nature genetics

LETTERS

https://doi.org/10.1038/s41588-019-0439-2

Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa

Hunna J. Watson et al.*

Characterized primarily by a low body-mass index, anorexia nervosa is a complex and serious illness[1], affecting 0.9–4% of women and 0.3% of men[2–4], with twin-based heritability estimates of 50–60%[5]. Mortality rates are higher than those in other psychiatric disorders[6], and outcomes are unacceptably poor[7]. Here we combine data from the Anorexia Nervosa Genetics Initiative (ANGI)[8,9] and the Eating Disorders Working Group of the Psychiatric Genomics Consortium (PGC-ED) and conduct a genome-wide association study of 16,992 cases of anorexia nervosa and 55,525 controls, identifying eight significant loci. The genetic architecture of anorexia nervosa mirrors its clinical presentation, showing significant genetic correlations with psychiatric disorders, physical activity, and metabolic (including glycemic), lipid and anthropometric traits, independent of the effects of common variants associated with body-mass index. These results further encourage a reconceptualization of anorexia nervosa as a metabo-psychiatric disorder. Elucidating the metabolic component is a critical direction for future research, and paying attention to both psychiatric and metabolic components may be key to improving outcomes.

The previous PGC-ED GWAS (3,495 cases, 10,982 controls) estimated the common genetic variant-based heritability of anorexia nervosa to be around 20%, identified the first genome-wide significant locus and reported significant genetic correlations (r_g) between anorexia nervosa and psychiatric and metabolic/anthropometric phenotypes[10]. These r_g analyses pointed toward metabolic etiological factors, as they are robust to reverse causation, although they could be mediated associations[11] or reflect confounding processes[12]. To advance genomic discovery in anorexia nervosa and further explore genetic correlations, we combined samples from ANGI[8,9], the Genetic Consortium for Anorexia Nervosa (GCAN)/Wellcome Trust Case Control Consortium-3 (WTCCC-3)[13] and the UK Biobank[14], quadrupling our sample size.

Our GWAS meta-analysis included 33 datasets comprising 16,992 cases and 55,525 controls of European ancestry from 17 countries (Supplementary Tables 1–4). We had 80% power to detect an odds ratio of 1.09–1.19 (additive model, 0.9% lifetime risk, $\alpha = 5 \times 10^{-8}$, minor allele frequency (MAF) = 0.05–0.5). Typical of complex-trait GWAS, we observed test statistic inflation ($\lambda = 1.22$) consistent with polygenicity, with no evidence of significant population stratification according to the linkage disequilibrium (LD) intercept and attenuation ratio (Supplementary Note and Supplementary Fig. 1). Meta-analysis results were completed for autosomes and the X chromosome. We identified eight loci that exceeded genome-wide significance ($P < 5 \times 10^{-8}$; Table 1 for loci; Fig. 1 for the Manhattan plot; Supplementary Figs. 2 and 3 for the forest and region plots, respectively). Many loci were near the threshold for significance, and no significant heterogeneity of SNP associations across cohorts was detected ($P = 0.15$–0.64; Supplementary Fig. 2). Conditional and joint analysis (GCTA-COJO)[15] confirmed independence of the lead SNPs within the significant loci (Supplementary Table 5). The eight loci were annotated to identify known protein-coding genes (Supplementary Table 6; Supplementary Table 7 reports a gene search restricted to the single-gene loci). The previously reported PGC-ED genome-wide significant variant (rs4622308)[10] on 12q13.2 did not reach genome-wide significance ($P = 7.02 \times 10^{-5}$); however, between-cohort heterogeneity was apparent ($I^2 = 53.7$; Supplementary Fig. 4 and Supplementary Note). The odds ratio was in the same direction in 22 (67%) of the cohorts ($z = 2.00$, $P = 0.05$, two-tailed test).

Although GWAS findings are informative genome-wide, identifying strong hypotheses about their connections to specific genes is not straightforward. We evaluated three ways to connect anorexia nervosa-associated loci identified by GWAS to genes: regulatory chromatin interactions, relationship to brain expression quantitative trait loci (eQTLs; using a superset of CommonMind[16] and GTEx[17]) and the standard approach of gene location within a GWAS locus. The significant anorexia nervosa-associated loci implicated 121 brain-expressed genes, 74% by location, 55% by adult brain eQTL analysis, 93% by regulatory chromatin interaction and 58 genes by all three methods. Supplementary Figure 5 shows the eight GWAS loci, GENCODE gene models, adult brain regulatory chromatin interactions, brain eQTLs and functional genomic annotations.

Four single-gene loci were confirmed by eQTL analyses, chromatin interaction studies or both. These were the locus-intersecting genes *CADM1* (locus 2, chromosome 11: 114.9–115.4 Mb, Supplementary Fig. 5b), *MGMT* (locus 4, chromosome 10: 131.2–131.4 Mb, Supplementary Fig. 5d), *FOXP1* (locus 5, chromosome 3: 70.6–71.0 Mb, Supplementary Fig. 5e) and *PTBP2* (locus 6, chromosome 1: 96.6–97.2 Mb, Supplementary Fig. 5f). For locus 5, eQTL data implicated a distal gene, *GPR27*. One intergenic locus (locus 7, chromosome 5: 24.9–25.3 Mb, Supplementary Fig. 5g) had no eQTL or chromatin interactions, whereas the other intergenic locus (locus 8, chromosome 3: 93.9–95.0 Mb, Supplementary Fig. 5h) had eQTL connections to *PROS1* and *ARL13B*. Two complex multigenic loci had many brain-expressed genes and dense chromatin and eQTL interactions that precluded identification of any single gene (locus 1, chromosome 3: 47.5–51.3 Mb; locus 3, chromosome 2: 53.8–54.3 Mb, Supplementary Fig. 5a,c). The clearest evidence and connections were for the single-gene loci that intersected with *CADM1*, *MGMT*, *FOXP1* and *PTBP2*, and we conclude that these genes may have a role in the etiology of anorexia nervosa (Supplementary Note).

*A full list of authors and affiliations appears at the end of the paper.

Table 1 | Newly associated genome-wide significant loci for anorexia nervosa

Locus	Chromosome	Base-pair region		Lead SNP	Base pair	*P* value	A1/A2	OR	s.e.	Frequency	Type	Number of genes	Nearest gene
		Start	End										
1	3	47,588,253	51,368,253	rs9821797	48,718,253	6.99×10^{-15}	A/T	1.17	0.02	0.12	Multigenic	111	*NCKIPSD*
2	11	114,997,256	115,424,956	rs6589488	115,096,956	6.31×10^{-11}	A/T	1.14	0.02	0.13	Single gene	1	*CADM1*
3	2	53,881,813	54,362,813	rs2287348	54,039,813	5.62×10^{-9}	T/C	1.11	0.02	0.16	Multigenic	13	*ASB3, ERLEC1*
4	10	131,269,764	131,463,964	rs2008387	131,448,764	1.73×10^{-8}	A/G	1.08	0.01	0.33	Single gene	2	*MGMT*
5	3	70,670,750	71,074,150	rs9874207	71,019,750	2.05×10^{-8}	C/T	1.08	0.01	0.49	Single gene	2	*FOXP1*
6	1	96,699,455	97,284,455	rs10747478	96,901,455	3.13×10^{-8}	T/G	1.08	0.01	0.41	Single gene	2	*PTBP2*
7	5	24,945,845	25,372,845	rs370838138	25,081,845	3.17×10^{-8}	G/C	1.08	0.01	0.56	Intergenic	0	*CDH10*
8	3	93,968,107	95,059,107	rs13100344	94,605,107	4.21×10^{-8}	T/A	1.08	0.01	0.54	Intergenic	2	*NSUN3*

The results of the GWAS meta-analysis of anorexia nervosa (16,992 cases and 55,525 controls) are shown, in which eight novel genome-wide significant loci were detected. Chromosome and region (based on hg19) are shown for SNPs with $P<1\times10^{-5}$ and LD-$r^2>0.1$ with the most associated lead SNP, the location of which is given (base pair). A1/A2 refers to allele 1/allele 2. The odds ratio (OR) and s.e. are shown for the association between allele 1 and the phenotype. Frequency indicates the frequency of allele 1 in controls. The number of genes was determined by genomic location, adult brain eQTL, regulatory chromatin interactions and MAGMA gene-wise analysis (see Methods). The nearest gene is the nearest gene within the region of LD 'friends' of the lead variant (LD-$r^2>0.6\pm500$ kb). The meta-analysis was restricted to variants with MAF ≥ 0.01 and information quality (INFO) score ≥ 0.70. All loci were confirmed using forest plots based on consistent direction of effect in the majority of cohorts and using region plots in which neighboring LD friends were required to show a similar effect. Chromosome X was analyzed but had no loci that reached genome-wide significance. Note, although lead variants are annotated to the nearest gene, this does not mean that the gene listed is a causal gene.

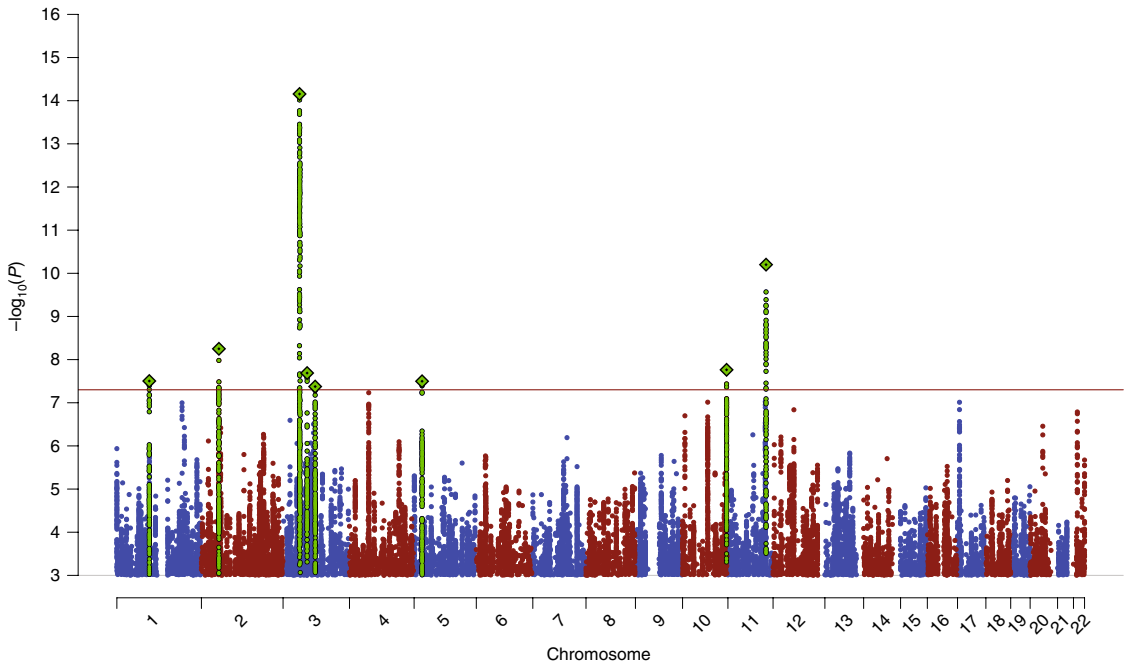

Fig. 1 | The Manhattan plot for the primary genome-wide association meta-analysis of anorexia nervosa with 33 case–control datasets (16,992 cases and 55,525 controls of European descent). The $-\log_{10}(P)$ values for the association tests (two-tailed) are shown on the *y* axis and the chromosomes are ordered on the *x* axis. Eight genetic loci surpassed the genome-wide significance threshold ($-\log_{10}(P)>7.3$; indicated by the line). The lead variant is indicated by a diamond, and green circles show the variants in LD. The blue and red colors differentiate adjacent chromosomes.

