic risk scores indicate additional genetic associations of AN risk variants with anthropometric, behavioral, and psychiatric traits (Hübel et al., 2021), and with weight trajectories even among healthy adults (Xu et al., 2021).

Although GWASs may provide powerful insights into genetic associations with AN, they cannot pinpoint gene- or tissue-level associations. Therefore, in this study we apply transcriptomic imputation (TI) to identify tissue-specific gene associations with AN. TI approaches leverage gene expression predictor models derived from large, well-curated gene expression datasets [e.g. the Genotype-Tissue Expression project (GTEx) (Lonsdale et al., 2013), CommonMind Consortium (CMC) (Fromer et al., 2016; Huckins et al., 2019), and Depression Genes and Networks (DGN) (Battle et al., 2014)]. These models may be applied to predict genetically regulated gene expression (GReX) in large genotyped cohorts, without the need to collect tissue samples (Huckins et al., 2019).

Observational genetic studies are powerful approaches to detect disease-associated variants, but cannot address subthreshold or prodromal disease states, and do not speak to clinical relevance. Therefore, we leverage a phenome-wide association study (pheWAS) design to probe clinical outcomes with AN-associations. pheWAS effectively query the full electronic health record (EHR) to identify diagnoses and traits associated with a gene or variant (Denny et al., 2010; Pendergrass et al., 2011; Smoller, 2018), and have been used to replicate associations and identify new pleiotropic consequences of GWAS variants, including psychiatric disorders (Denny et al., 2013; Leppert et al., 2020; Zheutlin et al., 2019). EHRs contain vast amount of longitudinal health data, including diagnoses, medications, laboratory tests, vital signs, and family medical history, and, coupled with genetic data through hospital-based biobanks, can provide an understanding of the clinical spectrum of disease and disease progression across the lifetime of the patient. Exploring the associations of AN-genetics with clinical phenotypes has the potential to clarify how some of these GWAS variants functionally contribute to AN disease risk, symptomatology, and clinical presentation.

Here, we explored the associations between AN-GReX and the clinical phenome. We performed TI using S-PrediXcan on the most recent PGC-ED AN GWAS (Watson et al., 2019) to first find GReX associated with AN, and then tested for clinical associations of these genes with structured EHR-encoded phenotypes using pheWAS. We further investigated the effects of BMI and sex on the GReX-phenotype associations through stratification of biobank individuals. Understanding the clinical associations of aberrant gene expression across the phenome may clarify the biological mechanisms of relevant AN GWAS risk variants.

Methods

Transcriptomic imputation

We performed TI using S-PrediXcan (Barbeira et al., 2018) on the largest available summary statistics of AN (16 992 cases/55 525 controls) (Watson et al., 2019). We tested for the association of GReX using available GTEx (Lonsdale et al., 2013), DGN (Battle et al., 2014), and CMC (Fromer et al., 2016; Hoffman et al., 2019) predictor models (Barbeira et al., 2018; Gamazon et al., 2015; Huckins et al., 2019) across 50 tissues with AN case–control status. We established two thresholds for significance; first, Bonferroni correction for all genes tested within each tissue (p_{tissue}, online Supplementary Table S1), and second, correcting for all tissues and genes tested ($p_{Experiment} = 3.75 \times 10^{-8}$). We performed a gene-set analysis for 53 significant AN-genes using FUMA (v1.3.6) GENE2FUNC (Watanabe, Taskesen, van Bochoven, & Posthuma, 2017), including all S-PrediXcan genes (N = 28 454) as the background genes. We defined significant gene sets using $p_{FDR} < 0.05$.

BioMe™

The Mount Sinai BioMe™ Biobank includes genotype and EHR data from 31 704 individuals (online Supplementary Table S2). Individuals were genotyped on the Illumina Global Screening Array; quality control (QC), and imputation of the genotyping data for BioMe™ is described elsewhere (Zheutlin et al., 2019). Ancestry QC of BioMe™ individuals using principal component analysis in PLINK (Chang et al., 2015; Purcell et al., 2007) is described in the online Supplemental Methods. After QC, a total of 31 585 individuals were available for analysis.

pheWAS

We calculated GReX for all S-PrediXcan p_{tissue} significant genes (N = 53) across 45 tissues from GTEx, DGN, and CMC in BioMe™ using PrediXcan-2 (Barbeira et al., 2018; Gamazon et al., 2015). We excluded sex-specific tissues (ovary, uterus, vagina, prostate, and testis) from this analysis, and genes with GReX variance (gVAR) less than 0.002 (online Supplementary Table S3C).

Logistic regressions between AN-GReX and categorical phenotypes were performed using the pheWAS R package (Carroll, Bastarache, & Denny, 2014), adjusting for age, sex, and the first five genotype-derived principal components. pheWAS were run individually for each BioMe™ ancestry cohort, and results were meta-analyzed using an inverse-variance based approach in METAL (Willer, Li, & Abecasis, 2010). We required at least 10 occurrences of each phenotype in each population group for

inclusion in the analysis, and an overall effective sample size N_{eff} > 100 (Equation (1)). We excluded associations with significant Cochran's Q-test for heterogeneity scores across cohorts (pHet > 0.001).

Effective sample size (N_{eff}):

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

'Encounter Diagnoses' were recorded at each patient visit using the International Classification of Disease (ICD) coding system. Phecodes were assigned from Encounter Diagnoses by grouping ICD-9 and ICD-10 diagnostic codes (Wu et al., 2019). We defined cases as individuals with at least two counts of a code. Those with zero counts were considered 'controls', and those with only one count were set to missing. After QC, our dataset included 2178 unique Encounter Diagnosis codes and 1093 unique phecodes. Due to the high correlation between Encounter Diagnosis and phecode files, we combined all results from both pheWAS and performed an FDR correction in R to determine significance ($p_{FDR} < 0.05$).

In addition to diagnostic codes, Bio*Me*™ EHR data include allergies, vital signs (weight, height, blood pressure, pulse, pulse oximetry, respirations, and temperature), lab results, family history, personal history, medication use, obstetrics and gynecology outcomes (OB/GYN), and other social and behavioral traits. Summaries, descriptive statistics, and QC descriptions for these phenotypes can be found in online Supplementary Methods. We tested for AN-GReX-pheWAS associations within each category, establishing significance using a Bonferroni correction (Equation (2)).

Bonferroni-correction for EHR data at tissue- and experiment-significance thresholds:

$$P_{Tissue} = \frac{0.05}{N_{Phenotypes}}$$

$$P_{Experiment} = \frac{0.05}{N_{Phenotypes} \times N_{Tissues}} \quad (2)$$

Stratification by BMI

Given the nature of AN symptomology, we tested whether our significant AN-GReX associations varied based on BMI. We stratified Bio*Me*™ individuals into three BMI categories: low (<first BMI quartile, $N = 6976$), mid (first—third BMI quartiles, $N = 13\,972$) and high (>third BMI quartile, $N = 9350$; online Supplementary Fig. S1A). Quartiles were defined based on ancestry- and sex-specific BMI distributions within our data (online Supplementary Fig. S1B; Table S2; Supplemental Methods). We repeated our pheWAS within these categories, for all AN-GReX associations reaching significance in our overall analysis. BMI-stratified pheWAS were adjusted for age, sex, and the first five genotype-derived PCs.

Stratification by sex

To determine whether our associations varied based on sex, we stratified the Bio*Me*™ cohort by females ($N = 17\,968$) and males ($N = 12\,417$), with sex determined by genotype, and repeated our AN-GReX pheWAS analyses within each group. Here, we included sex-specific tissues (ovary, uterus, vagina, prostate, and testis) for the respective sex. Sex-stratified pheWAS were adjusted for age and the first five genotype-derived PCs.

Testing for hidden case contamination

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

Full derivations and simulations are shown in our online Supplementary Material; briefly, we derive the expression among controls, and cases contaminated at rate p; the difference in expression between the two groups; the expected variance among cases, and pooled across all samples, and the expected statistical significance (T-score, and p value). Next, we simulated gene expression distributions for (i) 1000 cases and 1000 controls; (ii) 1000 cases and 10 000 controls; (iii) 1000 cases and 30 000 controls; 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, 1, 2, 3, 4, 5, 10) values. We repeated our simulations 10 000 times at each case–control proportion- p-β combination, and demonstrated that our formulae accurately estimate the desired values (online Supplementary Fig. S2).

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

(1) Diagnostic contamination occurs among our cases only, assuming the normal population rate for AN: 0.9% among women; 0.3% among men.
(2) Contamination occurs at significantly higher levels than might be expected in the population (0.6, 1.2, 3, 6%), and all of these samples fall into our case category.

For each scenario, we repeated our calculation using 500 and 1000 cases, matched with 1000, 5000, and 30 000 controls.

Results

S-PrediXcan identifies gene–tissue associations with AN

We applied S-PrediXcan to PGC-AN GWAS summary statistics, and identified 218 gene–tissue associations, including 53 unique genes across 12 loci (Fig. 1). Two loci on chromosomes 2 and 3 reached experiment-wide significance ($p < 3.75 \times 10^{-8}$), while 10 loci reached tissue-specific significance (Fig. 2, Table 1; online Supplementary Table S3A). Our S-PrediXcan AN-genes are significantly enriched in gene sets for traits including inflammatory bowel disease (IBD: Crohn's and ulcerative colitis) ($p_{adj} = 1.3 \times 10^{-25}$), sleep duration ($p_{adj} = 6.9 \times 10^{-17}$), blood protein levels ($p_{adj} = 3.2 \times 10^{-16}$), intelligence ($p_{adj} < 1.5 \times 10^{-6}$), regular attendance at a religious group, gym, or sports club ($p_{adj} < 1.46 \times 10^{-10}$), mood instability ($p_{adj} = 3.3 \times 10^{-9}$), AN ($p_{adj} = 7.4 \times 10^{-8}$), and body fat distribution ($P_{adj} < 5.0 \times 10^{-7}$) (online Supplementary Table S4A and B).

