Huckins, L. M., Hatzikotoulas, K., Southam, L., Thornton, L. M., Steinberg, J., Aguilera-McKay, F., ... Eating Disorder Working Group of the Psychiatric Genomics Consortium (2017). Investigation of common, low-frequency and rare genome-wide variation in anorexia nervosa. *Molecular Psychiatry*, 23(5), 1169-1180. https://doi.org/10.1038/mp.2017.88

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
10.1038/mp.2017.88

Link to publication record in Explore Bristol Research
PDF-document

This is the final published version of the article (version of record). It first appeared online via NATURE at http://www.nature.com/articles/mp201788#rightslink. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms
Anorexia nervosa (AN) is a complex neuropsychiatric disorder presenting with dangerously low body weight, and a deep and persistent fear of gaining weight. To date, only one genome-wide significant locus associated with AN has been identified. We performed an exome-chip based genome-wide association studies (GWAS) in 2158 cases from nine populations of European origin and 15 485 ancestrally matched controls. Unlike previous studies, this GWAS also probed association in low-frequency and rare variants. Sixteen independent variants were taken forward for in silico and de novo replication (11 common and 5 rare). No findings reached genome-wide significance. Two notable common variants were identified: rs10791286, an intronic variant in OPCML ($P = 9.89 \times 10^{-6}$), and rs7700147, an intergenic variant ($P = 2.93 \times 10^{-5}$). No low-frequency variant associations were identified at genome-wide significance, although the study was well-powered to detect low-frequency variants with large effect sizes, suggesting that there may be no AN loci in this genomic search space with large effect sizes.

Molecular Psychiatry (2018) 23, 1169–1180; doi:10.1038/mp.2017.88; published online 25 July 2017

INTRODUCTION

Family studies of anorexia nervosa (AN) have consistently shown that first-degree relatives of AN sufferers have an increased risk of AN, compared with relatives of unaffected individuals.1–4 Twin studies have estimated the heritability of AN at 56%,5 with the majority of remaining variance in liability attributed to non-shared environmental factors (38%).5

Three genome-wide association studies (GWAS) of AN have been conducted to date. The first comprised 1033 AN cases collected as part of the Price Foundation Genetic Study of

1Department of Human Genetics, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK; 2Division of Psychiatric Genomics, Icahn School of Medicine at Mount Sinai, New York, NY, USA; 3Department of Psychiatry and Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; 4Section of Eating Disorders, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, UK; 5NIHR BRC SLaM BioResource for Mental Health, SGDP Centre & Centre for Neuroimaging Sciences, Section of Eating Disorders, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, UK; 6Klinik für Kinder- und Jugendspsychiatrie, Psychotherapie und Psychosomatik Klinikum Frankfurt, Frankfurt, Germany; 7Division of Psychological and Social Medicine and Developmental Neurosciences, Faculty of Medicine, TU Dresden, Dresden, Germany; 8Eating Disorders Research and Treatment Center, Department of Child and Adolescent Psychiatry, Faculty of Medicine, TU Dresden, Dresden, Germany; 9Department of Child and Adolescent Psychiatry and Psychotherapy, University Hospital Essen, University of Duisburg-Essen, Essen, Germany; 10Helmholtz Zentrum München, Deutsches Forschungszentrum für Gesundheit und Umwelt, Neuherberg, Germany; 11Division of Psychological & Social Medicine and Developmental Neurosciences, Technische Universität Dresden, Faculty of Medicine, University Hospital C.G. Carus, Dresden, Germany; 12Department of Psychology, Georgia State University, Atlanta, GA, USA; 13Erasmus University Medical Center, Rotterdam, The Netherlands; 14Center for Human Genome Research at the Massachusetts General Hospital, Boston, MA, USA; 15Department of Public Health & Institute for Molecular Medicine FIMM, University of Helsinki, Helsinki, Finland; 16Brain Center Rudolf Magnus, Department of Neuroscience and Pharmacology, University Medical Center Utrecht, Utrecht, The Netherlands; 17INSERM U984, Centre of Psychiatry and Neuroscience, Paris, France; 18Mary Sue Coleman Director, Life Sciences Institute, Professor of Molecular and Integrative Physiology, University of Michigan, Ann Arbor, MI, USA; 19Department of Dietetics-Nutrition, Harokopio University, Athens, Greece; 20Rivierduinen Eating Disorders Ursula, Leiden, Zuid-Holland, The Netherlands; 21Eating Disorders Unit, 1st Department of Psychiatry, National and Kapodistrian University of Athens, Medical School, Athens, Greece; 22Department of Psychiatry, McLean Hospital/Harvard Medical School, Belmont, MA, USA; 23Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands; 24Department of Public Health, Clinicum, University of Helsinki, Helsinki, Finland; 25Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK; 26Health Data and Digitalisation, Norwegian Institute of Public Health, Oslo, Norway; 27Department of Psychiatry, University of Naples SUN, Naples, Italy; 28Department of Medicine and Surgery, Section of Neurosciences, University of Palermo, Palermo, Italy; 29Department of Genetics, Environment and Mental Health, Norwegian Institute of Public Health, Oslo, Norway; 30EHealth Lab-Computer Science Department, University of Cyprus, Nicosia, Cyprus; 31Adolescent Health Unit (A.H.U.), 2nd Department of Pediatrics – Medical School, University of Athens “P. & A. Kyriakou” Children’s Hospital, Athens, Greece; 32Center for Eating Disorders Rintveld, University of Utrecht, Utrecht, The Netherlands; 33Eli Lilly and Company, Erl Wood Manor, Windlesham, UK; 34Department of Genetics and Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; 35Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden and 36Social Genetic and Developmental Psychiatry, King’s College London, London, UK. Correspondence: Dr LM Huckins, Psychiatric Genomics, Icahn School of Medicine Mount Sinai, 1 Gustave Levy Place, New York City, NY 10029, USA or Dr CM Bulik, Departments of Psychiatry and Nutrition, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, USA or Dr E Zeggini, Department of Human Genetics, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. E-mail: laura.huckins@mssm.edu or cynthia_bulik@med.unc.edu or eleftheria@sanger.ac.uk