Supplementary Table 8 presents multi-trait analysis (GCTA-mtCOJO[18]; conditioning our genome-wide significant SNPs on associated variants in GWAS of body-mass index (BMI), type 2 diabetes, education years, high-density lipoprotein (HDL) cholesterol, neuroticism and schizophrenia. Seven loci appear to be independent. Locus 2 on chromosome 11 may not be unique to anorexia nervosa and may be driven by genetic variation also associated with type 2 diabetes.

Liability-scale SNP heritability (SNP-h^2) was estimated using LD score regression (LDSC)[19,20]. Assuming a lifetime prevalence[2–4] of 0.9–4%, SNP-h^2 was 11–17% (s.e. = 1%), supporting the polygenic nature of anorexia nervosa. Polygenic risk score (PRS) analyses using a leave-one-out approach indicated that the PRS captures approximately 1.7% of the phenotypic variance on the liability scale for discovery $P=0.5$. We did not observe differences in polygenic architecture between anorexia nervosa subtypes with binge eating (2,381 cases, 10,249 controls) or without (2,262 cases, 10,254 controls), or between males (447 cases, 20,347 controls) and females (14,898 cases, 27,545 controls) (Methods, Supplementary Note, Supplementary Fig. 6 and Supplementary Table 9). Similar to females, males in the highest PRS decile had 4.13 (95% confidence interval: 2.58–6.62) times the odds of anorexia nervosa than those

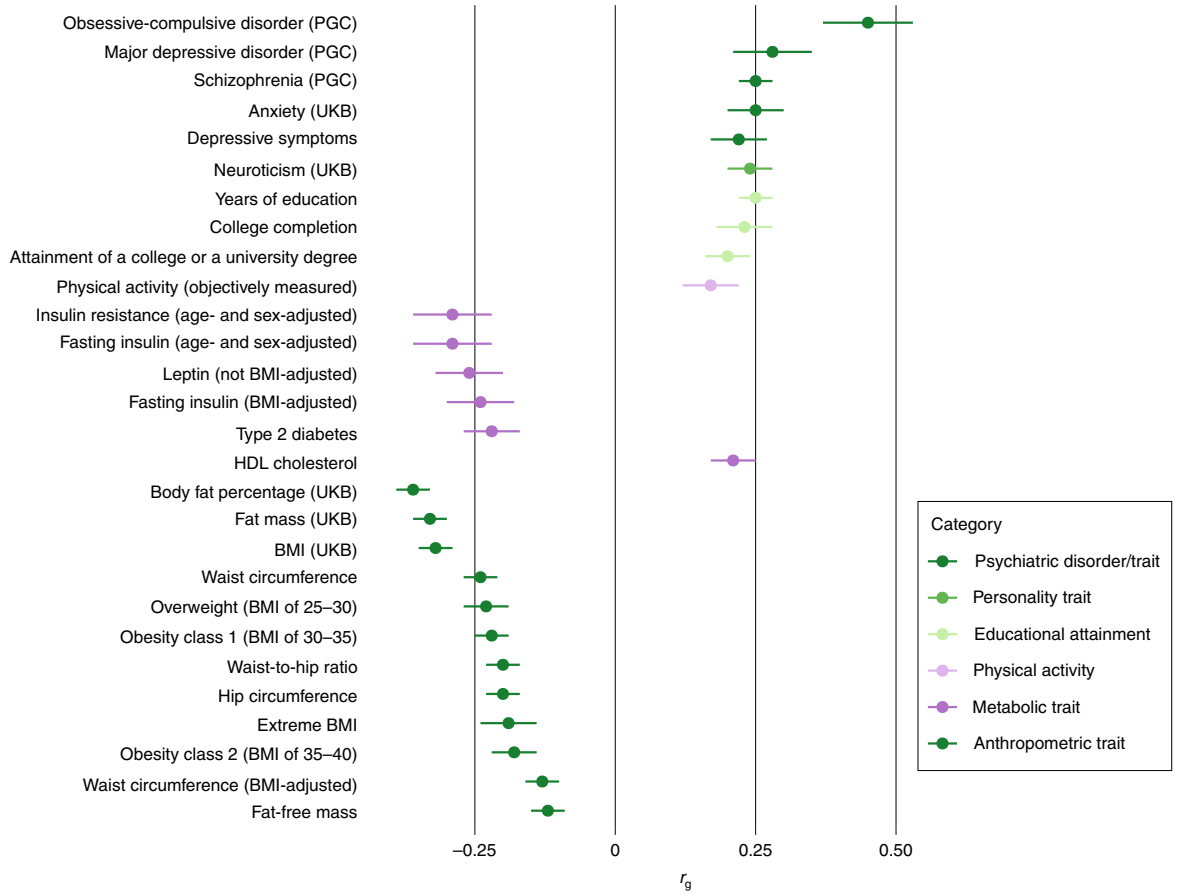

Fig. 2 | Bonferroni-significant genetic correlations (SNP-r_g) between anorexia nervosa and other phenotypes as estimated by LDSC. Only traits with significant *P* values following Bonferroni correction are shown. Error bars show the s.e. Correlations with 447 phenotypes were tested (Bonferroni-corrected significance threshold $P > 1.11 \times 10^{-4}$). Complete results are shown in Supplementary Table 10. Insulin resistance was analysed by the homeostatic model assessment of insulin resistance (HOMA-IR); UKB, UK Biobank.

in the lowest decile. However, confirmation of these results requires larger samples.

We tested SNP-based genetic correlations (SNP-r_g) with external traits using bivariate LDSC[19,20]. Bonferroni-significant SNP-r_g assorted into five trait categories: psychiatric and personality, physical activity, anthropometric traits, metabolic traits and educational attainment (Supplementary Table 10). Figure 2 presents Bonferroni-corrected positive SNP-r_g values associated with obsessive compulsive disorder (SNP-$r_g \pm$ s.e. $= 0.45 \pm 0.08$; $P = 4.97 \times 10^{-9}$), major depressive disorder (0.28 ± 0.07; $P = 8.95 \times 10^{-5}$), anxiety disorders (0.25 ± 0.05; $P = 8.90 \times 10^{-8}$) and schizophrenia (0.25 ± 0.03; $P = 4.61 \times 10^{-18}$). This pattern reflects observed comorbidities in clinical and epidemiological studies[21,22]. The newly identified positive SNP-r_g association with physical activity (0.17 ± 0.05; $P = 1.00 \times 10^{-4}$) encourages further exploration of the refractory symptom of pathologically elevated activity in anorexia nervosa[23]. We note that the significant SNP-r_g of anorexia nervosa with educational attainment (0.25 ± 0.03; $P = 1.69 \times 10^{-15}$) and related constructs was not seen for IQ[24].

Expanding our previous observations[10], we present a number of metabolic and anthropometric r_g with anorexia nervosa that are more pronounced than in other psychiatric disorders. We observed significant negative SNP-r_g with fat mass (-0.33 ± 0.03; $P = 7.23 \times 10^{-25}$), fat-free mass (-0.12 ± 0.03; $P = 4.65 \times 10^{-5}$), BMI (-0.32 ± 0.03; $P = 8.93 \times 10^{-25}$), obesity (-0.22 ± 0.03; $P = 2.96 \times 10^{-11}$), type 2 diabetes (-0.22 ± 0.05; $P = 3.82 \times 10^{-5}$), fasting insulin (-0.24 ± 0.06; $P = 2.31 \times 10^{-5}$), insulin resistance (-0.29 ± 0.07; $P = 2.83 \times 10^{-5}$) and leptin (-0.26 ± 0.06; $P = 4.98 \times 10^{-5}$), and a significant positive SNP-r_g with HDL cholesterol (0.21 ± 0.04; $P = 3.08 \times 10^{-7}$).

Systems biology analyses of our results revealed notable observations (Methods, Supplementary Tables 11–13 and Supplementary Figs. 7–15). Gene-wise analysis with MAGMA prioritized 79 Bonferroni-corrected significant genes, most within the multigenic locus on chromosome 3 (Supplementary Table 11). MAGMA indicated an association with *NCAM1* (Supplementary Table 11), the expression of which increases in response to food restriction in a rodent activity-based anorexia nervosa model[25]. Partitioned heritability analysis showed, as with other GWAS[26], considerable enrichment of SNP-h^2 in conserved regions[27] (fold enrichment $= 24.97$, s.e. $= 3.29$, $P = 3.32 \times 10^{-11}$; Supplementary Fig. 7). Cell type group-specific annotations revealed that the overall SNP-h^2 is significantly enriched for tissues of the central nervous system (Supplementary Fig. 8). One biological pathway was significant, Gene Ontology (GO): positive regulation of embryonic development (32 genes, $P = 1.39 \times 10^{-7}$; Supplementary Table 12), which contains two Bonferroni-corrected significant genes on chromosome 3, *CTNNB1* and *DAG1*. *CTNNB1* encodes catenin β-1, which is part of adherens junctions and a component of Wnt signaling, and *DAG1* encodes dystroglycan, a receptor that binds extracellular matrix proteins[28]. *DAG1* falls within locus 1 (47.5–51.3 Mb). This pathway points to a potential role of developmental processes in the etiology of this complex phenotype (although this is currently speculative). Genes associated with anorexia nervosa were enriched for expression in most brain tissues, particularly

the cerebellum, which has a notably high proportion of neurons[29] (Supplementary Fig. 9). Among 24 brain cell types from mouse brain, significant enrichment was found for medium spiny neurons and pyramidal neurons from hippocampal CA1 (Supplementary Fig. 10). Both medium spiny and pyramidal neurons are linked to feeding behaviors, including food motivation and reward[30,31] (Supplementary Note). Using PrediXcan (Supplementary Note), 36 genes were predicted to be differentially expressed in GTEx tissues or blood (Supplementary Table 13), with the expression of *MGMT* predicted to be downregulated in the caudate. We cautiously note that these results represent the first indications of specific pathways, tissues and cell types that may mediate genetic risk for anorexia nervosa.

Because low BMI is pathognomonic of anorexia nervosa, we investigated the extent to which genetic variants associated with BMI accounted for genetic correlations with metabolic and anthropometric traits. First, covarying for the genetic associations of BMI (Methods) led to a mild but statistically non-significant attenuation of the SNP-r_g between anorexia nervosa and fasting insulin, leptin, insulin resistance, type 2 diabetes and HDL cholesterol (Supplementary Tables 14, 15), suggesting that anorexia nervosa shares genetic variation with these metabolic phenotypes that may be independent of BMI. Second, we investigated bidirectional causality using generalized summary data-based Mendelian randomization (GSMR)[18]. This indicated a significant bidirectional causal relationship such that anorexia nervosa risk-increasing alleles may increase the risk for low BMI, and BMI-lowering alleles may increase the risk of anorexia nervosa (Supplementary Table 16). It is important to note that having only eight genome-wide significant loci for anorexia nervosa render this analysis marginally powered in the direction of anorexia nervosa to BMI, although this analysis is well-powered in the direction of BMI to anorexia nervosa.