AN-GReX is associated with autoimmune and autoinflammatory diseases

We performed a pheWAS to identify clinical associations with dysregulation of our AN-genes. We identified 16
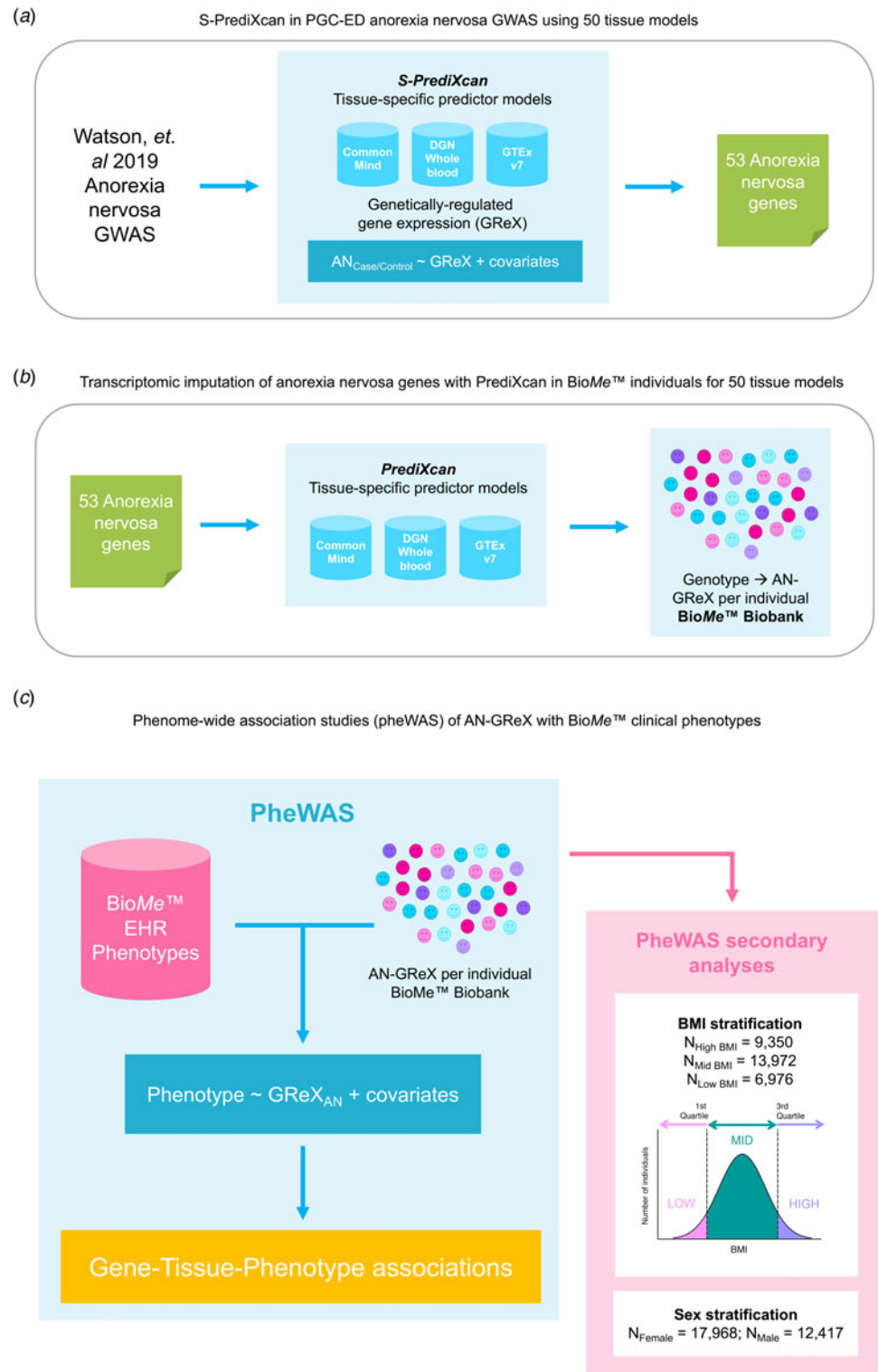
Fig. 1. Graphical depiction of S-PrediXcan and PrediXcan TI and pheWAS analyses. (*a*) We used S-PrediXcan predictor models for 50 different tissue types to impute GReX in Watson et al. (2019) AN GWAS and found 53 genes whose GReX was associated with AN. We then (*b*) imputed GReX for those 53 AN-genes in our BioMe™ cohort and (*c*) performed a pheWAS across available EHR phenotypes. The pheWAS analyses were run within each ancestral population and then meta-analyzed using an inverse-variance approach in METAL. Secondary analyses included stratifying individuals in BioMe™ by BMI and sex, and running the pheWAS analyses within each stratification group.

FDR-significant gene–tissue associations with four Phecode- and Encounter Diagnosis phenotypes (FDR < 0.05): type 1 diabetes (N_{Cases} = 408), celiac disease (N_{Cases} = 63), peptic ulcer (N_{Cases} = 254), and unspecified immunodeficiency (N_{Cases} = 55) (Fig. 3; online Supplementary Table S5). These associations are driven in part by the MHC-gene *CLIC1*; predicted upregulation of *CLIC1* in spleen was associated with celiac disease and type 1 diabetes ($p = 1.30 \times 10^{-11}$, $p = 5.08 \times 10^{-20}$, respectively) in the overall cohort. We did not observe any significant effect of patient BMI group on these associations (online Supplementary Fig. S3A and C). Upregulation of *CLIC1* in females was significantly associated with type 1 diabetes (N_{Cases} = 230) and celiac disease (N_{Cases} = 53) (Females-Spleen-*CLIC1*, $p < 1.49 \times 10^{-9}$) (online Supplementary Fig. S3B and D) .
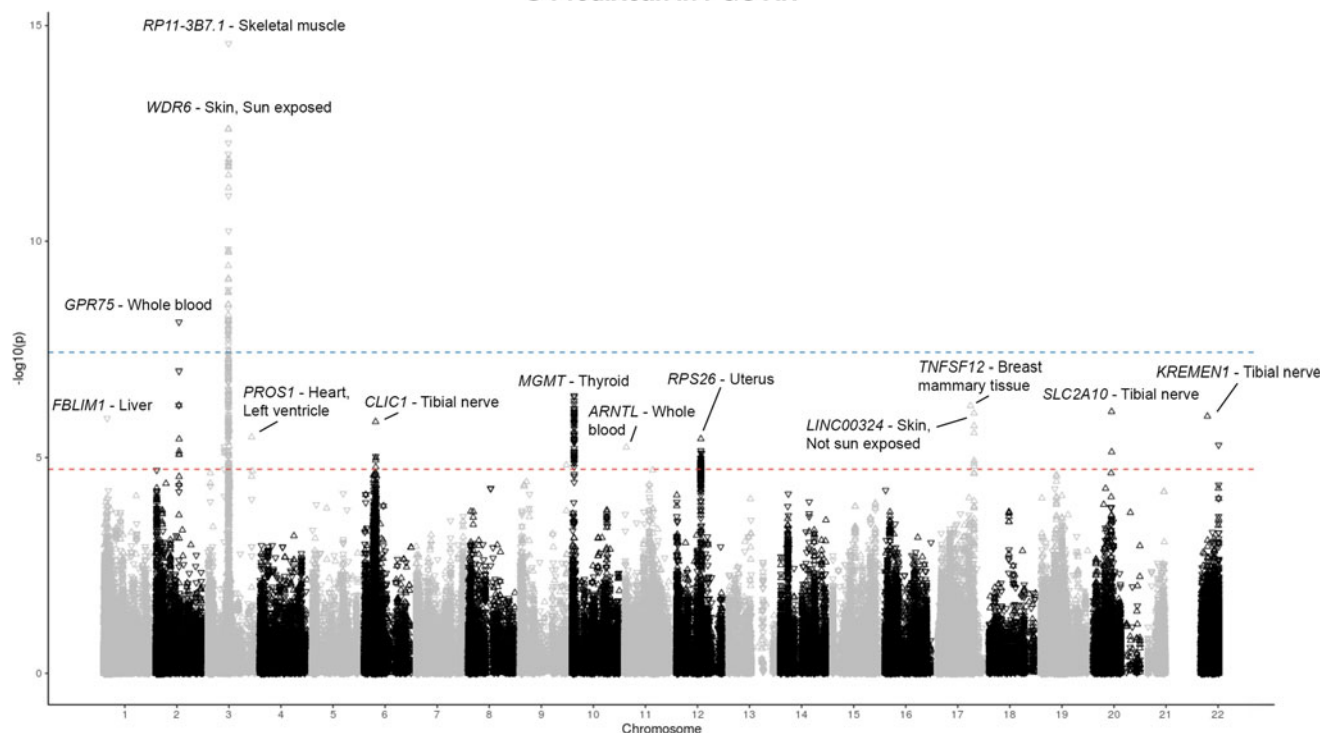
Fig. 2. S-PrediXcan results for PGC-ED AN GWAS (2019) (N_{Cases} = 16 992, $N_{Controls}$ = 55 525). S-PrediXcan TI of the PGC-ED AN GWAS summary statistics to determine GReX, with tests for association of GReX with AN disease status. Manhattan plot of S-PrediXcan gene–tissue associations with AN for 50 tissues. Each point represents a different gene–tissue association result; i.e. the same gene may have multiple points within a peak. Experiment-wide significance threshold of $p < 3.75 \times 10^{-8}$ (solid line); tissue-specific significance threshold of $p < 2.45 \times 10^{-5}$ (dotted line).

Table 1. S-PrediXcan loci results

Chromosome	Genes	Z score	p value	Top finding (gene)	Top finding (tissue)
1	FBLIM1	−4.851060	1.23×10^{-6}	FBLIM1	Liver
2	ASB3, CHAC2, GPR75	−5.780228	**7.46×10^{-9}**	GPR75	Whole Blood
3	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	7.321127	**2.46×10^{-13}**	WDR6	Skin, Sun Exposed
3	PROS1, DHFRL1	4.687382	2.77×10^{-6}	DHFRL1	DGN Whole Blood
6	CLIC1	4.814629	1.47×10^{-6}	CLIC1	Tibial Nerve
10	MGMT, EBF3	−5.086391	3.65×10^{-7}	MGMT	Thyroid
11	ARNTL	4.534957	5.76×10^{-6}	ARNTL	Whole Blood
12	SUOX, RPS26, SLC26A10	4.628307	3.69×10^{-6}	RPS26	Uterus
17	TNFSF12, LINC00324	4.982641	6.27×10^{-7}	TNFSF12	Breast Mammary Tissue
20	SLC2A10	4.921176	8.60×10^{-7}	SLC2A10	Tibial Nerve
22	KREMEN1	4.872670	1.10×10^{-6}	KREMEN1	Tibial Nerve
22	SCO2	−4.560015	5.11×10^{-6}	SCO2	Brain, Cerebelium

Summary of S-PrediXcan of PGC-ED AN GWAS results. Genes associated with AN from each locus are shown, as well as the top finding Z score, p value, gene and tissue for each locus. p values shaded in bold indicate experiment-wide significant loci ($p < 3.75 \times 10^{-8}$).