Joint first authors.

228A full list of consortium members appears before the References section.

229Joint last authors.

Received 3 August 2016; revised 16 February 2017; accepted 17 February 2017; published online 25 July 2017
Anorexia Nervosa and 3733 pediatric controls from the Children’s Hospital of Philadelphia. This study focused on common variation and identified 11 suggestive variants ($P < 1 \times 10^{-5}$). None reached genome-wide significance in the primary analysis, although one variant (rs4479806) approached genome-wide significance in an associated secondary analysis. The second study (comprising 2907 cases and 14 860 controls) was carried out by the Genetic Consortium for AN, as part of the Wellcome Trust Case Control Consortium 3 (WTCCC3) effort. This study identified two suggestively associated variants ($P < 1 \times 10^{-5}$). Notably, signals at $P = 4 \times 10^{-6}$ were significantly more likely to have the same direction of effect in the replication as in the discovery cohorts ($P = 4 \times 10^{-6}$), which implies that true signals exist within this data set, but that the study was underpowered for detection. Recently, a third study-meta-analyzed samples from both of these studies, as well as some novel cases, comprising a total of 3495 cases and 10 982 controls. To our knowledge, this study identified the first genome-wide significant locus for AN (index variant rs4622308, $P = 4.3 \times 10^{-9}$).

Both previous studies focused on common variation. Here, we conducted, to our knowledge, the first association study that also considered low frequency (minor allele frequency (MAF) < 5%) and rare exonic variants in addition to common variation.

MATERIALS AND METHODS

Sample collections

We conducted a GWAS across nine discovery data sets (the majority overlapping with Genetic Consortium for AN, as part of the Wellcome Trust Case Control Consortium 3 (WTCCC3) effort). This study identified two suggestively associated variants ($P < 1 \times 10^{-5}$). Notably, signals at $P < 1 \times 10^{-5}$ were significantly more likely to have the same direction of effect in the replication as in the discovery cohorts ($P = 4 \times 10^{-6}$), which implies that true signals exist within this data set, but that the study was underpowered for detection. Recently, a third study-meta-analyzed samples from both of these studies, as well as some novel cases, comprising a total of 3495 cases and 10 982 controls. To our knowledge, this study identified the first genome-wide significant locus for AN (index variant rs4622308, $P = 4.3 \times 10^{-9}$).

Both previous studies focused on common variation. Here, we conducted, to our knowledge, the first association study that also considered low frequency (minor allele frequency (MAF) < 5%) and rare exonic variants in addition to common variation.

Figure 1. Geographical distribution of samples across Europe. (a) Distribution of cases across Europe; 375 USA cases are not shown in this diagram. (b) Distribution of controls across Europe; 873 USA controls are not shown in this diagram.
Population prevalence of AN in these populations ranged from 0.4 to 3% (refs 12–18; Table 1).

Genotyping
Cases were genotyped on either the Infinium HumanCoreExome-12 BeadChip Kit (Illumina, San Diego, CA, USA),19 or the Infinium HumanCoreExome-24 BeadChip Kit (Illumina),20 at the Wellcome Trust Sanger Institute. Where possible, controls were selected from existing studies with matching genotyping platforms to cases. Three control cohorts had been genotyped on the Infinium HumanExome-12 BeadChip Kit (Table 1). To ameliorate potential confounding due to chip effects,21 chip-type quality control (QC) was carried out, and ~14 000 single-nucleotide polymorphisms (SNPs) removed.

Quality control
Genotypes were called using the GenCall22 and Zcall23 algorithms. At each of these genotype-calling stages, QC was performed for each population and for cases and controls separately (Supplementary Table 1). The final number of SNPs included in the analyses is given in Table 2.

Controlling for population stratification
In order to account for population stratification, a principal components analysis was carried out for each cohort separately using the smartpca software.24 Population outliers were identified by merging each population with central European 1000 Genomes data.25 Variance explained by each PC was plotted for each population. In order to be both conservative and consistent across populations, the first 10 principal components were included as covariates in the association testing.