Replication is challenging with GWAS of low-prevalence conditions, such as anorexia nervosa, as replication samples must be sufficiently powered to detect the initial findings. We included all available samples in our analysis to maximize chances of reaching the GWAS inflection point, after which there might be a linear increase in hits[32]. The PRS leave-one-out analyses provide evidence of replication by demonstrating a higher burden of common risk variants associated with anorexia nervosa cases, compared with controls, across all the cohorts (Supplementary Fig. 16).

In conclusion, we report multiple genetic loci alongside promising clinical and functional analyses and enrichments. The increased sample size in the present GWAS has allowed us to characterize more fully the metabolic contribution to anorexia nervosa than our previous report[10] by revealing significant r_g with metabolism-related phenotypes, including glycemic and anthropometric traits, and by demonstrating that the effect is robust to correction for the effects of common variants significantly associated with BMI. Low BMI has traditionally been viewed as a consequence of the psychological features of anorexia nervosa (that is, drive for thinness and body dissatisfaction). This perspective has failed to yield interventions that reliably lead to sustained weight gain and psychological recovery[7]. Fundamental metabolic dysregulation may contribute to the exceptional difficulty that individuals with anorexia nervosa have in maintaining a healthy BMI (even after therapeutic renourishment). Our results encourage consideration of both metabolic and psychological drivers of anorexia nervosa when exploring new avenues for treating this frequently lethal illness.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0439-2.

Received: 9 June 2018; Accepted: 14 May 2019;
Published online: 15 July 2019

References

1. Schaumberg, K. et al. The science behind the Academy for Eating Disorders' nine truths about eating disorders. *Eur. Eat. Disord. Rev.* **25**, 432–450 (2017).
2. Keski-Rahkonen, A. & Mustelin, L. Epidemiology of eating disorders in Europe: prevalence, incidence, comorbidity, course, consequences, and risk factors. *Curr. Opin. Psychiatry* **29**, 340–345 (2016).
3. Hudson, J. I., Hiripi, E., Pope, H. G. & Kessler, R. C. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biol. Psychiatry* **61**, 348–358 (2007).
4. Micali, N. et al. Lifetime and 12-month prevalence of eating disorders amongst women in mid-life: a population-based study of diagnoses and risk factors. *BMC Med.* **15**, 12 (2017).
5. Yilmaz, Z., Hardaway, J. A. & Bulik, C. M. Genetics and epigenetics of eating disorders. *Adv. Genomics Genet.* **5**, 131–150 (2015).
6. Arcelus, J., Mitchell, A. J., Wales, J. & Nielsen, S. Mortality rates in patients with anorexia nervosa and other eating disorders: a meta-analysis of 36 studies. *Arch. Gen. Psychiatry* **68**, 724–731 (2011).
7. Watson, H. & Bulik, C. Update on the treatment of anorexia nervosa: review of clinical trials, practice guidelines and emerging interventions. *Psychol. Med.* **43**, 2477–2500 (2013).
8. Kirk, K. M. et al. The Anorexia Nervosa Genetics Initiative: study description and sample characteristics of the Australian and New Zealand arm. *Aust. N. Z. J. Psychiatry* **51**, 583–594 (2017).
9. Thornton, L. M. et al. The Anorexia Nervosa Genetics Initiative (ANGI): overview and methods. *Contemp. Clin. Trials* **74**, 61–69 (2018).
10. Duncan, L. et al. Significant locus and metabolic genetic correlations revealed in genome-wide association study of anorexia nervosa. *Am. J. Psychiatry* **174**, 850–858 (2017).
11. Pickrell, J. K. et al. Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* **48**, 709–717 (2016).
12. Martin, J., Taylor, M. J. & Lichtenstein, P. Assessing the evidence for shared genetic risks across psychiatric disorders and traits. *Psychol. Med.* **48**, 1759–1774 (2018).
13. Boraska, V. et al. A genome-wide association study of anorexia nervosa. *Mol. Psychiatry* **19**, 1085–1094 (2014).
14. Sudlow, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).
15. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
16. Fromer, M. et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.* **19**, 1442–1453 (2016).
17. GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).
18. Zhu, Z. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).
19. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
20. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
21. Cederlöf, M. et al. Etiological overlap between obsessive-compulsive disorder and anorexia nervosa: a longitudinal cohort, multigenerational family and twin study. *World Psychiatry* **14**, 333–338 (2015).
22. Kaye, W. H., Bulik, C. M., Thornton, L., Barbarich, N. & Masters, K. Comorbidity of anxiety disorders with anorexia and bulimia nervosa. *Am. J. Psychiatry* **161**, 2215–2221 (2004).
23. Dalle Grave, R., Calugi, S. & Marchesini, G. Compulsive exercise to control shape or weight in eating disorders: prevalence, associated features, and treatment outcome. *Compr. Psychiatry* **49**, 346–352 (2008).
24. Savage, J. E. et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
25. Ho, E. V., Klenotich, S. J., McMurray, M. S. & Dulawa, S. C. Activity-based anorexia alters the expression of BDNF transcripts in the mesocorticolimbic reward circuit. *PLoS ONE* **11**, e0166756 (2016).
26. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
27. Lindblad-Toh, K. et al. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478**, 476–482 (2011).
28. Bello, V. et al. The dystroglycan: nestled in an adhesome during embryonic development. *Dev. Biol.* **401**, 132–142 (2015).
29. Azevedo, F. A. C. et al. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* **513**, 532–541 (2009).

30. O'Connor, E. C. et al. Accumbal D1R neurons projecting to lateral hypothalamus authorize feeding. *Neuron* **88**, 553–564 (2015).
31. Kim, J., Zhang, X., Muralidhar, S., LeBlanc, S. A. & Tonegawa, S. Basolateral to central amygdala neural circuits for appetitive behaviors. *Neuron* **93**, 1464–1479 (2017).
32. Levinson, D. F. et al. Genetic studies of major depressive disorder: why are there no genome-wide association study findings and what can we do about it? *Biol. Psychiatry* **76**, 510–512 (2014).

Acknowledgements

Grant support for ANGI, the PGC-ED and its component groups is shown in Supplementary Table 17. We thank all study volunteers, study coordinators and research staff who enabled this study. ANGI: The Anorexia Nervosa Genetics Initiative is an initiative of the Klarman Family Foundation. Additional support was offered by the National Institute of Mental Health. We are deeply grateful to all of the individuals who, through their participation, made ANGI a success. The goodwill that permeated the eating disorders community fueled by the enthusiasm of prominent bloggers, advocates, clinicians, treatment centers, scientists, organizations, families and especially those who have suffered from anorexia nervosa, yielded in an unprecedented and inspired global movement to complete this science. ANGI (United States): We thank Walden Behavioral Care, McCallum Place and the Renfrew Center for assisting with recruitment. We express our gratitude to J. Alexander (http://www.junealexander.com/) and C. Arnold (http://carriearnold.com/), who helped us disseminate information about ANGI. We acknowledge support from the North Carolina Translational and Clinical Sciences Institute (NC TraCS), the Carolina Data Warehouse and the Foundation of Hope, Raleigh, North Carolina. ANGI (Australia and New Zealand): We thank the Australia & and New Zealand Academy for Eating Disorders for assistance with recruitment and publicity. We thank VIVA! Communications for their efforts in promoting the study and the Butterfly Foundation for their ongoing support of anorexia nervosa research in Australia and EDANZ in New Zealand. We thank the QSkin Sun and Health Study for controls. We also acknowledge the assistance of S. Maguire and J. Russell (University of Sydney), P. Hay (Western Sydney University), S. Madden (Western Sydney University and the Sydney Children's Hospital Network), S. Sawyer and E. Hughes (Royal Children's Hospital, Melbourne), K. Fairweather-Schmidt (Flinders University), A. Fursland (Centre for Clinical Interventions and Curtin University), J. McCormack (Princess Margaret Hospital for Children), F. Wagg (Royal Hobart Hospital) and W. Ward (Royal Brisbane and Women's Hospital) in recruitment. We also thank L. Nunn for validation work on the ED100Kv1 Questionnaire. Additionally, administrative support for data collection was received from the Australian Twin Registry, which is supported by an Enabling Grant (ID 310667) from the NHMRC administered by the University of Melbourne. In New Zealand, we also acknowledge assistance with recruitment from M. Roberts (University of Auckland), R. Lawson (South Island Eating Disorders Service), M. Meiklejohn (Auckland District Health Board) and R. Mysliwiec. Special thanks to those who provided their stories in relation to publicity about ANGI. ANGI (Sweden): We acknowledge the assistance of the Stockholm Centre for Eating Disorders (SCÄ) and thank the Swedish National Quality Register for Eating Disorders (Riksät) and Lifegene. We would also like to thank the research nurses and data collectors at the Department of Medical Epidemiology and Biostatistics who worked on ANGI. ANGI (Denmark): We thank the Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH). PGC: We are deeply indebted to the investigators who comprise the PGC and to the hundreds of thousands of individuals who have shared their life experiences with PGC investigators and the contributing studies. We are grateful to the Children's Hospital of Philadelphia (CHOP), the Price Foundation Collaborative Group (PFCG), Genetic Consortium for Anorexia Nervosa (GCAN), Wellcome Trust Case-Control Consortium-3 (WTCCC-3), the UK Biobank and all PGC-ED members for their support in providing individual samples used in this study. We thank SURFsara (http://www.surf.nl) for support in using the Lisa Compute Cluster. We thank M. Lam for Ricopili consultation. This study also represents independent research partly funded by the English National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London and the NIHR BioResource. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the English Department of Health and Social Care. High performance computing facilities were funded with capital equipment grants from the GSTT Charity (TR130505) and Maudsley Charity (980). Research reported in this publication was also supported by the National Institute of Mental Health of the US National Institutes of Health under Award Numbers U01MH109528 and U01MH109514. The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health.