BMI moderates the effect of MGMT-GReX on cholesterol levels

Next, we looked at the association of AN-GReX and measures of highest recorded, lowest recorded and mean total cholesterol (mg/dl), high-density lipoprotein (HDL) cholesterol (mg/dl), and low-density lipoprotein (LDL) cholesterol (mg/dl) (Table 2). In the full cohort, *RPS26* and *SUOX* GReX were significantly associated with the highest HDL measures, and *MGMT* GReX
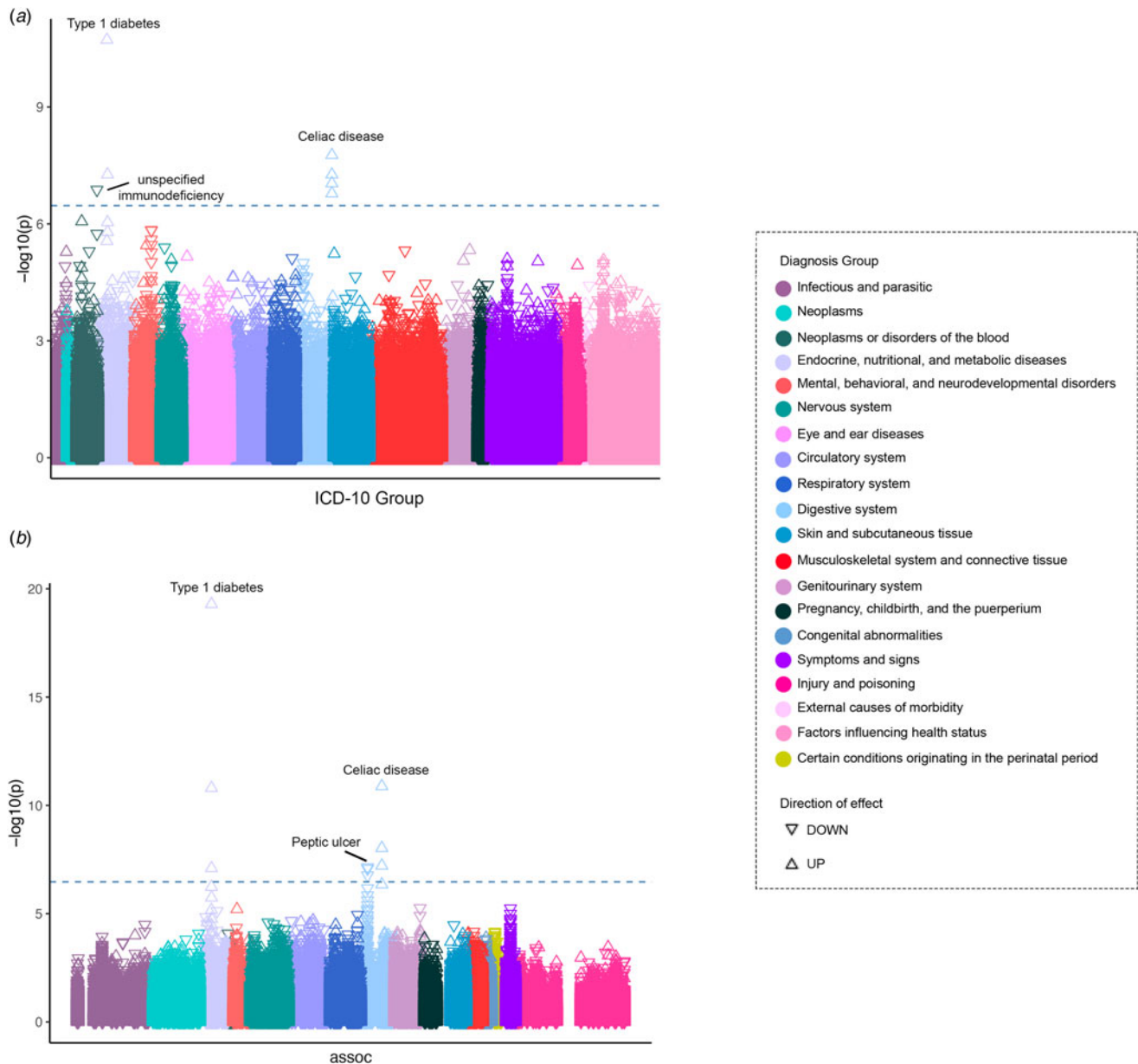
Fig. 3. AN-GReX associations with Bio*Me*™ Diagnosis codes. GReX-Tissue-Phenotype associations for (*a*) encounter diagnosis ICD-10 codes (*N* = 2178) and (*b*) phe-codes (*N* = 1093) for the Bio*Me*™ cohort (*N* = 30 585). Diagnosis codes are plotted along the *x*-axis and grouped by category, with the −log(10) *p* value associations along the *y*-axis. FDR-significant diagnosis codes are labeled (FDR-adjusted *p* < 0.05). FDR-significant *p* value threshold *p* = 3.4 × 10^{-7} (blue dashed line).

was associated with the highest cholesterol at within-tissue significance (*N* = 23 357) (*p* < 5.1 × 10^{-4}) (Fig. 4C). Associations of AN-GReX with lipid measures varied across BMI-stratified groups; predicted downregulation of *MGMT* in the stomach, liver, esophagus, and cells was associated with the highest total cholesterol and LDL levels, and in dorsolateral prefrontal cortex and hippocampus with mean cholesterol and LDL among high-BMI individuals (*N*$_{High\ BMI}$ = 6910) (High-BMI-Stomach-*MGMT*, *p* = 1.55 × 10^{-7}) (Fig. 4C; online Supplementary Table S6).

AN-GReX is associated with tobacco use

Multiple genes were associated with categorical and continuous measures of tobacco use in the overall cohort at within-tissue

significance (online Supplementary Fig. S4, Table S7, Supplementary Results). We found moderating effects of patient BMI on tobacco use: among High-BMI groups, different genes were associated with the number of years of tobacco use at experiment-wide significance (*N*$_{High\ BMI}$ = 8129) (High-BMI-Pituitary-*CCDC36*, *p* = 4.2 × 10^{-6}; High-BMI-Colon, Transverse-*P4HTM*, *p* = 3.4 × 10^{-5}; High-BMI-Stomach-*P4HTM*, *p* = 2.2 × 10^{-5}) (online Supplementary Fig. S4, Table S7).

AN-GReX is associated with measures of pain score and pain location

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

Table 2. pheWAS AN-GReX tissue associations with cholesterol phenotypes

Stratification category	Gene	Tissue	Effect	s.e.	p
(A) AN-GReX pheWAS associations with highest recorded total cholesterol (mg/dl)					
High BMI	*MGMT*	**Stomach**	**−7.2**	**1.4**	**1.6×10^{-7}**
High BMI	*MGMT*	**Brain, nucleus accumbens basal ganglia**	**−11.5**	**2.3**	**4.0×10^{-7}**
High BMI	*MGMT*	**Colon, transverse**	**−6.1**	**1.3**	**1.2×10^{-6}**
High BMI	*MGMT*	**Brain, caudate basal ganglia**	**−7.2**	**1.5**	**2.1×10^{-6}**
High BMI	*MGMT*	**Cells, EBV-transformed lymphocytes**	**−8.0**	**1.7**	**2.8×10^{-6}**
High BMI	*MGMT*	**Thyroid**	**−5.6**	**1.2**	**2.9×10^{-6}**
High BMI	*MGMT*	**Visceral omentum adipose**	**−5.7**	**1.2**	**2.9×10^{-6}**
High BMI	*MGMT*	**CommonMind DLPFC**	**−13.6**	**2.9**	**3.7×10^{-6}**
High BMI	*MGMT*	**Lung**	**−5.9**	**1.3**	**6.3×10^{-6}**
High BMI	*MGMT*	**Esophagus, mucosa**	**−5.6**	**1.3**	**9.3×10^{-6}**
High BMI	*MGMT*	**Brain, hippocampus**	**−7.6**	**1.7**	**1.0×10^{-5}**
High BMI	*MGMT*	**Brain, cerebellum**	**−7.8**	**1.8**	**1.9×10^{-5}**
High BMI	*MGMT*	**Spleen**	**−3.7**	**0.9**	**2.7×10^{-5}**
High BMI	*MGMT*	**Tibial nerve**	**−5.2**	**1.3**	**3.7×10^{-5}**
High BMI	*MGMT*	Liver	−5.5	1.4	1.3×10^{-4}
High BMI	*MGMT*	Brain, cerebellar hemisphere	−4.2	1.2	2.9×10^{-4}
Overall	*MGMT*	Stomach	−2.7	0.7	2.9×10^{-4}
(B) AN-GReX pheWAS associations with highest recorded LDL cholesterol (mg/dl)					
High BMI	*MGMT*	**Stomach**	**−5.2**	**1.1**	**2.6×10^{-6}**
High BMI	*MGMT*	**Brain, nucleus accumbens basal ganglia**	**−7.9**	**1.8**	**1.5×10^{-5}**
High BMI	*MGMT*	**Brain, hippocampus**	**−5.8**	**1.4**	**3.1×10^{-5}**
High BMI	*MGMT*	**Colon, transverse**	**−4.1**	**1.0**	**4.0×10^{-5}**
High BMI	*MGMT*	**Cells, EBV-transformed lymphocytes**	**−5.7**	**1.4**	**4.2×10^{-5}**
High BMI	*MGMT*	**Brain, caudate basal ganglia**	**−4.9**	**1.2**	**7.4×10^{-5}**
High BMI	*MGMT*	**Thyroid**	**−3.8**	**1.0**	**7.4×10^{-5}**
High BMI	*MGMT*	**Visceral omentum adipose**	**−3.8**	**1.0**	**9.4×10^{-5}**
High BMI	*MGMT*	**CommonMind DLPFC**	**−9.1**	**2.4**	**1.2×10^{-4}**
High BMI	*MGMT*	**Esophagus, mucosa**	**−3.9**	**1.0**	**1.2×10^{-4}**
High BMI	*MGMT*	Lung	−3.8	1.1	3.5×10^{-4}
High BMI	*MGMT*	Brain, cerebellum	−5.2	1.5	4.1×10^{-4}
High BMI	*MGMT*	Tibial nerve	−3.5	1.0	6.3×10^{-4}
High BMI	*MGMT*	Liver	−3.8	1.1	8.2×10^{-4}
Mid BMI	*SLC26A10*	Aorta artery	−4.7	1.4	8.8×10^{-4}
(C) AN-GReX pheWAS associations with highest recorded HDL cholesterol (mg/dl)					
Overall	*SUOX*	Thyroid	0.9	0.3	1.7×10^{-4}
Overall	*SUOX*	Brain, caudate basal ganglia	1.4	0.4	1.9×10^{-4}
Overall	*RPS26*	Breast mammary tissue	−0.7	0.2	1.9×10^{-4}
Overall	*SUOX*	Tibial artery	1.4	0.4	2.0×10^{-4}
Overall	*RPS26*	Brain, nucleus accumbens basal ganglia	−0.7	0.2	3.1×10^{-4}
Overall	*SUOX*	Tibial nerve	0.8	0.2	3.5×10^{-4}
Overall	*SUOX*	Pituitary	1.0	0.3	4.6×10^{-4}
Female	*SUOX*	Brain, caudate basal ganglia	1.8	0.5	4.7×10^{-4}

(Continued)

Table 2. (*Continued.*)

Stratification category	Gene	Tissue	Effect	S.E.	*p*
Overall	*SUOX*	Brain, anterior cingulate cortex (BA24)	2.0	0.6	5.1×10^{-4}
Overall	*RPS26*	CommonMind DLPFC	−1.2	0.3	5.2×10^{-4}
Overall	*RPS26*	Skeletal muscle	−0.6	0.2	5.4×10^{-4}
Overall	*SUOX*	Esophagus, mucosa	4.2	1.2	5.5×10^{-4}
Overall	*SUOX*	Skin, not sun exposed	1.2	0.3	6.1×10^{-4}
Overall	*RPS26*	Brain, putamen basal ganglia	−0.7	0.2	6.3×10^{-4}
Overall	*RPS26*	Minor salivary gland	−1.0	0.3	6.5×10^{-4}
Overall	*SUOX*	Small intestine, terminal ileum	1.3	0.4	6.8×10^{-4}
Overall	*SUOX*	Esophagus, muscularis	1.4	0.4	7.6×10^{-4}
Overall	*RPS26*	Tibial nerve	−0.6	0.2	7.6×10^{-4}
Overall	*SUOX*	Lung	1.0	0.3	7.7×10^{-4}
Overall	*SUOX*	Pancreas	1.0	0.3	7.9×10^{-4}
Overall	*RPS26*	Brain, substantia nigra	−0.9	0.3	8.0×10^{-4}
Overall	*RPS26*	Brain, anterior cingulate cortex (BA24)	−0.6	0.2	8.1×10^{-4}
Overall	*RPS26*	Brain, hypothalamus	−0.7	0.2	8.1×10^{-4}
Overall	*SUOX*	Heart, left ventricle	2.2	0.7	8.1×10^{-4}
Overall	*RPS26*	Aorta artery	−0.7	0.2	8.2×10^{-4}
Overall	*RPS26*	Brain, spinal cord (cervical c-1)	−0.6	0.2	8.5×10^{-4}
Overall	*SUOX*	Spleen	0.9	0.3	8.8×10^{-4}
Overall	*SUOX*	Breast mammary tissue	3.7	1.1	9.2×10^{-4}
Overall	*RPS26*	Whole blood	−0.7	0.2	9.5×10^{-4}
Overall	*RPS26*	Lung	−0.6	0.2	9.9×10^{-4}
(D) AN-GReX pheWAS associations with mean total cholesterol (mg/dl)					
High BMI	**MGMT**	**CommonMind DLPFC**	**−9.5**	**2.1**	$\mathbf{4.8 \times 10^{-6}}$
High BMI	**MGMT**	**Brain, spinal cord (cervical c-1)**	**−4.9**	**1.1**	$\mathbf{6.6 \times 10^{-6}}$
High BMI	**MGMT**	**Brain, caudate basal ganglia**	**−4.8**	**1.1**	$\mathbf{7.9 \times 10^{-6}}$
High BMI	**MGMT**	**Brain, hippocampus**	**−5.2**	**1.2**	$\mathbf{1.8 \times 10^{-5}}$
High BMI	**MGMT**	**Lung**	**−3.9**	**0.9**	$\mathbf{2.2 \times 10^{-5}}$
High BMI	**MGMT**	**Esophagus, mucosa**	**−3.7**	**0.9**	$\mathbf{2.6 \times 10^{-5}}$
High BMI	**MGMT**	**Brain, cerebellum**	**−5.1**	**1.3**	$\mathbf{6.8 \times 10^{-5}}$
High BMI	**MGMT**	**Liver**	**−3.9**	**1.0**	$\mathbf{9.1 \times 10^{-5}}$
High BMI	*MGMT*	Spleen	−2.4	0.6	1.4×10^{-4}
High BMI	*APEH*	Skin, sun exposed	8.9	2.6	5.8×10^{-4}
(E) AN-GReX pheWAS associations with mean LDL cholesterol (mg/dl)					
High BMI	**MGMT**	**Brain, hippocampus**	**−4.1**	**1.0**	$\mathbf{2.6 \times 10^{-5}}$
High BMI	**MGMT**	**CommonMind DLPFC**	**−7.1**	**1.7**	$\mathbf{4.1 \times 10^{-5}}$
High BMI	**MGMT**	**Brain, spinal cord (cervical c-1)**	**−3.7**	**0.9**	$\mathbf{4.2 \times 10^{-5}}$
High BMI	**MGMT**	**Esophagus, mucosa**	**−2.9**	**0.7**	$\mathbf{1.0 \times 10^{-4}}$
High BMI	**MGMT**	**Brain, caudate basal ganglia**	**−3.4**	**0.9**	$\mathbf{1.2 \times 10^{-4}}$
High BMI	*MGMT*	Brain, cerebellum	−3.9	1.1	2.8×10^{-4}
High BMI	*MGMT*	Lung	−2.8	0.8	2.8×10^{-4}
High BMI	*MGMT*	Liver	−3.0	0.8	3.1×10^{-4}
Mid BMI	*SLC26A10*	Aorta artery	−3.4	1.0	7.8×10^{-4}

(*Continued*)

Table 2. (Continued.)

Stratification category	Gene	Tissue	Effect	S.E.	p
(F) AN-GReX pheWAS associations with mean HDL cholesterol (mg/dl)					
Mid BMI	RP11-3B7.1	Lung	−2.1	0.6	4.2×10^{-4}
Male	RP11-3B7.1	Lung	−2.0	0.6	4.9×10^{-4}
Overall	RP11-3B7.1	Lung	−1.5	0.4	6.2×10^{-4}
Male	GPX1	Brain, caudate basal ganglia	1.2	0.3	8.3×10^{-4}
Overall	RPS26	Brain, nucleus accumbens basal ganglia	−0.5	0.1	9.4×10^{-4}
Overall	RPS26	Breast mammary tissue	−0.5	0.1	9.4×10^{-4}
(G) AN-GReX pheWAS associations with lowest total cholesterol (mg/dl)					
Female	PFKFB4	Skin, sun exposed	5.7	1.7	7.4×10^{-4}
High BMI	MGMT	CommonMind DLPFC	−7.3	2.2	7.5×10^{-4}
High BMI	MST1	Stomach	5.0	1.5	8.7×10^{-4}
(H) AN-GReX pheWAS associations with lowest LDL cholesterol (mg/dl)					
Mid BMI	RBM5	Pituitary	8.3	2.5	7.2×10^{-4}
Overall	MST1	Liver	1.4	0.4	8.6×10^{-4}
(I) AN-GReX pheWAS associations with lowest HDL cholesterol (mg/dl)					
High BMI	RP11-477N3.1	Lung	4.2	1.0	3.4×10^{-5}
Female	SEMA3F	Spleen	3.3	0.9	3.9×10^{-4}
High BMI	RP11-477N3.1	Skeletal muscle	2.4	0.7	3.9×10^{-4}
Mid BMI	RP11-3B7.1	Lung	−2.0	0.6	4.5×10^{-4}
Overall	SEMA3F	Spleen	2.3	0.7	7.0×10^{-4}
Overall	RP11-3B7.1	Lung	−1.4	0.4	8.0×10^{-4}
High BMI	RP11-477N3.1	Tibial artery	2.0	0.6	9.1×10^{-4}
Mid BMI	CCDC36	Lung	−1.7	0.5	9.2×10^{-4}

pheWAS associations with highest, lowest, and mean total cholesterol, LDL cholesterol and HDL cholesterol measures for all stratification groups: overall ($N = 23\,661$), high-BMI ($N = 6910$), mid-BMI ($N = 11\,243$), low-BMI ($N = 5324$), females ($N = 13\,907$), and males ($N = 9669$). AN-GReX association with phenotypes of (A) highest recorded total cholesterol (mg/dl), (B) highest recorded LDL cholesterol (mg/dl), (C) highest recorded HDL cholesterol (mg/dl), (D) mean total cholesterol (mg/dl), (E) mean LDL cholesterol (mg/dl), (F) mean HDL cholesterol (mg/dl), (G) lowest recorded total cholesterol (mg/dl), (H) lowest recorded LDL cholesterol (mg/dl), and (I) lowest recorded HDL cholesterol (mg/dl). Bonferroni experiment-wide significant associations are marked in bold ($p_{\text{Experiment}} = 0.05/(9 \text{ phenotypes} \times 45 \text{ tissues}) = 1.23 \times 10^{-4}$). Bonferroni tissue-specific value threshold $p < 0.0056$ (0.05/9 phenotypes).

location across all stratification groups (Fig. 4D–F; online Supplementary Table S8). Our results revealed a large number of AN-GReX-pain associations specific to the Low-BMI category (Fig. 4D–F; online Supplementary Table S8); in particular, we note 60 gene–tissue associations with presence of foot pain ($N_{\text{Cases}} = 433$), one of which passed experiment-wide significance (Low-BMI-Small intestine-GPX1, $p = 7.6 \times 10^{-6}$), which may indicate a propensity to excess exercise, or exercise-related injuries among these individuals.