Author contributions

C.M.B. and P.F.S. conceived and designed the study. L.M.T., C.M.B. and G.B. performed overall study coordination. C.M.B. was the lead principal investigator of ANGI, and P.F.S. was a co-investigator of ANGI. N.G.M., M.L. and P.B.M. were site principal investigators of ANGI. H.J.W., Z.Y., J.R.I.C., C.H., J.B., H.A.G., S.Y., V.M.L., M. Mattheisen, P.G.-R. and S.E.M. performed the statistical analyses. H.J.W., Z.Y., C.H., J.R.I.C., H.A.G., J.B., A.H., P.G.-R., P.F.S., G.B. and C.M.B. comprised the writing group. C.M.B. and G.B. were PGC-ED co-chairs. S. Ripke provided statistical consultation. A.H. assisted with data interpretation. A.W.B., C.M.B., J.J., M.K., K.M.K., P.L., N.G.M., C.N., R.P., L.M.T. and T.D.W. collected and managed the ANGI samples at sites and assisted with site-specific study co-ordination. A.W.B., J.M.B., H.B., S. Crawford, K.A.H., L.J.H., C.J., A.S.K., W.H.K., J.M., C.M.O., J.F.P., N.L.P., M.S., T.W., D.C.W. and D.B.W. provided ANGI controls and extra samples. L.E.D. provided data expertise. S. Gordon, J. Grove, A.K.H., A. Juréus, K.M.K., J.T.L., R.P. and L. Petersen contributed to the ANGI study. S. Gordon, J. Grove, K.K., J.T.L., M. Mattheisen, S. Medland and L. Petersen were ANGI site analysts. K.B.H. and K.L.P. conducted additional secondary analyses. G.W.M., T.D.W., A.B., P.L. and C.N. were ANGI investigators. J.J. and M.K. assisted with ANGI recruitment in New Zealand. PGC-ED members and other individuals contributed to sample acquisition and made individual data from subjects available: R.A.H.A., L.A., T.A., O.A.A., J.H.B., A.W.B., W.H.B., A.B., I.B., C.B., J.M.B., H.B., G.B., K.B., C.M.B., R.B., M. Cassina, S. Cichon, M. Clementi, J.R.I.C., R.D.C., P.C., S. Crawford, S. Crow, J.J.C., U.N.D., O.S.P.D., M.D.Z., G.D., D. Degortes, D.M.D., D. Dikeos, C.D., M.D.W., E.D., K.E., S.E., G.E., T.E., X.E., A. Farmer, A. Favaro, F.F.A., M.M.F., K.F., M. Föcker, L.F., A.J.F., M. Forzan, S. Gallinger, I.G., J. Giuranna, F.G., P.G., M.G.M., J. Grove, S. Guillaume, K.A.H., K.H., J. Hauser, J. Hebebrand, S.G.H., A.K.H., S.H., B.H.D., W.H., A.H., L.J.H., J.I.H., H. Imgart, H. Inoko, V.J., S.J.M., C.J., J.J., A. Julià, G.K., D.K., A.S.K., J.K., L. Karhunen, A.K., M.J.H.K., W.H.K., J.L.K., M.K., A.K., K.K., Y.K., L. Klareskog, G.P.S.K., M.C.L., M.L., S.L.H., R.D.L., P.L., L.L., B.D.L., J. Lissowska, J. Luykx, P.J.M., M. Maj, K. Mannik, S. Marsal, C.R.M., N.G.M., M. Mattheisen, M. Mattingsdal, S. McDevitt, P. McGuffin, A.M., I.M., N.M., J.M., A.M.M., P. Monteleone, P.B.M., M.A.M.C., B.N., M.N., C.N., I.N., C.M.O., J.K.O., R.A.O., L. Padyukov, A.P., J.P., H.P., N.L.P., J.F.P., D.P., R.R., A. Raevuori, N.R., T.R.K., V.R., S. Ripatti, F. Ritschel, M.R., A. Rotondo, D.R., F. Rybakowski, P.S., S.W.S., U.S., A. Schosser, J.S., L.S., P.E.S., M.C.T.S.L., A. Slopien, S.S., M.S., G.D.S., P.F.S., B.Ś., J.P.S., I.T., E.T., A. Tortorella, F.T., J.T., A. Tsitsika, M.T.N., K.T., A.A.V.E., E.V.F.E., T.D.W., G.W., E. Walton, H.J.W., T.W., D.C.W., E. Widen, D.B.W., S. Zerwas and S. Zipfel.

Competing interests

O.A.A. received a speaker's honorarium from Lundbeck. G.B. received grant funding and consultancy fees in preclinical genetics from Eli Lilly, consultancy fees from Otsuka and has received honoraria from Illumina. C.M.B. is a grant recipient from Shire Pharmaceuticals and served on Shire Scientific Advisory Board; she receives author royalties from Pearson. D.D. served as a speaker and on advisory boards, and has received consultancy fees for participation in research from various pharmaceutical industry companies including: AstraZeneca, Boehringer, Bristol Myers Squibb, Eli Lilly, Genesis Pharma, GlaxoSmithKline, Janssen, Lundbeck, Organon, Sanofi, UniPharma and Wyeth; he has received unrestricted grants from Lilly and AstraZeneca as director of the Sleep Research Unit of Eginition Hospital (National and Kapodistrian University of Athens, Greece). J.I.H. has received grant support from Shire and Sunovion, and has received consulting fees from DiaMentis, Shire, and Sunovion. A.S.K. is a member of the Shire Canadian BED Advisory Board and is on the steering committee for the Shire B/educated Educational Symposium: 15–16 June 2018. J.L.K. served as an unpaid member of the scientific advisory board of AssurexHealth Inc. M.L. declares that, over the past 36 months, he has received lecture honoraria from Lundbeck and served as scientific consultant for EPID Research Oy, but has received no other equity ownership, profit-sharing agreements, royalties or patents. P.F.S. is on the Lundbeck advisory committee and is a Lundbeck grant recipient; he has served on the scientific advisory board for Pfizer, has received a consultation fee from Element Genomics, and a speaker reimbursement fee from Roche. J.T. has received an honorarium for participation in an EAP meeting and has received royalties from several books from Routledge, Wiley and Oxford University Press. T.W. has acted as a lecturer and scientific advisor to H. Lundbeck A/S. All other authors have no conflicts of interest to disclose.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-019-0439-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to C.M.B.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Hunna J. Watson[1,2,3], Zeynep Yilmaz[1,4,173], Laura M. Thornton[1,173], Christopher Hübel[5,6,173], Jonathan R. I. Coleman[5,7,173], Héléna A. Gaspar[5,7], Julien Bryois[6], Anke Hinney[8], Virpi M. Leppä[6], Manuel Mattheisen[9,10,11,12], Sarah E. Medland[13], Stephan Ripke[14,15,16], Shuyang Yao[6], Paola Giusti-Rodríguez[4], Anorexia Nervosa Genetics Initiative[17], Ken B. Hanscombe[18], Kirstin L. Purves[5], Eating Disorders Working Group of the Psychiatric Genomics Consortium[17], Roger A. H. Adan[19,20,21], Lars Alfredsson[22], Tetsuya Ando[23], Ole A. Andreassen[24], Jessica H. Baker[1], Wade H. Berrettini[25], Ilka Boehm[26], Claudette Boni[27], Vesna Boraska Perica[28,29], Katharina Buehren[30], Roland Burghardt[31], Matteo Cassina[32], Sven Cichon[33], Maurizio Clementi[32], Roger D. Cone[34], Philippe Courtet[35], Scott Crow[36], James J. Crowley[4,10], Unna N. Danner[20], Oliver S. P. Davis[37,38], Martina de Zwaan[39], George Dedoussis[40], Daniela Degortes[41], Janiece E. DeSocio[42], Danielle M. Dick[43], Dimitris Dikeos[44], Christian Dina[45], Monika Dmitrzak-Weglarz[46], Elisa Docampo[47,48,49], Laramie E. Duncan[50], Karin Egberts[51], Stefan Ehrlich[26], Geòrgia Escaramís[47,48,49], Tõnu Esko[52,53], Xavier Estivill[47,48,49,54], Anne Farmer[5], Angela Favaro[41], Fernando Fernández-Aranda[55,56], Manfred M. Fichter[57,58], Krista Fischer[52], Manuel Föcker[8], Lenka Foretova[59], Andreas J. Forstner[33,60,61,62,63], Monica Forzan[32], Christopher S. Franklin[28], Steven Gallinger[64], Ina Giegling[65], Johanna Giuranna[8], Fragiskos Gonidakis[66], Philip Gorwood[27,67], Monica Gratacos Mayora[47,48,49], Sébastien Guillaume[35], Yiran Guo[68], Hakon Hakonarson[68,69], Konstantinos Hatzikotoulas[28,70], Joanna Hauser[71], Johannes Hebebrand[8], Sietske G. Helder[5,72], Stefan Herms[33,61,63], Beate Herpertz-Dahlmann[30], Wolfgang Herzog[73], Laura M. Huckins[28,74], James I. Hudson[75], Hartmut Imgart[76], Hidetoshi Inoko[77], Vladimir Janout[78], Susana Jiménez-Murcia[55,56], Antonio Julià[79], Gursharan Kalsi[5], Deborah Kaminská[80], Jaakko Kaprio[81,82], Leila Karhunen[83], Andreas Karwautz[84], Martien J. H. Kas[19,85], James L. Kennedy[86,87,88], Anna Keski-Rahkonen[81], Kirsty Kiezebrink[89], Youl-Ri Kim[90], Lars Klareskog[91], Kelly L. Klump[92], Gun Peggy S. Knudsen[93], Maria C. La Via[1], Stephanie Le Hellard[94,95,96], Robert D. Levitan[86,87,88], Dong Li[68], Lisa Lilenfeld[97], Bochao Danae Lin[19], Jolanta Lissowska[98], Jurjen Luykx[19], Pierre J. Magistretti[99,100], Mario Maj[101], Katrin Mannik[52,102], Sara Marsal[79], Christian R. Marshall[103], Morten Mattingsdal[24], Sara McDevitt[104,105], Peter McGuffin[5], Andres Metspalu[52,106], Ingrid Meulenbelt[107], Nadia Micali[108,109,110], Karen Mitchell[111], Alessio Maria Monteleone[101], Palmiero Monteleone[112], Melissa A. Munn-Chernoff[1], Benedetta Nacmias[113], Marie Navratilova[59], Ioanna Ntalla[40], Julie K. O'Toole[114], Roel A. Ophoff[19,115], Leonid Padyukov[91], Aarno Palotie[53,82,116], Jacques Pantel[27], Hana Papezova[80], Dalila Pinto[74], Raquel Rabionet[117,118,119], Anu Raevuori[81], Nicolas Ramoz[27], Ted Reichborn-Kjennerud[93,120], Valdo Ricca[113,121], Samuli Ripatti[53,81,122], Franziska Ritschel[26,123], Marion Roberts[5,124,125], Alessandro Rotondo[126], Dan Rujescu[57,65], Filip Rybakowski[127], Paolo Santonastaso[128], André Scherag[129], Stephen W. Scherer[130], Ulrike Schmidt[7,131], Nicholas J. Schork[132], Alexandra Schosser[133], Jochen Seitz[30], Lenka Slachtova[134], P. Eline Slagboom[107], Margarita C. T. Slof-Op 't Landt[135,136], Agnieszka Slopien[137], Sandro Sorbi[113,138], Beata Świątkowska[139], Jin P. Szatkiewicz[4], Ioanna Tachmazidou[28], Elena Tenconi[41], Alfonso Tortorella[140,141], Federica Tozzi[142], Janet Treasure[7,131], Artemis Tsitsika[143], Marta Tyszkiewicz-Nwafor[137], Konstantinos Tziouvas[144], Annemarie A. van Elburg[20,145], Eric F. van Furth[135,136], Gudrun Wagner[84], Esther Walton[26], Elisabeth Widen[82], Eleftheria Zeggini[28,70], Stephanie Zerwas[1], Stephan Zipfel[146], Andrew W. Bergen[147,148], Joseph M. Boden[149], Harry Brandt[150], Steven Crawford[150], Katherine A. Halmi[151], L. John Horwood[149], Craig Johnson[152], Allan S. Kaplan[86,87,88], Walter H. Kaye[153], James E. Mitchell[154], Catherine M. Olsen[13],