Upregulation of CLIC1 is associated with glucagon medication

Upregulation of CLIC1 was associated with glucagon, a hormone used to treat severe hypoglycemia, in the overall cohort ($N_{\text{Cases}} = 129$) (Overall-Spleen-CLIC1, $p = 4.20 \times 10^{-9}$) and in females ($N_{\text{Cases}} = 78$) (Females-Spleen-CLIC1, $p = 1.4 \times 10^{-8}$) (online Supplementary Fig. S5, Table S9). We additionally find significant associations between the upregulation of CLIC1 and insulin medications in females (Females-Subcutaneous adipose-CLIC1, $p < 9.6 \times 10^{-7}$).

Identifying potential case contamination effects

Our BioMe™ patients have not all been explicitly assessed for eating disorders, and information regarding eating disorder diagnoses and assessments earlier in life may be omitted from the records due to incomplete clinical history assessments. Therefore, it is possible that diagnostic contamination within some of our sample is responsible for the associations observed within our data. We tested whether such contamination may drive the associations observed in our study, assuming the following different contamination scenarios (online Supplementary Table S10), for two genes with large S-PrediXcan effect sizes: (i) that contamination occurs among cases, assuming the highest common estimate for AN prevalence (0.9% among women, 0.3% among men); (ii) that contamination occurs within our biobank, assuming the highest common estimate for AN prevalence (0.6%, 190 cases), and that all of these samples fall into our 'case' category.

Our model indicates that diagnostic contamination among cases is unlikely to result in significant pheWAS associations in our data. Contamination occurring at population ED prevalence (0.6%) did not result in significant associations for any of the
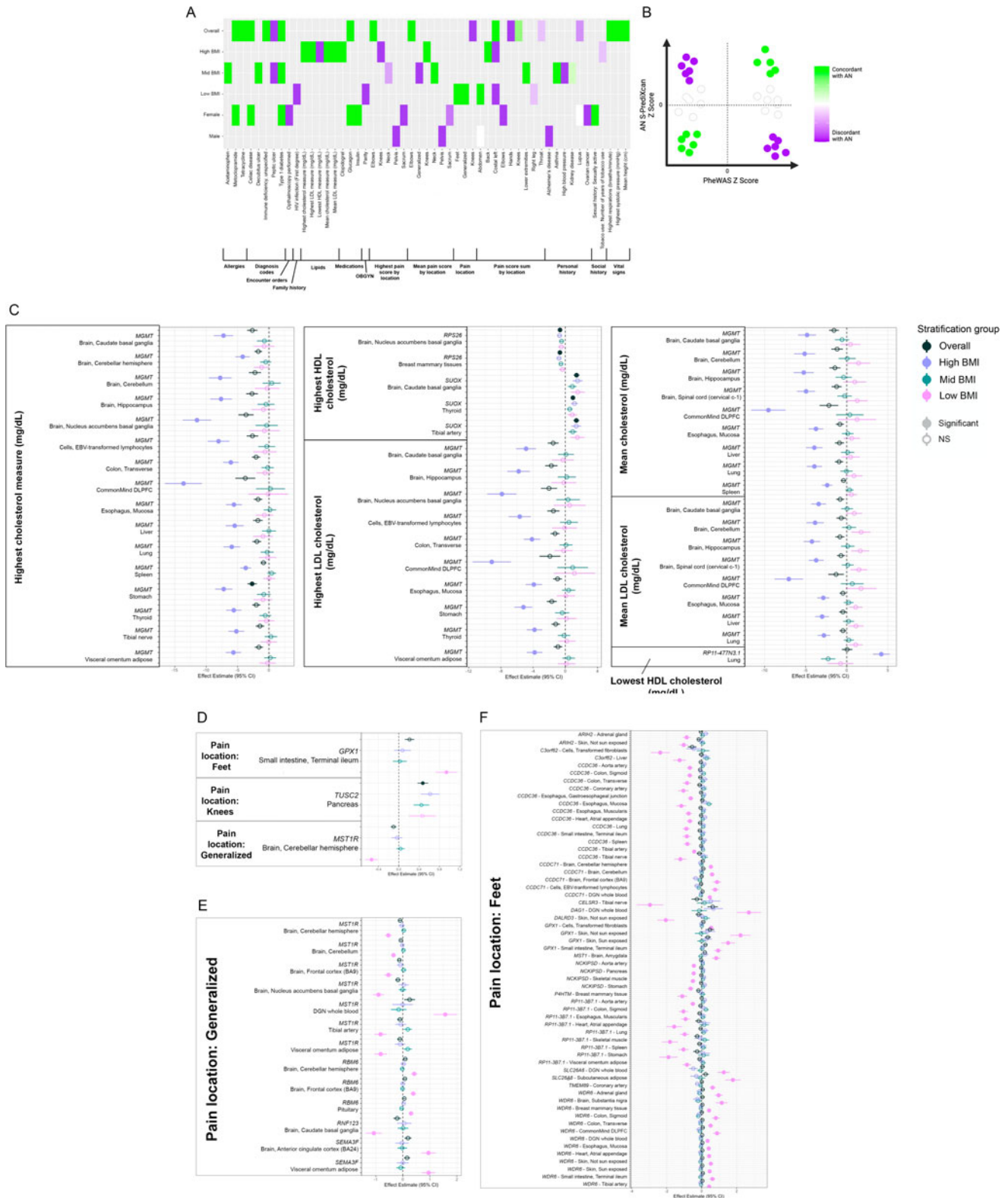
Fig. 4. Context-specific pheWAS associations. (*a*) Concordance of context-specific experiment-wide significant AN-GReX clinical associations with AN direction of effect. We compared the direction of effect (DoE) for each experiment-wide gene–tissue-pheWAS association with the DoE of that gene–tissue pair for AN from our S-PrediXcan results (online Supplemental Methods). For those phenotypes concordant with AN, this may indicate that genetic regulation of those AN-genes is more similar to individuals with AN in individuals with those clinical outcomes. (*b*) Schematic of the proportion of concordance of AN-GReX pheWAS associations with AN S-PrediXcan associations. Associations with similar direction of effect to AN (green) identified as 'concordant', associations with opposite direction of effect (purple) identified as 'discordant'. (*c*) Context-specific associations of AN-GReX with lipid phenotypes of highest recorded, lowest recorded, and mean measures of total cholesterol (mg/dl), HDL cholesterol (mg/dl), and LDL cholesterol (mg/dl). Experiment-wide significant threshold set at $p = 0.05/(9 \text{ phenotypes} \times 45 \text{ tissues}) = 1.2 \times 10^{-4}$. Tissue-specific threshold set at $0.05/(9 \text{ phenotypes}) = 0.0056$. Context-specific associations of AN-GReX with pain location for (*d*) experiment-wide significant associations ($p = 0.05/(99 \text{ phenotypes} \times 45 \text{ tissues}) = 1.1 \times 10^{-5}$), (*e*) all context-specific associations (experiment-wide and tissue-specific) with generalized pain, and (*f*) with foot pain. Tissue-specific threshold for pain location set at $0.05/(99 \text{ phenotypes}) = 5.0 \times 10^{-4}$.

case–control scenarios in our model (online Supplementary Table S10); nor did contamination at 1.2% or 3%. Assuming 6% contamination resulted in potentially significant associations (estimated $p < 2.8 \times 10^{-8}$; see Methods and online Supplementary Methods for more details).

Discussion

Our analysis identified novel AN-genes associated with metabolic, anthropometric, autoimmune, and psychiatric phenotypes (online Supplementary Table S4B). In particular, the experiment-wide significant locus on chromosome 3 overlaps with a known GWAS peak for IBD (de Lange et al., 2017), and includes genes associated with Crohn's and ulcerative colitis (online Supplementary Table S4B) (Hedman et al., 2019; Mårild et al., 2017; Raevuori et al., 2014; Zerwas et al., 2017), in line with earlier AN GWAS showing genetic overlap with autoimmune disease (Gibson & Mehler, 2019). Variants in one of our significant S-PrediXcan genes, GPR75, have been recently shown to have a protective effect against obesity, and have been associated with lower body weight overall (Akbari et al., 2021).

To our knowledge, our results include the first AN-association within the MHC locus (CLIC1, chloride intracellular 1), a region that has been associated with many other psychiatric (Ripke et al., 2014; Stahl et al., 2019) and autoimmune disorders. In particular, CLIC1 has previously been identified as associated with schizophrenia (The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium, 2017), autism (The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium, 2017), MDD (Zhu et al., 2019), post-traumatic stress disorder (Marchese et al., 2021), neuroticism (Baselmans et al., 2019) and depressive phenotypes (Baselmans et al., 2019). Importantly, CLIC1 variants have also been associated with complement component C4 and C3 protein levels in the blood (Yang et al., 2012), which through the complement cascade (Sekar et al., 2016) are involved in immunological functions of pathogen clearance and in synaptic pruning and neuronal connectivity (Stephan, Barres, & Stevens, 2012). CLIC1 encodes a chloride ion channel protein involved in many necessary cellular functions including the regulation of cell membrane potential, and the proliferation and differentiation of cells (Li et al., 2018), including a role in axonal outgrowth of neurons (Carlini et al., 2020).

We also identified multiple genes that have been previously associated with decreased sleep and insomnia phenotypes (including a core circadian clock gene ARNTL), and with a range of psychiatric disorders (Alloy, Ng, Titone, & Boland, 2017; Boivin, 2000; Bunney & Bunney, 2013; Bunney et al., 2015; Etain et al., 2014; Huckins et al., 2017; Karatsoreos, 2014; Kripke, Mullaney, Atkinson, & Wolf, 1978; Mansour et al., 2009; Melo, Abreu, Linhares Neto, de Bruin, & de Bruin, 2017; Meyrer, Demling, Kornhuber, & Nowak, 2009; Murray, Allen, Trinder, & Burgess, 2002), including eating disorders (Allison, Spaeth, & Hopkins, 2016), as well as with satiety and hunger (Arble, Bass, Laposky, Vitaterna, & Turek, 2009; Herpertz et al., 2000).