John F. Pearson[155], Nancy L. Pedersen[6], Michael Strober[156,157], Thomas Werge[158], David C. Whiteman[13], D. Blake Woodside[87,88,159,160], Garret D. Stuber[1,161], Scott Gordon[13], Jakob Grove[9,162,163,164], Anjali K. Henders[165], Anders Juréus[6], Katherine M. Kirk[13], Janne T. Larsen[162,166,167], Richard Parker[13], Liselotte Petersen[162,166,167], Jennifer Jordan[124,168], Martin Kennedy[169], Grant W. Montgomery[13,165,170], Tracey D. Wade[171], Andreas Birgegård[10,11], Paul Lichtenstein[6], Claes Norring[10,11], Mikael Landén[6,172,173], Nicholas G. Martin[13,173], Preben Bo Mortensen[162,166,167,173], Patrick F. Sullivan[1,4,6,173], Gerome Breen[5,7,174] and Cynthia M. Bulik[1,6,173,174]*

[1]Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [2]School of Psychology, Curtin University, Perth, Western Australia, Australia. [3]School of Paediatrics and Child Health, University of Western Australia, Perth, Western Australia, Australia. [4]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [5]Institute of Psychiatry, Psychology and Neuroscience, Social, Genetic and Developmental Psychiatry (SGDP) Centre, King's College London, London, UK. [6]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [7]National Institute for Health Research Biomedical Research Centre, King's College London and South London and Maudsley National Health Service Foundation Trust, London, UK. [8]Department of Child and Adolescent Psychiatry, University Hospital Essen, University of Duisburg-Essen, Essen, Germany. [9]Department of Biomedicine, Aarhus University, Aarhus, Denmark. [10]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden. [11]Center for Psychiatry Research, Stockholm Health Care Services, Stockholm City Council, Stockholm, Sweden. [12]Department of Psychiatry, Psychosomatics and Psychotherapy, University of Würzburg, Würzburg, Germany. [13]QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [14]Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA, USA. [15]Stanley Center for Psychiatric Research, Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA. [16]Department of Psychiatry and Psychotherapy, Charité - Universitätsmedizin, Berlin, Germany. [17]A list of members and affiliations appears in the Supplementary Note. [18]Department of Medical and Molecular Genetics, King's College London, Guy's Hospital, London, UK. [19]Brain Center Rudolf Magnus, Department of Translational Neuroscience, University Medical Center Utrecht, Utrecht, the Netherlands. [20]Center for Eating Disorders Rintveld, Altrecht Mental Health Institute, Zeist, the Netherlands. [21]Institute of Neuroscience and Physiology, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden. [22]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [23]Department of Behavioral Medicine, National Institute of Mental Health, National Center of Neurology and Psychiatry, Tokyo, Japan. [24]NORMENT KG Jebsen Centre, Division of Mental Health and Addiction, University of Oslo, Oslo University Hospital, Oslo, Norway. [25]Department of Psychiatry, Center for Neurobiology and Behavior, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, USA. [26]Division of Psychological and Social Medicine and Developmental Neurosciences, Faculty of Medicine, Technische Universität Dresden, Dresden, Germany. [27]INSERM 1266, Institute of Psychiatry and Neuroscience of Paris, Paris, France. [28]Wellcome Sanger Institute, Hinxton, UK. [29]Department of Medical Biology, School of Medicine, University of Split, Split, Croatia. [30]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, RWTH Aachen University, Aachen, Germany. [31]Department of Child and Adolescent Psychiatry, Klinikum Frankfurt/Oder, Frankfurt, Germany. [32]Clinical Genetics Unit, Department of Woman and Child Health, University of Padova, Padova, Italy. [33]Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, Switzerland. [34]Life Sciences Institute and Department of Molecular and Integrative Physiology, University of Michigan, Ann Arbor, MI, USA. [35]Department of Emergency Psychiatry and Post-Acute Care, CHRU Montpellier, University of Montpellier, Montpellier, France. [36]Department of Psychiatry, University of Minnesota, Minneapolis, MN, USA. [37]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK. [38]School of Social and Community Medicine, University of Bristol, Bristol, UK. [39]Department of Psychosomatic Medicine and Psychotherapy, Hannover Medical School, Hannover, Germany. [40]Department of Nutrition and Dietetics, Harokopio University, Athens, Greece. [41]Department of Neurosciences, University of Padova, Padova, Italy. [42]College of Nursing, Seattle University, Seattle, WA, USA. [43]Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA. [44]Department of Psychiatry, Athens University Medical School, Athens University, Athens, Greece. [45]L'institut du thorax, INSERM, CNRS, UNIV Nantes, CHU Nantes, Nantes, France. [46]Department of Psychiatric Genetics, Poznan University of Medical Sciences, Poznan, Poland. [47]Barcelona Institute of Science and Technology, Barcelona, Spain. [48]Universitat Pompeu Fabra, Barcelona, Spain. [49]Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain. [50]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA. [51]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital of Würzburg, Centre for Mental Health, Würzburg, Germany. [52]Estonian Genome Center, University of Tartu, Tartu, Estonia. [53]Program in Medical and Population Genetics, Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA. [54]Genomics and Disease, Bioinformatics and Genomics Programme, Centre for Genomic Regulation, Barcelona, Spain. [55]Department of Psychiatry, University Hospital of Bellvitge –IDIBELL and CIBERobn, Barcelona, Spain. [56]Department of Clinical Sciences, School of Medicine, University of Barcelona, Barcelona, Spain. [57]Department of Psychiatry and Psychotherapy, Ludwig-Maximilians-University (LMU), Munich, Germany. [58]Schön Klinik Roseneck affiliated with the Medical Faculty of the University of Munich (LMU), Munich, Germany. [59]Department of Cancer, Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic. [60]Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany. [61]Department of Genomics, Life and Brain Center, University of Bonn, Bonn, Germany. [62]Department of Psychiatry (UPK), University of Basel, Basel, Switzerland. [63]Department of Biomedicine, University of Basel, Basel, Switzerland. [64]Department of Surgery, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada. [65]Department of Psychiatry, Psychotherapy and Psychosomatics, Martin Luther University of Halle-Wittenberg, Halle, Germany. [66]First Psychiatric Department, National and Kapodistrian University of Athens, Medical School, Eginition Hospital, Athens, Greece. [67]CMME, Hôpital Sainte-Anne (GHU Paris Psychiatrie et Neurosciences), Paris Descartes University, Paris, France. [68]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA. [69]Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [70]Institute of Translational Genomics, Helmholtz Zentrum München, Neuherberg, Germany. [71]Department of Adult Psychiatry, Poznan University of Medical Sciences, Poznan, Poland. [72]Zorg op Orde, Leidschendam, the Netherlands. [73]Department of General Internal Medicine and Psychosomatics, Heidelberg University Hospital, Heidelberg University, Heidelberg, Germany. [74]Department of Psychiatry, and Genetics and Genomics Sciences, Division of Psychiatric Genomics, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [75]Biological Psychiatry Laboratory, McLean Hospital/Harvard Medical School, Boston, MA, USA. [76]Eating Disorders Unit, Parklandklinik, Bad Wildungen, Germany. [77]Department of Molecular Life Science, Division of Basic Medical Science and Molecular Medicine, School of Medicine, Tokai University, Isehara, Japan. [78]Faculty of Health Sciences, Palacky University, Olomouc, Czech Republic. [79]Rheumatology Research Group, Vall d'Hebron Research Institute, Barcelona, Spain. [80]Department of Psychiatry, First Faculty of Medicine,

Charles University, Prague, Czech Republic. [81]Department of Public Health, University of Helsinki, Helsinki, Finland. [82]Institute for Molecular Medicine Finland, Helsinki Institute of Life Science, University of Helsinki, Helsinki, Finland. [83]Institute of Public Health and Clinical Nutrition, Department of Clinical Nutrition, University of Eastern Finland, Kuopio, Finland. [84]Eating Disorders Unit, Department of Child and Adolescent Psychiatry, Medical University of Vienna, Vienna, Austria. [85]Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, the Netherlands. [86]Centre for Addiction and Mental Health, Toronto, Ontario, Canada. [87]Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada. [88]Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada. [89]Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK. [90]Department of Psychiatry, Seoul Paik Hospital, Inje University, Seoul, Korea. [91]Rheumatology Unit, Department of Medicine, Center for Molecular Medicine, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden. [92]Department of Psychology, Michigan State University, East Lansing, MI, USA. [93]Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway. [94]Department of Clinical Science, K.G. Jebsen Centre for Psychosis Research, Norwegian Centre for Mental Disorders Research (NORMENT), University of Bergen, Bergen, Norway. [95]Dr. Einar Martens Research Group for Biological Psychiatry, Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway. [96]Department of Clinical Medicine, Laboratory Building, Haukeland University Hospital, Bergen, Norway. [97]American School of Professional Psychology, Argosy University, Northern Virginia, Arlington, VA, USA. [98]Department of Cancer Epidemiology and Prevention, M Skłodowska-Curie Cancer Center - Oncology Center, Warsaw, Poland. [99]BESE Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. [100]Department of Psychiatry, University of Lausanne-University Hospital of Lausanne (UNIL-CHUV), Lausanne, Switzerland. [101]Department of Psychiatry, University of Campania "Luigi Vanvitelli", Naples, Italy. [102]Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland. [103]Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada. [104]Department of Psychiatry, University College Cork, Cork, Ireland. [105]HSE National Clinical Programme for Eating Disorders, Cork, Ireland. [106]Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. [107]Department of Biomedical Data Science, Leiden University Medical Centre, Leiden, the Netherlands. [108]Department of Psychiatry, Faculty of Medicine, University of Geneva, Geneva, Switzerland. [109]Division of Child and Adolescent Psychiatry, Geneva University Hospital, Geneva, Switzerland. [110]Great Ormond Street Institute of Child Health, University College London, London, UK. [111]National Center for PTSD, VA Boston Healthcare System, Department of Psychiatry, Boston University School of Medicine, Boston, MA, USA. [112]Department of Medicine, Surgery and Dentistry "Scuola Medica Salernitana", University of Salerno, Salerno, Italy. [113]Department of Neuroscience, Psychology, Drug Research and Child Health (NEUROFARBA), University of Florence, Florence, Italy. [114]Kartini Clinic, Portland, OR, USA. [115]Center for Neurobehavioral Genetics, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA, USA. [116]Center for Human Genome Research at the Massachusetts General Hospital, Boston, MA, USA. [117]Saint Joan de Déu Research Institute, Saint Joan de Déu Barcelona Children's Hospital, Barcelona, Spain. [118]Institute of Biomedicine (IBUB), University of Barcelona, Barcelona, Spain. [119]Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Spain. [120]Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [121]Department of Health Science, University of Florence, Florence, Italy. [122]Institute for Molecular Medicine Finland (FIMM), HiLIFE Unit, University of Helsinki, Helsinki, Finland. [123]Eating Disorders Research and Treatment Center, Department of Child and Adolescent Psychiatry, Faculty of Medicine, Technische Universität Dresden, Dresden, Germany. [124]Department of Psychological Medicine, University of Otago, Christchurch, New Zealand. [125]Faculty of Medicine & Health Sciences, University of Aukland, Aukland, New Zealand. [126]Department of Psychiatry, Neurobiology, Pharmacology, and Biotechnologies, University of Pisa, Pisa, Italy. [127]Department of Psychiatry, Poznan University of Medical Sciences, Poznan, Poland. [128]Department of Neurosciences, Padua Neuroscience Center, University of Padova, Padova, Italy. [129]Institute of Medical Statistics, Computer and Data Sciences, Jena University Hospital, Jena, Germany. [130]Department of Genetics and Genomic Biology, The Hospital for Sick Children, Toronto, Ontario, Canada. [131]Institute of Psychiatry, Psychology and Neuroscience, Department of Psychological Medicine, King's College London, London, UK. [132]J. Craig Venter Institute (JCVI), La Jolla, CA, USA. [133]Department of Psychiatry and Psychotherapy, Medical University of Vienna, Vienna, Austria. [134]Department of Pediatrics and Center of Applied Genomics, First Faculty of Medicine, Charles University, Prague, Czech Republic. [135]Center for Eating Disorders Ursula, Rivierduinen, Leiden, the Netherlands. [136]Department of Psychiatry, Leiden University Medical Centre, Leiden, the Netherlands. [137]Department of Child and Adolescent Psychiatry, Poznan University of Medical Sciences, Poznan, Poland. [138]IRCSS Fondazione Don Carlo Gnocchi, Florence, Italy. [139]Department of Environmental Epidemiology, Nofer Institute of Occupational Medicine, Lodz, Poland. [140]Department of Psychiatry, University of Naples SUN, Naples, Italy. [141]Department of Psychiatry, University of Perugia, Perugia, Italy. [142]Brain Sciences Department, Stremble Ventures, Limassol, Cyprus. [143]Adolescent Health Unit, Second Department of Pediatrics, "P. & A. Kyriakou" Children's Hospital, University of Athens, Athens, Greece. [144]Pediatric Intensive Care Unit, "P. & A. Kyriakou" Children's Hospital, University of Athens, Athens, Greece. [145]Faculty of Social and Behavioral Sciences, Utrecht University, Utrecht, the Netherlands. [146]Department of Internal Medicine VI, Psychosomatic Medicine and Psychotherapy, University Medical Hospital Tuebingen, Tuebingen, Germany. [147]BioRealm, LLC, Walnut, CA, USA. [148]Oregon Research Institute, Eugene, OR, USA. [149]Christchurch Health and Development Study, University of Otago, Christchurch, New Zealand. [150]The Center for Eating Disorders at Sheppard Pratt, Baltimore, MD, USA. [151]Department of Psychiatry, Weill Cornell Medical College, New York, NY, USA. [152]Eating Recovery Center, Denver, CO, USA. [153]Department of Psychiatry, University of California San Diego, San Diego, CA, USA. [154]Department of Psychiatry and Behavioral Science, University of North Dakota School of Medicine and Health Sciences, Fargo, ND, USA. [155]Biostatistics and Computational Biology Unit, University of Otago, Christchurch, New Zealand. [156]Department of Psychiatry and Biobehavioral Science, Semel Institute for Neuroscience and Human Behavior, University of California Los Angeles, Los Angeles, CA, USA. [157]David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA. [158]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark. [159]Centre for Mental Health, University Health Network, Toronto, Ontario, Canada. [160]Program for Eating Disorders, University Health Network, Toronto, Ontario, Canada. [161]Department of Cell Biology and Physiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [162]The Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH), Aarhus, Denmark. [163]Centre for Integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark. [164]Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark. [165]Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland, Australia. [166]National Centre for Register-Based Research, Aarhus BSS, Aarhus University, Aarhus, Denmark. [167]Centre for Integrated Register-based Research (CIRRAU), Aarhus University, Aarhus, Denmark. [168]Canterbury District Health Board, Christchurch, New Zealand. [169]Department of Pathology and Biomedical Science, University of Otago, Christchurch, New Zealand. [170]Queensland Brain Institute, University of Queensland, Brisbane, Queensland, Australia. [171]School of Psychology, Flinders University, Adelaide, South Australia, Australia. [172]Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden. [173]Department of Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [173]These authors contributed equally: Zeynep Yilmaz, Laura M. Thornton, Christopher Hübel, Jonathan R. I. Coleman, Mikael Landén, Nicholas G. Martin, Preben Bo Mortensen, Patrick F. Sullivan. [174]These authors jointly supervised this work: Gerome Breen, Cynthia M. Bulik. *e-mail: cynthia_bulik@med.unc.edu

Methods

Samples and study design. Thirty-three datasets with 16,992 cases of anorexia nervosa and 55,525 controls were included in the primary GWAS. We included individuals from the PGC-ED Freeze 1[10], newly collected samples from the ANGI[8,9], archived samples from the GCAN/WTCCC3[13], samples from cases of anorexia nervosa from the UK Biobank[14], and additional controls from Poland. Case definitions established a lifetime diagnosis of anorexia nervosa via hospital or register records, structured clinical interviews, or online questionnaires based on standardized criteria (Diagnostic and Statistical Manual of Mental Disorders (DSM) III-R, DSM-IV, International Classification of Diseases (ICD) 8, ICD-9 or ICD-10), whereas in the UK Biobank, cases self-reported a diagnosis of anorexia nervosa. Controls were carefully matched for ancestry, and some, but not all, control cohorts were screened for lifetime eating and/or some or all psychiatric disorders. Given the relative rarity of anorexia nervosa, large unscreened control cohorts were deemed appropriate for inclusion[33].

The cohorts are described in the Supplementary Note. Ethical approvals and consent forms were reviewed and archived for all participating cohorts (see Supplementary Note for Danish methods). Summary details about ascertainment (Supplementary Table 2), the genotyping platforms used (Supplementary Table 3) and genotype availability (Supplementary Table 4) are provided.

Statistical analysis. Data processing and analysis were done on the Lisa Compute Cluster hosted by SURFsara (http://www.surfsara.nl) and the GenomeDK high-performance computing cluster (http://genome.au.dk).

Meta-analysis of genome-wide association data. Quality control, imputation, GWAS and meta-analysis followed the standardized pipeline of the PGC, Ricopili (Rapid Imputation Consortium Pipeline). Ricopili versions used were 2017_Oct_11.002 and 2017_Nov_30.003. Quality control included SNP and sample quality control, population stratification and ancestry outliers, and familial and cryptic relatedness. Further information about the Ricopili pipeline is available from the website (https://sites.google.com/a/broadinstitute.org/ricopili) and GitHub repository (https://github.com/Nealelab/ricopili/tree/master/rp_bin). Further details of the quality control procedures can be found in the Supplementary Note.

Imputation. Imputation of SNPs and insertions–deletions was based on the 1000 Genomes Phase 3 (http://www.internationalgenome.org) data[34].

GWAS. GWASs were conducted separately for each cohort using imputed variant dosages and an additive model. Covariates nominally associated with the phenotype in univariate analysis ($P<0.05$) and five ancestry principal components were included in the GWAS (Supplementary Table 18). These analyses used the tests and methods programmed in the Ricopili pipeline. To the extent that national laws and regulations permitted, we examined sample overlap across cohorts by performing LD score bivariate regressions and estimating genetic covariance intercepts to assess sample overlap[19,20] (Supplementary Table 19). Genomic inflation factors (λ) of the final datasets indicated no evidence of inflation of the test statistics due to population stratification or other sources (Supplementary Table 1). The 33 cohorts were meta-analyzed with the Ricopili pipeline, which uses an inverse-variance weighted fixed-effect model. We filtered our GWAS results with MAF ≥ 0.01 and INFO score ≥ 0.70 (indicating 'high quality').

Analysis of chromosome X. Several cohorts in the primary GWAS did not have X chromosome variant data, specifically, some GCAN-based cohorts (fre1, ukd1, usa1, gns2) and were excluded. Imputation was performed separately from the autosomes[35]. Chromosome X variants in the pseudoautosomal regions were excluded before imputation. SNPs exceeding MAF and INFO score thresholds of 0.01 and 0.70 were retained and analysis was performed with PLINK v1.9 (https://www.cog-genomics.org/plink2) and Ricopili.

Female-only GWAS. A supplementary GWAS analysis was conducted on females only to determine the similarity of the results to the primary GWAS analysis, which included both females and males. The cohorts that did not have chromosome X variants to verify sex could not be included (fre1, ukd1, usa1, gns2).

Distance- and LD-based clumping. The GWAS results implicate genomic regions (loci). To define a locus, first SNPs that met the genome-wide significant threshold of $P<5\times10^{-8}$ were identified. Second, clumping was used to convert significant SNPs to regions. The SNP with the smallest P value in a genomic window was kept as the index SNP and SNPs in high LD with the index SNP defined the left and right end of the region (SNPs with $P<0.0001$ and $r^2>0.1$ within 3-Mb windows). Third, partially or wholly overlapping clumps within 50 kb were identified and merged into one region. Fourth, only loci with additional evidence of association from variants in high LD as depicted by regional plots were retained; furthermore, forest plots needed to confirm the associations based on the majority of cohorts. Finally, conditional analyses were conducted to identify SNPs with associations independent of the top SNP within the genomic section of interest.

Annotation. Genome-wide significant loci were annotated with RegionAnnotator (https://github.com/ivankosmos/RegionAnnotator) to identify known protein-coding genes within loci (Supplementary Table 6).

Conditional and joint analyses. Conditional and joint analyses were conducted using GCTA-COJO[15]. GCTA-COJO investigates every locus with a joint combination of independent markers using a genome-wide SNP selection procedure. It takes into account the LD correlations between SNPs and runs a conditional and joint analysis on the basis of conditional P values. After a model optimizing process, the joint effects of all selected SNPs are calculated. The largest subsample from our GWAS (sedk) was used to approximate the underlying LD structure of the investigated lead SNPs. The conditional regression was performed in a stepwise manner using the GCTA software[36]. We analyzed SNPs that had $P<5\times10^{-8}$ (Supplementary Table 5).