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

Unlike GWAS, which include carefully constructed case–control cohorts, pheWAS encompass all individuals within a healthcare system, including patients with subthreshold or partial presentations of a disorder, or individuals with commonly comorbid or co-diagnosed conditions. Importantly, because individuals are not ascertained for any specific disorder, they represent a more comprehensive clinical picture of the comorbidities and symptomatology associated with AN gene expression.

Examining the consequences of aberrant predicted gene expression among these patients (i.e. testing for GReX associations) may reveal clinical and biological consequences of these genes; for example, studying whether AN-associated genes have anthropometric and metabolic consequences among adults with no evidence for previous AN- or ED-diagnoses may disentangle whether certain endophenotypes present as a cause or consequence of AN. Food avoidance and restriction may arise due to gastrointestinal (GI) complaints and distress that provoke these behaviors and precede development of AN (Zucker & Bulik, 2020). Similarly, autoimmune disorders of the GI tract, such as celiac disease and Crohn's disease, show a bidirectional relationship with AN, with previous diagnosis of a GI-associated autoimmune disorders increasing the risk of AN and vice versa (Hedman et al., 2019). Our pheWAS results of AN-GReX associations with GI symptoms such as abdominal pain, ascites, and peptic ulcer, as well as GI-related autoimmune disorders such as celiac disease, suggest AN-GReX may contribute directly to these diseases and phenotypes, and that food aversive behaviors and gastric distress may be genetically regulated in these individuals, rather than occurring as a consequence of AN. We found a sex-specific association of CLIC1-GReX with celiac disease in females (online Supplementary Fig. S3B), indicating sex-specific regulation of AN-genes may be contributing to the disparity in autoimmune diagnoses and symptomatology between the sexes (Fuchs et al., 2014; Ludvigsson & Murray, 2019). Our main pheWAS diagnosis results of celiac disease and type 1 diabetes further show concordance with AN in the direction of effect of the tissue GReX, indicating a similarity in the genetic regulation of AN-genes between patients with these disorders and individuals with AN (online Supplementary Fig. 4A; Supplemental Results). Although these associations could be the consequence of undiagnosed AN-individuals in our biobank, they more likely reflect real biological associations of expression of those particular genes with the phenotype.

Our results further confirm the contribution of metabolic factors to AN etiology; we see very robust association of AN-GReX with type 1 diabetes, the hyperglycemic hormone glucagon and various forms of insulin. The association of AN with metabolic traits and abnormalities has been fairly well established, including with insulin resistance ($r_g = -0.29$), fasting insulin ($r_g = -0.24$), leptin ($r_g = -0.26$) and type 2 diabetes ($r_g = -0.22$), and a positive genetic correlation with HDL cholesterol ($r_g = 0.21$) (Adams, Reay, Geaghan, & Cairns, 2021; Ilyas et al., 2019; Watson et al., 2019). Notably, our results point to a similar role of aberrant glycemic regulation in the etiology of AN. Future analyses including EHR-derived lab results (LabWAS) studies may further elucidate AN-genes associated with abnormal metabolic regulation and clinical features.

Context-specific associations reveal a role for BMI in the regulation of AN-gene expression

We stratified our pheWAS analyses by BMI in order to observe whether the effects of predicted AN-gene expression on clinical outcomes were moderated by BMI status. Understanding how

BMI influences associations between AN-GReX and the clinical phenome may give us further insight into the biological mechanisms leading to those outcomes and the conferral of AN disease risk. The context-specific associations we see between BMI status and the association of AN-GReX with multiple phenotypes indicate the potential differences in AN-gene regulation in an environment of higher or lower BMI.

Among our BMI-stratified pheWAS analyses, we identified a number of associations between AN-GReX and foot pain, exclusive to the Low-BMI group. We hypothesize that these associations may stem from exercise-related injuries among this group. Multiple genes in our study, many of which were associated with foot pain in Low-BMI individuals, have been previously associated with regular gym attendance (online Supplementary Table S4B, $p_{adj} = 3.4 \times 10^{-10}$), or measures of physical activity. Excessive and compulsive exercise is a behavior often seen in individuals with AN (Bulik, 2004; Shroff et al., 2006), and evidence is suggestive that hyperactivity increases risk of chronic AN (Achamrah, Coëffier, & Déchelotte, 2016; Strober, Freeman, & Morrell, 1997). Similarly, studies have shown a genetic correlation between AN and physical activity ($r_g = 0.17$) (Hübel et al., 2019b; Watson et al., 2019). These genes may reflect a genetic liability to compulsive physical activity. Given the known association of AN disease with low bone mineral density and propensity for bone fracture (Solmi et al., 2016), our results may reflect the result of genes associated with compulsive exercise and bone mineral density contributing to increased osteoarticular pain.

We saw distinct modulation by BMI of *MGMT*-GReX-cholesterol associations in our cohort. Variants in the *MGMT* gene have been shown to be associated with AN (Watson et al., 2019), as well as anthropometric phenotypes such as waist-to-hip ratio and body height (Kichaev et al., 2019). *CTNNB1*, which was associated with both highest and lowest weight measures in our cohort, distinctly with measures of highest weight in High-BMI individuals, has been previously shown to be associated with lean body mass (Hübel et al., 2019a). We saw distinct associations of *PROS1*-GReX with lowest weight in Low-BMI individuals, which is a gene involved in many biological processes, including inflammation (Suleiman, Négrier, & Boukerche, 2013). Our results appear to be due to the effect of BMI on AN-GReX, rather than the direct effect of BMI/weight on the phenotypes (online Supplementary Fig. S4B). Future studies are needed to assess the specific role of BMI on AN-gene regulation.

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

First, we make use of ICD codes within the medical record in order to infer diagnoses; since these are primarily used by clinicians for billing purposes, they likely provide an imperfect proxy for true disease state. In order to mitigate spurious results stemming from imperfectly assigned codes, we require 'cases' to have at least two instances of an ICD code; individuals with only one code were set to missing. Furthermore, for each code

we restricted our meta-analysis to include only phenotypes with at least 10 cases and 10 controls, and required a total effective sample size >100 for inclusion in our final analysis.

Second, it is possible that our patients regularly receive treatment at multiple different hospital systems; as such, we may be capturing only partial data for each of our patients. In order to mitigate this, we restricted our analysis to patients with multiple data points within our EHR. Future analyses that seek to harmonize or meta-analyze patient data across EHR (e.g. NYC-CDRN, PsycheMERGE) are ongoing, and will be vital to disentangling this effect further.

Importantly, our patients have not all been explicitly assessed for EDs, and information regarding ED diagnoses and assessments earlier in life may be omitted from the records due to incomplete clinical history assessments. Therefore, it is possible that case-contamination within some of our sample is responsible for the associations observed within our data. In order to address this, we performed a simple experiment to simulate the effect sizes expected if undiagnosed AN contamination drives our result. We estimated the expected association statistics observed due to possible levels of contamination among pheWAS cases. Our model indicates that diagnostic contamination at 0.06, 1.2, or 3% is insufficient to account for the effect sizes observed within our pheWAS analysis. Moreover, since these genes are selected due to their large S-PrediXcan effect sizes, we expect that the contamination effects observed for these two genes will be greater than any others in our study; as such, we do not believe that our findings are attributable to hidden case contamination.

Our results illustrate the clinical outcomes associated with differences in AN-gene expression. Characterization of the phenotypic associations with AN-gene expression in a clinical setting can give us more insight into the biological mechanisms underlying AN and, consequently, how to diagnose and treat the disorder. By understanding the associations of AN-gene expression with symptomatology, prodromal or subthreshold disease states, we may gain insights into the biology of the disease, and perhaps identify therapeutic targets and opportunities for clinical interventions. For example, if GI complaints are truly the consequence of aberrant AN-gene expression, and contribute to disordered eating due to GI distress, treatment of those symptoms may help alleviate other AN symptoms or prevent development of later AN (Riedlinger et al., 2020; Wiklund et al., 2019). An understanding of the clinical associations of AN-gene expression can further augment the definition of AN, and could allow clinicians to more broadly identify individuals at greater risk of AN, or those who present with symptom constellations that do not yet meet the established diagnostic threshold.

References

Achamrah, N., Coëffier, M., & Déchelotte, P. (2016). Physical activity in patients with anorexia nervosa. *Nutrition Reviews*, 74(5), 301–311. https://doi.org/10.1093/nutrit/nuw001.

Adams, D. M., Reay, W. R., Geaghan, M. P., & Cairns, M. J. (2021). Investigation of glycaemic traits in psychiatric disorders using Mendelian randomisation revealed a causal relationship with anorexia nervosa. *Neuropsychopharmacology*, 46(6), 1093–1102. https://doi.org/10.1038/s41386-020-00847-w.

Akbari, P., Gilani, A., Sosina, O., Kosmicki, J. A., Khrimian, L., Fang, Y.-Y., … Lotta, L. A. (2021). Sequencing of 640 000 exomes identifies GPR75 variants associated with protection from obesity. *Science*, 373(6550), eabf8683. https://doi.org/10.1126/science.abf8683.

Allison, K. C., Spaeth, A., & Hopkins, C. M. (2016). Sleep and eating disorders. *Current Psychiatry Reports*, 18(10), 92. https://doi.org/10.1007/s11920-016-0728-8.

Alloy, L. B., Ng, T. H., Titone, M. K., & Boland, E. M. (2017). Circadian rhythm dysregulation in bipolar Spectrum disorders. *Current Psychiatry Reports*, 19(4), 21. https://doi.org/10.1007/s11920-017-0772-z.

Arble, D. M., Bass, J., Laposky, A. D., Vitaterna, M. H., & Turek, F. W. (2009). Circadian timing of food intake contributes to weight gain. *Obesity*, 17(11), 2100–2102. https://doi.org/10.1038/oby.2009.264.

Barbeira, A. N., Dickinson, S. P., Bonazzola, R., Zheng, J., Wheeler, H. E., Torres, J. M., … Im, H. K. (2018). Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nature Communications*, 9(1), 1825. https://doi.org/10.1038/s41467-018-03621-1.

Baselmans, B. M. L., Jansen, R., Ip, H. F., van Dongen, J., Abdellaoui, A., van de Weijer, M. P., … Bartels, M. (2019). Multivariate genome-wide analyses of the well-being spectrum. *Nature Genetics*, 51(3), 445–451. https://doi.org/10.1038/s41588-018-0320-8.

Battle, A., Mostafavi, S., Zhu, X., Potash, J. B., Weissman, M. M., McCormick, C., … Koller, D. (2014). Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. *Genome Research*, 24(1), 14–24. https://doi.org/10.1101/gr.155192.113.

Boivin, D. B. (2000). Influence of sleep-wake and circadian rhythm disturbances in psychiatric disorders. *Journal of Psychiatry & Neuroscience: JPN*, 25(5), 446–458.

Bulik, C. M. (2004). Genetic and biological risk factors. In J. K. Thompson (Ed.), *Handbook of eating disorders and obesity* (p. 3). New Jersey: Wiley.

Bulik, C. M., Flatt, R., Abbaspour, A., & Carroll, I. (2019). Reconceptualizing anorexia nervosa. *Psychiatry and Clinical Neurosciences*, 73(9), 518–525. https://doi.org/10.1111/pcn.12857.

Bunney, B. G., & Bunney, W. E. (2013). Mechanisms of rapid antidepressant effects of sleep deprivation therapy: Clock genes and circadian rhythms. *Biological Psychiatry*, 73(12), 1164–1171. https://doi.org/10.1016/j.biopsych.2012.07.020.

Bunney, B. G., Li, J. Z., Walsh, D. M., Stein, R., Vawter, M. P., Cartagena, P., … Bunney, W. E. (2015). Circadian dysregulation of clock genes: Clues to rapid treatments in major depressive disorder. *Molecular Psychiatry*, 20(1), 48–55. https://doi.org/10.1038/mp.2014.138.

Carlini, V., Verduci, I., Cianci, F., Cannavale, G., Fenoglio, C., Galimberti, D., & Mazzanti, M. (2020). CLIC1 protein accumulates in circulating monocyte membrane during neurodegeneration. *International Journal of Molecular Sciences*, 21(4), 1484. https://doi.org/10.3390/ijms21041484.

Carroll, R. J., Bastarache, L., & Denny, J. C. (2014). R pheWAS: Data analysis and plotting tools for phenome-wide association studies in the R environment. *Bioinformatics*, 30(16), 2375–2376. https://doi.org/10.1093/bioinformatics/btu197.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7. https://doi.org/10.1186/s13742-015-0047-8.

Chesney, E., Goodwin, G. M., & Fazel, S. (2014). Risks of all-cause and suicide mortality in mental disorders: A meta-review. *World Psychiatry*, 13(2), 153–160. https://doi.org/10.1002/wps.20128.

de Lange, K. M., Moutsianas, L., Lee, J. C., Lamb, C. A., Luo, Y., Kennedy, N. A., … Barrett, J. C. (2017). Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nature Genetics*, 49(2), 256–261. https://doi.org/10.1038/ng.3760.

Denny, J. C., Bastarache, L., Ritchie, M. D., Carroll, R. J., Zink, R., Mosley, J. D., … Roden, D. M. (2013). Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nature Biotechnology*, 31(12), 1102–1110. https://doi.org/10.1038/nbt.2749.

Denny, J. C., Ritchie, M. D., Basford, M. A., Pulley, J. M., Bastarache, L., Brown-Gentry, K., … Crawford, D. C. (2010). pheWAS: Demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics*, 26(9), 1205–1210. https://doi.org/10.1093/bioinformatics/btq126.

Duncan, L., Yilmaz, Z., Gaspar, H., Walters, R., Goldstein, J., Anttila, V., … Bulik, C. M. (2017). Significant locus and metabolic genetic correlations revealed in genome-wide association study of anorexia Nervosa. *The American Journal of Psychiatry*, 174(9), 850–858. https://doi.org/10.1176/appi.ajp.2017.16121402.

Etain, B., Jamain, S., Milhiet, V., Lajnef, M., Boudebesse, C., Dumaine, A., … Bellivier, F. (2014). Association between circadian genes, bipolar disorders and chronotypes. *Chronobiology International*, 31(7), 807–814. https://doi.org/10.3109/07420528.2014.906445.

Fromer, M., Roussos, P., Sieberts, S. K., Johnson, J. S., Kavanagh, D. H., Perumal, T. M., … Sklar, P. (2016). Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nature Neuroscience*, 19(11), 1442–1453. https://doi.org/10.1038/nn.4399.

Fuchs, V., Kurppa, K., Huhtala, H., Collin, P., Mäki, M., & Kaukinen, K. (2014). Factors associated with long diagnostic delay in celiac disease. *Scandinavian Journal of Gastroenterology*, 49(11), 1304–1310. https://doi.org/10.3109/00365521.2014.923502.

Gamazon, E. R., Wheeler, H. E., Shah, K. P., Mozaffari, S. V., Aquino-Michaels, K., Carroll, R. J., … Im, H. K. (2015). A gene-based association method for mapping traits using reference transcriptome data. *Nature Genetics*, 47(9), 1091–1098. https://doi.org/10.1038/ng.3367.

Gibson, D., & Mehler, P. S. (2019). Anorexia nervosa and the immune system – A narrative review. *Journal of Clinical Medicine*, 8(11), 1915. https://doi.org/10.3390/jcm8111915.

Hedman, A., Breithaupt, L., Hübel, C., Thornton, L. M., Tillander, A., Norring, C., … Bulik, C. M. (2019). Bidirectional relationship between eating disorders and autoimmune diseases. *Journal of Child Psychology and Psychiatry*, 60(7), 803–812. https://doi.org/10.1111/jcpp.12958.

Herpertz, S., Albers, N., Wagner, R., Pelz, B., Köpp, W., Mann, K., … Hebebrand, J. (2000). Longitudinal changes of circadian leptin, insulin and cortisol plasma levels and their correlation during refeeding in patients with anorexia nervosa. *European Journal of Endocrinology*, 142(4), 373–379. https://doi.org/10.1530/eje.0.1420373.

Hoffman, G. E., Bendl, J., Voloudakis, G., Montgomery, K. S., Sloofman, L., Wang, Y.-C., … Roussos, P. (2019). CommonMind Consortium provides transcriptomic and epigenomic data for schizophrenia and bipolar disorder. *Scientific Data*, 6(1), 180. https://doi.org/10.1038/s41597-019-0183-6.

Hübel, C., Abdulkadir, M., Herle, M., Loos, R. J. F., Breen, G., Bulik, C. M., & Micali, N. (2021). One size does not fit all. Genomics differentiates among anorexia nervosa, bulimia nervosa, and binge-eating disorder. *The International Journal of Eating Disorders*, 54(5), 785–793. https://doi.org/10.1002/eat.23481.

Hübel, C., Gaspar, H. A., Coleman, J. R. I., Finucane, H., Purves, K. L., Hanscombe, K. B., … Breen, G. (2019a). Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, 180(6), 428–438. https://doi.org/10.1002/ajmg.b.32709.

Hübel, C., Gaspar, H. A., Coleman, J. R. I., Hanscombe, K. B., Purves, K., Prokopenko, I., … Breen, G. (2019b). Genetic correlations of psychiatric traits with body composition and glycemic traits are sex- and age-dependent. *Nature Communications*, 10(1), 5765. https://doi.org/10.1038/s41467-019-13544-0.

Huckins, L. M., Dobbyn, A., McFadden, W., Wang, W., Ruderfer, D. M., Hoffman, G., … Stahl, E. A. (2017). Transcriptomic imputation of bipolar disorder and bipolar subtypes reveals 29 novel associated genes. *BioRxiv*. https://doi.org/10.1101/222786.

Huckins, L. M., Dobbyn, A., Ruderfer, D. M., Hoffman, G., Wang, W., Pardiñas, A. F., … Stahl, E. A. (2019). Gene expression imputation across multiple brain regions provides insights into schizophrenia risk. *Nature Genetics*, 51(4), 659–674. https://doi.org/10.1038/s41588-019-0364-4.

Ilyas, A., Hübel, C., Stahl, D., Stadler, M., Ismail, K., Breen, G., … Kan, C. (2019). The metabolic underpinning of eating disorders: A systematic review and meta-analysis of insulin sensitivity. *Molecular and Cellular Endocrinology*, 497, 110307. https://doi.org/10.1016/j.mce.2018.10.005.

Karatsoreos, I. N. (2014). Links between circadian rhythms and psychiatric disease. *Frontiers in Behavioral Neuroscience*, 8, 162. https://doi.org/10.3389/fnbeh.2014.00162.

Kichaev, G., Bhatia, G., Loh, P.-R., Gazal, S., Burch, K., Freund, M. K., … Price, A. L. (2019). Leveraging polygenic functional enrichment to improve GWAS power. *American Journal of Human Genetics*, 104(1), 65–75. https://doi.org/10.1016/j.ajhg.2018.11.008.

Kripke, D. F., Mullaney, D. J., Atkinson, M., & Wolf, S. (1978). Circadian rhythm disorders in manic-depressives. *Biological Psychiatry*, 13(3), 335–351.

Leppert, B., Millard, L. A. C., Riglin, L., Davey Smith, G., Thapar, A., Tilling, K., … Stergiakouli, E. (2020). A cross-disorder PRS-pheWAS of 5 major psychiatric disorders in UK biobank. *PLoS Genetics*, 16(5), e1008185. https://doi.org/10.1371/journal.pgen.1008185.

Li, B., Mao, Y., Wang, Z., Chen, Y., Wang, Y., Zhai, C., … Chen, J. (2018). CLIC1 promotes the progression of gastric cancer by regulating the MAPK/AKT pathways. *Cellular Physiology and Biochemistry*, 46(3), 907–924. https://doi.org/10.1159/000488822.

Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., … Moore, H. F. (2013). The genotype-tissue expression (GTEx) project. *Nature Genetics*, 45(6), 580–585. https://doi.org/10.1038/ng.2653.

Ludvigsson, J. F., & Murray, J. A. (2019). Epidemiology of celiac disease. *Gastroenterology Clinics of North America*, 48(1), 1–18. https://doi.org/10.1016/j.gtc.2018.09.004.

Mansour, H. A., Talkowski, M. E., Wood, J., Chowdari, K. V., McClain, L., Prasad, K., … Nimgaonkar, V. L. (2009). Association study of 21 circadian genes with bipolar I disorder, schizoaffective disorder, and schizophrenia. *Bipolar Disorders*, 11(7), 701–710. https://doi.org/10.1111/j.1399-5618.2009.00756.x.