Multi-trait-based conditional and joint analyses. To separate marginal effects from conditional effects (that is, the effect of a risk factor on an outcome controlling for the effect of another risk factor), we performed a multi-trait-based conditional and joint analysis (GCTA-mtCOJO)[18] using an extension of the GCTA software (http://cnsgenomics.com/software/gcta)[36] (Supplementary Table 8). This method uses summary-level data to perform the conditional analysis. We conditioned the results of our anorexia nervosa GWAS on GWAS results for education years[37], type 2 diabetes[38], HDL cholesterol[39], BMI (C.H. et al., manuscript in preparation), schizophrenia[40] and neuroticism[41]. We again used the individual-level genotype data from our largest cohort (sedk) to approximate the underlying LD structure. As a first step, the method performs a generalized summary data-based Mendelian randomization (GSMR) analysis (http://cnsgenomics.com/software/gsmr) to test for causal association between the outcome (that is, anorexia nervosa) and the risk factor (for example, schizophrenia). We removed potentially pleiotropic SNPs from this analysis by the heterogeneity in dependent instruments (HEIDI) outlier method[18]. Pleiotropy is the phenomenon when a single locus directly affects several phenotypes. The power of the HEIDI outlier method is dependent on the sample size of the GWAS. Pleiotropic SNPs are defined as the SNPs that show an effect on the outcome that significantly diverges from that expected under a causal model. Second, the GCTA-mtCOJO method calculates the genetic correlation between the exposure and the outcome using LDSC (https://github.com/bulik/ldsc) to adjust for genetic overlap[19,20]. It also uses the intercept of the bivariate LDSC to account for potential sample overlap[19,20]. As a result, GCTA-mtCOJO calculates conditional betas, conditional standard errors and conditional P values. Subsequently, we clumped the conditional GWAS results using the standard PLINK v.1.9[42] algorithm (SNPs with $P<0.0001$ and $r^2>0.1$ within 3-Mb windows) to investigate whether any of the genome-wide significant loci showed dependency on genetic variation associated with other phenotypes. As described previously[18], the GCTA-mtCOJO analysis requires the estimates of b_{xy} of the covariate risk factors on the target risk factor and disease, r_g of the covariate risk factors, heritability (h^2_{SNP}) for the covariate risk factors and the sampling covariance between SNP effects estimated from potentially overlapping samples.

eQTL and chromosome conformation capture (Hi-C) interactions. Although GWAS findings are informative genome-wide, identifying strong hypotheses about their connections to specific genes is not straightforward. The lack of direct connections to genes constrains subsequent experimental modeling and efforts to develop improved therapeutics. Genomic location is often used to connect significant SNPs to genes, but this is problematic because GWAS loci usually contain many correlated and highly significant SNP associations over hundreds of kilobases. Moreover, the three-dimensional arrangement of chromosomes in cell nuclei enables regulatory interactions between genomic regions that are located far apart[43]. Chromosome conformation capture methods, such as Hi-C, enable identification of three-dimensional interactions in vivo[44,45] and can clarify GWAS findings. For example, an intergenic region associated with multiple cancers was shown to be an enhancer for *MYC* through a long-range chromatin loop[46,47], intronic *FTO* variants are robustly associated with body mass but influence expression of distal genes through long-range interactions[48], and Hi-C was used previously[49] to assess the three-dimensional chromatin interactome in fetal brain and connections of some schizophrenia associations to specific genes were found in the study.

To gain a better understanding of the three-dimensional organization of chromatin in the brain and to evaluate disease relevance, we applied Hi-C[50] to post-mortem samples ($n=3$ samples of the adult temporal cortex). Details on methodology, data processing, quality control and statistical models used for these analyses have been published elsewhere[51]. We generated sufficient reads to enable a kilobase-resolution map of the chromatin interactome from adult humans. We generated tissue RNA-sequencing, total-stranded RNA-sequencing, chromatin immunprecipitation followed by sequencing (H3K27ac, H3K4me3 and CTCF) and open chromatin data (assay for transpose-accessible chromatin using sequencing; ATAC-seq) for the adult brain to help to interpret the Hi-C results. We also integrated brain expression and eQTL data from GTEx to aid these analyses. The Hi-C analysis is unbiased in that all chromatin interactions that pass a confidence threshold are considered when evaluating the associations between SNPs and genes (that is, it is not a capture experiment where only candidate SNP-to-gene associations are evaluated).

Similar to a previous study[49], we used Hi-C data generated from human adult brain to identify genes implicated by three-dimensional functional interactomics

(Supplementary Fig. 5). These Hi-C data ($n = 3$, anterior temporal cortex) contain more than 103,000 high-confidence, regulatory chromatin interactions[51]. These interactions capture the physical proximity of two regions of the genome in brain nuclei (anchors, 10 kb resolution), although they are separated by 20 kb to 2 Mb in genomic distance. We focused on the regulatory subset of E–P or P–P (E, enhancer; P, promoter) chromatin interactions (with P defined by the location of an open chromatin anchor near the transcription start site of an adult brain-expressed transcript and E defined by the overlap with open chromatin in adult brain plus either H3K27ac or H3K4me3 histone marks). The presence of a regulatory chromatin interaction from a GWAS locus to a gene provides a strong hypothesis about SNP-to-gene regulatory functional interactions.

SNP-based heritability estimation. LDSC software (https://github.com/bulik/ldsc) and methods were used to estimate SNP-based heritabilities for each cohort and overall[19,20]. We used precomputed LD scores based on the 1000 Genomes Project European ancestry samples[34] (directly downloaded from https://github.com/bulik/ldsc). The liability scale estimate assumed a population prevalence of 0.9–4% for anorexia nervosa[2,3].

Polygenic risk scoring for within-trait predictions. Polygenic leave-one-dataset-out analysis, using PRSice v2.1.3[52], was conducted in the first instance to identify any extreme outlying datasets. In addition, it enabled the evaluation of the association between anorexia nervosa PRS and anorexia nervosa risk in an independent cohort as a means of replication of the GWAS results. We derived a PRS for anorexia nervosa from the meta-analysis of all datasets except for the target cohort, and then applied the PRS to the target cohort to predict affected status (Supplementary Fig. 16). Logistic regression was performed, including as covariates the first five ancestry components and any other principal components that were significantly associated with the phenotype in the target cohort, and the target cohort was split into deciles based on anorexia nervosa PRS, for which decile 1, which consisted of those with the lowest anorexia nervosa PRS, served as the reference.

Anorexia nervosa subtype analysis. PRS analyses were conducted with anorexia nervosa subgroups to investigate prediction of case status across the subtypes. For this, we split the cases of anorexia nervosa into two groups based on whether binge eating was present. First, GWAS meta-analyses were conducted for anorexia nervosa with binge eating compared to controls (2,381 cases and 10,249 controls; $k = 3$ datasets: aunz, chop, usa2) and anorexia nervosa with no binge eating compared to controls (2,262 cases and 10,254 controls; $k = 3$ datasets: aunz, chop, usa2). Controls were randomly split between analyses to maintain independence (Supplementary Fig. 6). Genetic correlation analysis using LDSC[19,20] was conducted to examine the potential genetic overlap of the two anorexia nervosa subtypes (Supplementary Table 9). Second, using PRSice[52], we calculated PRS for each anorexia nervosa subtype separately in the three target cohorts for which anorexia nervosa subtype data were available. Finally, mean PRS scores were estimated for each subtype by cohort after accounting for covariates in R. Subtype phenotyping is described in the Supplementary Note.

Males. To assess whether sex-specific differences in genetic risk load exist for anorexia nervosa, we calculated PRS, using PRSice[52], from a GWAS meta-analysis performed on females only (14,898 cases and 27,545 controls) and applied it to a male-only target cohort (447 cases and 20,347 controls) to predict affected status.

Genetic correlations in the cross-trait analysis. Common variant-based genetic correlation (SNP-r_g) analysis measures the extent to which two traits or disorders share common genetic variation. SNP-r_g between anorexia nervosa and 447 traits (422 from an internally curated dataset and 25 from LDHub[53]) were tested using GWAS summary statistics using an analytical extension of LDSC[19,20]. The sources of the summary statistics files (PMID, DOI or unpublished results) used in the LDSC are provided in Supplementary Table 10. When there were multiple summary statistics files available for a trait, significant SNP-r_g reported in the main text were chosen based on the largest sample size and/or matching ancestry with our sample (that is, European ancestry).

Genetic correlations with anorexia nervosa corrected for BMI were carried out to investigate whether the observed genetic correlations between anorexia nervosa and metabolic phenotypes were attributable to BMI or partially independent. We used GCTA-mtCOJO[18] to perform a GWAS analysis for anorexia nervosa conditioning on BMI using BMI summary data from our UK Biobank analysis (described in the next section) to derive anorexia nervosa GWAS summary statistics corrected for the common variants genetic component of BMI (Supplementary Tables 14, 15).

GWAS of related traits in UK Biobank. Several GWAS analyses were carried out for traits using data from the UK Biobank to allow us to investigate body composition genetics in healthy individuals without a psychiatric or weight-altering disorder or individuals who were taking weight-altering medication. We also used UK Biobank data to carry out GWAS of physical activity level, anxiety and neuroticism (Supplementary Table 20). For details, see the Supplementary Note.

GSMR analyses. We performed two bidirectional GSMR analyses[18] to test for the causal association first between BMI and anorexia nervosa, and second between type 2 diabetes and anorexia nervosa, using an extension of the GCTA software[36] (Supplementary Table 16). We used the individual-level genotype data from our largest cohort (sedk) to approximate the underlying LD structure. We removed potentially pleiotropic SNPs from this analysis by the HEIDI outlier method[18]. Pleiotropic SNPs are defined as the SNPs which show an effect on the outcome that significantly diverges from the one expected under a causal model. The method uses the intercept of the bivariate LD score regression to account for potential sample overlap[19,20]. As a rule of thumb, GSMR requires GWAS to have at least ten genome-wide significant hits. We lowered the threshold for this requirement to eight SNPs in our analyses of anorexia nervosa as an exposure and BMI or type 2 diabetes as an outcome. Results, therefore, should be interpreted with caution. Moreover, we investigated bidirectional conditional effects between BMI or type 2 diabetes and anorexia nervosa. We used GCTA-mtCOJO to perform a GWAS analysis for anorexia nervosa conditioning on BMI using summary data from our UK Biobank analysis or type 2 diabetes using summary data[38]. Our anorexia nervosa GWAS and the BMI and type 2 diabetes GWAS analyses are based on independent samples. For BMI, we also reran the GSMR analysis using the BMI-adjusted anorexia nervosa GWAS summary data from the GCTA-mtCOJO analysis.

Gene-wise analysis. MAGMA v.1.06[54] (http://ctg.cncr.nl/software/magma) was used to perform a gene-wise test of association with anorexia nervosa based on GWAS summary statistics. MAGMA generates gene-based *P* values by combining SNP-based *P* values within a gene while accounting for LD. To include regulatory regions, SNPs are mapped to genes within a 35-kb upstream and 10-kb downstream window, and the gene *P* value is obtained using the multi = snp-wise model, which aggregates mean and top SNP association models. We tested 19,846 ENSEMBL genes, including the X chromosome (Supplementary Table 11). As a reference panel for the underlying LD structure, we used 1000 Genomes European data phase 3[34].

Pathway analysis. MAGMA v.1.06[54] was used to perform a competitive pathway analysis, testing whether genes associated with anorexia nervosa were more enriched in a given pathway than all other pathways. The analysis included chromosome X. Biological pathways were defined using gene ontology pathways and canonical pathways from MSigDB v.6.1[55], and psychiatric pathways mined from the literature. A total of 7,268 pathways were tested (Supplementary Table 12).