Marchese, S., Cancelmo, L., Diab, O., Cahn, L., Aaronson, C., Daskalakis, N. P., … Feder, A. (2021). Altered gene expression and PTSD symptom dimensions in world trade center responders. *MedRxiv*. https://doi.org/10.1101/2021.03.05.21252989.

Mårild, K., Størdal, K., Bulik, C. M., Rewers, M., Ekbom, A., Liu, E., … Ludvigsson, J. F. (2017). Celiac disease and anorexia nervosa: A nationwide study. *Pediatrics*, 139(5), 1–9. https://doi.org/10.1542/peds.2016-4367.

Melo, M. C. A., Abreu, R. L. C., Linhares Neto, V. B., de Bruin, P. F. C., & de Bruin, V. M. S. (2017). Chronotype and circadian rhythm in bipolar disorder: A systematic review. *Sleep Medicine Reviews*, 34, 46–58. https://doi.org/10.1016/j.smrv.2016.06.007.

Meyrer, R., Demling, J., Kornhuber, J., & Nowak, M. (2009). Effects of night shifts in bipolar disorders and extreme morningness. *Bipolar Disorders*, 11(8), 897–899. https://doi.org/10.1111/j.1399-5618.2009.00767.x.

Mitchell, J. E., & Peterson, C. B. (2020). Anorexia Nervosa. *The New England Journal of Medicine*, 382(14), 1343–1351. https://doi.org/10.1056/NEJMcp1803175.

Munn-Chernoff, M. A., Johnson, E. C., Chou, Y.-L., Coleman, J. R. I., Thornton, L. M., Walters, R. K., … Agrawal, A. (2021). Shared genetic risk between eating disorder- and substance-use-related phenotypes: Evidence from genome-wide association studies. *Addiction Biology*, 26 (e12880), 1–20. https://doi.org/10.1111/adb.12880.

Murray, G., Allen, N. B., Trinder, J., & Burgess, H. (2002). Is weakened circadian rhythmicity a characteristic of neuroticism? *Journal of Affective Disorders*, 72(3), 281–289. https://doi.org/10.1016/s0165-0327(01)00465-7.

Pendergrass, S. A., Brown-Gentry, K., Dudek, S. M., Torstenson, E. S., Ambite, J. L., Avery, C. L., … Ritchie, M. D. (2011). The use of phenome-wide association studies (pheWAS) for exploration of novel genotype-phenotype relationships and pleiotropy discovery. *Genetic Epidemiology*, 35(5), 410–422. https://doi.org/10.1002/gepi.20589.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81(3), 559–575. https://doi.org/10.1086/519795.

Raevuori, A., Haukka, J., Vaarala, O., Suvisaari, J. M., Gissler, M., Grainger, M., … Suokas, J. T. (2014). The increased risk for autoimmune diseases in patients with eating disorders. *PLoS One*, 9(8), e104845. https://doi.org/10.1371/journal.pone.0104845.

Riedlinger, C., Schmidt, G., Weiland, A., Stengel, A., Giel, K. E., Zipfel, S., … Mack, I. (2020). Which symptoms, complaints and complications of the gastrointestinal tract occur in patients with eating disorders? A systematic review and quantitative analysis. *Frontiers in Psychiatry*, 11, 195. https://doi.org/10.3389/fpsyt.2020.00195.

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K.-H., Holmans, P. A., … O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, 511(7510), 421–427. https://doi.org/10.1038/nature13595.

Sekar, A., Bialas, A. R., de Rivera, H., Davis, A., Hammond, T. R., Kamitaki, N., … McCarroll, S. A. (2016). Schizophrenia risk from complex variation of complement component 4. *Nature*, 530(7589), 177–183. https://doi.org/10.1038/nature16549.

Shroff, H., Reba, L., Thornton, L. M., Tozzi, F., Klump, K. L., Berrettini, W. H., … Bulik, C. M. (2006). Features associated with excessive exercise in women with eating disorders. *The International Journal of Eating Disorders*, 39(6), 454–461. https://doi.org/10.1002/eat.20247.

Smoller, J. W. (2018). The use of electronic health records for psychiatric phenotyping and genomics. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, 177(7), 601–612. https://doi.org/10.1002/ajmg.b.32548.

Solmi, M., Veronese, N., Correll, C. U., Favaro, A., Santonastaso, P., Caregaro, L., … Stubbs, B. (2016). Bone mineral density, osteoporosis, and fractures among people with eating disorders: A systematic review and meta-analysis. *Acta Psychiatrica Scandinavica*, 133(5), 341–351. https://doi.org/10.1111/acps.12556.

Stahl, E. A., Breen, G., Forstner, A. J., McQuillin, A., Ripke, S., Trubetskoy, V., … Sklar, P. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature Genetics*, 51(5), 793. https://doi.org/10.1038/s41588-019-0397-8.

Stephan, A. H., Barres, B. A., & Stevens, B. (2012). The complement system: An unexpected role in synaptic pruning during development and disease. *Annual Review of Neuroscience*, 35, 369–389. https://doi.org/10.1146/annurev-neuro-061010-113810.

Strober, M., Freeman, R., & Morrell, W. (1997). The long-term course of severe anorexia nervosa in adolescents: Survival analysis of recovery, relapse, and outcome predictors over 10–15 years in a prospective study. *International Journal of Eating Disorders*, 22(4), 339–360. https://doi.org/10.1002/(SICI)1098-108X(199712)22:4<339::AID-EAT1>3.0.CO;2-N.

Suleiman, L., Négrier, C., & Boukerche, H. (2013). Protein S: A multifunctional anticoagulant vitamin K-dependent protein at the crossroads of

coagulation, inflammation, angiogenesis, and cancer. *Critical Reviews in Oncology/Hematology*, 88(3), 637–654. https://doi.org/10.1016/j.critrevonc.2013.07.004.

The Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium. (2017). Meta-analysis of GWAS of over 16000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Molecular Autism*, 8(1), 21. https://doi.org/10.1186/s13229-017-0137-9.

Watanabe, K., Taskesen, E., van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, 8(1), 1826. https://doi.org/10.1038/s41467-017-01261-5.

Watson, H. J., & Bulik, C. M. (2013). Update on the treatment of anorexia nervosa: Review of clinical trials, practice guidelines and emerging interventions. *Psychological Medicine*, 43(12), 2477–2500. https://doi.org/10.1017/S0033291712002620.

Watson, H. J., Yilmaz, Z., Thornton, L. M., Hübel, C., Coleman, J. R. I., Gaspar, H. A., … Bulik, C. M. (2019). Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nature Genetics*, 51(8), 1207–1214. https://doi.org/10.1038/s41588-019-0439-2.

Watson, Hunna J, Yilmaz, Zeynep, Thornton, Laura M., Hübel, Christopher, Coleman, Jonathan R I, Gaspar, Héléna A., … Bulik, Cynthia M. (2019). Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nature Genetics*, 51(8), 1207–1214. http://dx.doi.org/10.1038/s41588-019-0439-2.

Wiklund, C. A., Kuja-Halkola, R., Thornton, L. M., Hübel, C., Leppä, V., & Bulik, C. M. (2019). Prolonged constipation and diarrhea in childhood and disordered eating in adolescence. *Journal of Psychosomatic Research*, 126, 109797. https://doi.org/10.1016/j.jpsychores.2019.109797.

Willer, C. J., Li, Y., & Abecasis, G. R. (2010). METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, 26(17), 2190–2191. https://doi.org/10.1093/bioinformatics/btq340.

Wu, P., Gifford, A., Meng, X., Li, X., Campbell, H., Varley, T., … Wei, W.-Q. (2019). Mapping ICD-10 and ICD-10-CM codes to phecodes: Workflow development and initial evaluation. *JMIR Medical Informatics*, 7(4), e14325. https://doi.org/10.2196/14325.

Xu, J., Johnson, J. S., Birgegård, A., Jordan, J., Kennedy, M. A., Landén, M., … Huckins, L. M. (2021). Exploring the clinical consequences and genetic aetiology of adult weight trajectories. *MedRxiv*. https://doi.org/10.1101/2021.10.04.21264526.

Yang, X., Sun, J., Gao, Y., Tan, A., Zhang, H., Hu, Y., … Mo, Z. (2012). Genome-wide association study for serum complement C3 and C4 levels in healthy Chinese subjects. *PLoS Genetics*, 8(9), e1002916. https://doi.org/10.1371/journal.pgen.1002916.

Yilmaz, Z., Hardaway, J. A., & Bulik, C. M. (2015). Genetics and epigenetics of eating disorders. *Advances in Genomics and Genetics*, 5, 131–150. https://doi.org/10.2147/AGG.S55776.

Zerwas, S., Larsen, J. T., Petersen, L., Thornton, L. M., Quaranta, M., Koch, S. V., … Bulik, C. M. (2017). Eating disorders, autoimmune, and autoinflammatory disease. *Pediatrics*, 140(6), e20162089. https://doi.org/10.1542/peds.2016-2089.

Zheutlin, A. B., Dennis, J., Karlsson Linnér, R., Moscati, A., Restrepo, N., Straub, P., … Smoller, J. W. (2019). Penetrance and pleiotropy of polygenic risk scores for schizophrenia in 106 160 patients across four health care systems. *American Journal of Psychiatry*, 176(10), 846–855. https://doi.org/10.1176/appi.ajp.2019.18091085.

Zhu, Z., Zhu, X., Liu, C.-L., Shi, H., Shen, S., Yang, Y., … Liang, L. (2019). Shared genetics of asthma and mental health disorders: A large-scale genome-wide cross-trait analysis. *European Respiratory Journal*, 54(6), 1–12. https://doi.org/10.1183/13993003.01507-2019.

Zipfel, S., Giel, K. E., Bulik, C. M., Hay, P., & Schmidt, U. (2015). Anorexia nervosa: Aetiology, assessment, and treatment. *The Lancet. Psychiatry*, 2(12), 1099–1111. https://doi.org/10.1016/S2215-0366(15)00356-9.

Zucker, N. L., & Bulik, C. M. (2020). On bells, saliva, and abdominal pain or discomfort: Early aversive visceral conditioning and vulnerability for anorexia nervosa. *International Journal of Eating Disorders*, 53(4), 508–512. https://doi.org/10.1002/eat.23255.