Partitioned heritability. Partitioned heritability was investigated using stratified LDSC[26], which estimates the per-SNP contribution to overall SNP-heritability (SNP-h^2) across various functional annotation categories of the genome (Supplementary Fig. 7). It accounts for linked markers and uses a 'full baseline model' of 24 annotations that are not specific to any cell type. We excluded the MHC region in our analysis. SNP-h^2 can be partitioned in two different ways: a non-cell type-specific and a cell type-specific manner. Partitioned heritability analysis was used to test for cell type-specific enrichment in the GWAS of anorexia nervosa among 10 cell type groups: adrenal tissue and pancreas, cardiovascular tissue, central nervous system, connective tissue and bone, gastrointestinal tissue, immune and hematopoietic tissues, kidney, liver, skeletal muscle and other tissues, which includes adipose tissue (Supplementary Fig. 8).

Gene expression. We conducted a series of gene expression analyses as described in the Supplementary Note.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The policy of the PGC is to make genome-wide summary results public. Genome-wide summary statistics for the meta-analysis are freely downloadable from the website of the PGC (http://www.med.unc.edu/pgc/results-and-downloads). Individual-level data are deposited in dbGaP (http://www.ncbi.nlm.nih.gov/gap) for ANGI-ANZ/SE/US (accession number phs001541.v1.p1) and CHOP/PFCG (accession number phs000679.v1.p1). ANGI-DK individual-level data are not available in dbGaP owing to Danish laws, but are available through collaboration with principal investigators of the Danish institutions. GCAN/WTCCC3 individual-level data are deposited in EGA (https://www.ebi.ac.uk/ega) (accession number EGAS00001000913) with the exception of the Netherlands and USA/Canada; data from these countries are available through collaboration with principal investigators of institutions in these countries. UK Biobank individual-level data can be applied for on the UK Biobank website (http://www.ukbiobank.ac.uk/register-apply).

References

33. Moskvina, V., Holmans, P., Schmidt, K. M. & Craddock, N. Design of case–controls studies with unscreened controls. *Ann. Hum. Genet.* **69**, 566–576 (2005).

34. The 1000 Genomes Project Consortium An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
35. Chang, D. et al. Accounting for eXentricities: analysis of the X chromosome in GWAS reveals X-linked genes implicated in autoimmune diseases. *PLoS ONE* **9**, e113684 (2014).
36. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
37. Okbay, A. et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539–542 (2016).
38. The DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet*. **44**, 981–990 (2012)..
39. Teslovich, T. M. et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).
40. Schizophrenia Working Group of the Psychiatric Genomics Consortium Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
41. Hübel, C. et al. Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *Am. J. Med. Genet. B* https://doi.org/10.1002/ajmg.b.32709 (2018).
42. Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
43. Dekker, J. Mapping the 3D genome: aiming for consilience. *Nat. Rev. Mol. Cell Biol.* **17**, 741–742 (2016).
44. Dekker, J. Gene regulation in the third dimension. *Science* **319**, 1793–1794 (2008).
45. Ethier, S. D., Miura, H. & Dostie, J. Discovering genome regulation with 3C and 3C-related technologies. *Biochim. Biophys. Acta* **1819**, 401–410 (2012).
46. Pomerantz, M. M. et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with *MYC* in colorectal cancer. *Nat. Genet.* **41**, 882–884 (2009).
47. Wright, J. B., Brown, S. J. & Cole, M. D. Upregulation of c-MYC in *cis* through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells. *Mol. Cell Biol.* **30**, 1411–1420 (2010).
48. Smemo, S. et al. Obesity-associated variants within *FTO* form long-range functional connections with *IRX3*. *Nature* **507**, 371–375 (2014).
49. Won, H. et al. Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**, 523–527 (2016).
50. Lu, L., Liu, X., Peng, J., Li, Y. & Jin, F. Easy Hi-C: a simple efficient protocol for 3D genome mapping in small cell populations. Preprint at *bioRxiv* https://doi.org/10.1101/245688 (2018).
51. Giusti-Rodriguez, P. M. D. & Sullivan, P. F. Using three-dimensional regulatory chromatin interactions from adult and fetal cortex to interpret genetic results for psychiatric disorders and cognitive traits. Preprint at *bioRxiv* https://doi.org/10.1101/406330 (2018).
52. Euesden, J., Lewis, C. M. & O'Reilly, P. F. PRSice: Polygenic Risk Score software. *Bioinformatics* **31**, 1466–1468 (2015).
53. Zheng, J. et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* **33**, 272–279 (2017).
54. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
55. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

nature research

Corresponding author(s): Cynthia Bulik

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	No software was used to compile the code. All code is available upon reasonable requests from the analysts and senior author.
Data analysis	The manuscript includes information on all software and versions used. All software used is publicly available, including Plink (https://www.coggenomics.org/plink2/), Ricopili (https://sites.google.com/a/broadinstitute.org/ricopili/download), LDSC (https://github.com/bulik/ldsc), METAL (http://csg.sph.umich.edu/abecasis/metal/download/), R (https://www.r-project.org/), MAGMA (https://ctg.cncr.nl/software/magma), S-PrediXcan (https://github.com/hakyimlab/PrediXcan), GCTA (https://cnsgenomics.com/software/gcta/#Download), PRSice (http://prsice.info/), and GSMR (http://cnsgenomics.com/software/gsmr/).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The Psychiatric Genomics Consortium's (PGC) policy is to make genome-wide summary results public. Genome-wide summary statistics for the meta-analysis are freely downloadable from PGCs download website (http://www.med.unc.edu/pgc/results-and-downloads). Individual-level data are deposited in dbGaP (http://www.ncbi.nlm.nih.gov/gap) for ANGI-ANZ/SE/US (accession number phs001541.v1.p1) and CHOP/PFCG (accession number phs000679.v1.p1). ANGI-DK individual-level data are not available in dbGaP owing to Danish laws, but are available via collaboration with PIs. GCAN/WTCCC3 individual-level data are deposited in EGA (https://www.ebi.ac.uk/ega) (accession number EGAS00001000913) with the exception of Netherlands and US/Canada, which are available via collaboration with PIs. UK Biobank individual-level data can be applied for on the UK Biobank website (http://www.ukbiobank.ac.uk/register-apply).

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Genome-wide association study meta-analysis
Research sample	Cohorts from the Psychiatric Genomics Consortium (PGC), cohorts collected as part of the Anorexia Nervosa Genetics Initiative (bioRxiv, doi: https://doi.org/10.1101/234013), and the UK Biobank cohort. These are all case-control cohorts where cases have anorexia nervosa diagnoses. Further detailed information about the studies and cohorts included is provided in the Supplementary Material, Tables S1-S4. In total, 33 case-control cohorts were studied. Case definitions established a lifetime diagnosis of anorexia nervosa via hospital or register records, structured clinical interviews, or on-line questionnaires based on standardized criteria (DSM-III-R, DSM-IV, ICD-8, ICD-9, or ICD-10), whereas in the UK Biobank cases self-reported a diagnosis of anorexia nervosa. Controls were carefully matched for ancestry, and some, but not all control cohorts were screened for lifetime eating and/or some or all psychiatric disorders.
Sampling strategy	In the last published GWAS of anorexia nervosa (Psychiatric Genomics Consortium Freeze 1), one significant loci was detected (Duncan et al. 2016). Power analyses in psychiatric genetics have shown that once there are enough samples to detect the first significant loci, an inflection point is attained such that with increasing sample size more loci will become detectable. We quadrupled the number of cases in this GWAS and were confident we would be able to identify further genome-wide significant loci ($p < 5e08$)
Data collection	Phenotype data were ascertained through a variety of methods, through hospital or register records, structured clinical interviews, or self-report questionnaires. Information about sample ascertainment for each of the 33 cohorts included in the study are given in the Supplementary Material and Table S2. Phenotype data from the Genetic Consortium for Anorexia Nervosa/Wellcome Trust Case Control Consortium-3 (GCAN/WTCCC3) samples were from: the Structured Clinical Interview for DSM-IV Disorders (SCID), Child Semi-Structured Assessment for the Genetics of Alcoholism, Adolescent version (C-SSAGA-A), Diagnostic Interview for Genetic Studies (DIGS), Mini International Neuropsychiatric Interview (MINI), Structured Interview for Anorexic and Bulimic Disorders (SIAB), Eating Disorder Examination (EDE), Eating Disorder Examination Questionnaire (EDEQ), Composite International Diagnostic Interview (CIDI), Structured Interview for Anorexic and Bulimic Syndromes for DSM-IV and ICD-10 (SIAB-EX), EATATE Lifetime Diagnostic Interview, and/or chart review with algorithms. Phenotype data for the Anorexia Nervosa Genetics Initiative (ANGI) samples were from: the ED100K diagnostic questionnaire, clinic/register diagnosis, disease checklist, and/or a questionnaire harmonized to the ED100K; and phenotype data for UK Biobank sample were from the self-report Mental Health Questionnaire. Samples were predominantly genotyped on the Illumina GSA chip. Non-typed genetic variants were imputed to the 1000 Genomes Phase 3 reference.
Timing	Anorexia Nervosa Genetics Initiative (ANGI) study recruitment for Australia/New Zealand and Sweden was between 2013-2016. Samples for ANGI Denmark came from a national register (birth years 1981-2005). UK Biobank recruited their participants between 2006-2010. The other cohorts were obtained from the Psychiatric Genomics Consortium and have been used in previous studies. These cohorts are derived from many different study sources (shown in Table S2) and precise recruitment start/end times are not known by us.
Data exclusions	Samples were excluded from the GWAS with pre-established criteria if they were dropped during routine GWAS quality control (i.e. call rate < 98%, heterozygosity inbreeding coefficient <0.2), if they were not of European ancestry, failed sex checks, or showed cryptic relatedness. These methods are all outlined in our Methods and Supplementary material. We present raw (pre-QC) sample sizes for each of the 33 cohorts, and all post-QC sample sizes in Table S1. 883 cases (4.9%) and 8,629 (13.4%) controls were excluded.
Non-participation	Non-participation is not relevant to this study because data used were de-identified and involved secondary analysis of samples collected from primary studies. In other words, there was no direct contact with participants.

Randomization	Randomization is not applicable since this was a case-control study design.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	This study is a secondary analysis of phenotype and genotype data obtained from contributing studies and researchers. No human participants were directly recruited or contacted for the present study. The cases had a lifetime diagnosis of anorexia nervosa (see Table S2 for ascertainment characteristics) and the controls had no diagnosis of anorexia nervosa. Details about each cohort can be obtained from the primary study source. In Table S2 we list the PMIDs of each study that contributed genotypes. Ascertainment information is provided in the Supplementary Material. The covariates were genomic (ancestry) principal components generated through principal components analysis. 33 cohorts were which included in the present study, with samples collected from 17 countries.
Recruitment	This study represents secondary data analysis of samples already collected in other studies. Recruitment and sample ascertainment information for the cohorts is provided in Table S2 and the Supplementary Material. We provide detailed information in the Supplement for the newly added cohorts that are in this current data freeze (Freeze 2 of the Eating Disorders Working Group of the Psychiatric Genomics Consortium) and direct the reader to the Freeze 1 publication (Duncan et al., 2017) for information on how the previously analyzed cohorts were obtained. Briefly, the original studies used methods including national registers, hospital records, clinic-based recruitment, community-based recruitment, and biobank collections.