Association testing
Unbalanced case–control ratios can lead to anticonservative P-value estimates.26 This study includes a number of unbalanced strata (Table 1). The likelihood ratio test has been shown to have low type-I error rate across both balanced and unbalanced cohorts,27 and was chosen as the association test for this study.

A lower cutoff of minor allele count of 5 and MAF of 0.1% was used. Association testing was performed for each cohort separately using SNPtest.27 In the cohorts with mixed sex controls (all except Italy and Norway), sex was also included as a covariate.

The standard genome-wide significance threshold of $P \leq 5 \times 10^{-8}$ was applied.

Meta-analysis
Summary statistics across cohort were meta-analyzed using an inverse variance-based test in METAL.28 In order to test the heterogeneity of the results, Cochran’s $Q$ and the $I^2$ statistic were computed.

Assigning variants to genes
Variants identified associated at $P \leq 1 \times 10^{-4}$ were assigned to genes using Ensembl (release 83; Ensembl Genome Browser).29,30 For each variant, all predicted consequences (for example, missense, non-synonymous, and so on) and associated gene transcripts were downloaded and compared. Each variant was associated with only one predicted consequence and one Ensembl gene ID (Ensembl Genome Browser).29

Cluster plot checking
Cluster plots were created for all SNPs reaching $P \leq 1 \times 10^{-4}$ in any analysis (cohort-specific or meta-analysis) using ScatterShot.31 SNPs were visually inspected for each cohort, and for cases and controls separately. In instances where multiple cohorts were merged (for example, UK cases), cluster plots were checked separately for each original cohort.

Burden testing
The potential aggregation of rare variants in cases compared with controls was investigated using a gene-based approach. Burden tests were carried out using the Zeggini–Morris burden test32 as implemented in rvtests (Rvtests - Genome Analysis Wiki).

All SNPs with MAF between 0.1 and 5% were included; similar to the single-point analysis, a lower bound of minor allele count = 5 was used. A list of genes and locations was obtained from the UCSC genome browser (Table Browser: www.genome.ucsc.edu). All genes with at least two qualifying variants in at least two populations were used, resulting in a total of 9083 genes.

Burden tests were carried out for each population individually, and the results meta-analyzed using Stouffer’s method, weighted according to effective sample size.33

The genome-wide significance threshold for burden testing is computed in a similar manner to that for single-point analysis, using Bonferroni correction for the number of genes tested. This result in a genome-wide significance threshold of $5.5 \times 10^{-6}$.

Pathway analysis
One of the key motivations of studying complex psychiatric disorders such as AN is the desire to unearth biological pathways underlying disease development. Pathway analysis was performed using summary statistics from the meta-analysis for the full data set.

Four pathway databases were used: the Kyoto Encyclopedia of Genes and Genomes (KEGG),34,35 the Reactome pathway database (REACTOME),36 PANTHER pathway (PANTHER),37,38 and the Gene Ontology database (GO).39,40 These were curated to remove redundancy, resulting in a total set of 1836 pathways.

The analysis was run once on a merged set of 235 KEGG,34,35 REACTOME36 and PANTHER37,38 pathways, and once for the 1601 GO pathways.39,40

Pathway analysis was carried out using MAGMA.41 MAGMA was selected for its ability to deal robustly with linkage disequilibrium (LD) between markers, correct for gene length and deal accurately with rare variants. To our knowledge, MAGMA was first used to annotate SNPs to genes. This analysis was repeated twice. In the first analysis, variants were assigned only to the gene they were in, resulting in 68.73% of the variants being assigned to 13 400 genes. In the second analysis, variants were assigned allowing a 20 kb window in both directions from the gene. This procedure included 75.44% of variants across 18 118 genes.

SNP P-values were used to create gene scores. The European panel of the 1000 Genomes project was used as a reference set to estimate LD between SNPs. The analysis also requires the sample size of the study to be specified; because of the unbalanced nature of the study, the effective sample sizes were given here.

Gene P-values were calculated using MAGMA.41 The top 10% of SNPs per gene were used. Significance was defined using a false discovery rate of $5\%$.42

There is a risk when assigning SNPs to genes using MAGMA that some highly associated SNP might be assigned to multiple overlapping genes, and thus distort pathway results. SNP–gene assignments were checked for all pathways that reached false discovery rate-corrected significance. No instances of SNPs being assigned to multiple genes were found across these pathways.
Replication
SNPs reaching \( P < 1 \times 10^{-4} \) in the discovery stage were prioritized for replication. In total, 16 SNPs were selected.

Replication was carried out using two data sets: one existing in silico data set and one for de novo genotyping. The in silico data set came from an existing GWAS of AN, genotyped on the Illumina HumanHap610 platform. This data set included 1033 cases and 3733 controls. All cases included in this study were female. Controls were both male and female. The de novo replication cohort consisted of 266 self-volunteered female UK cases, collected through the charity Charlotte’s Helix (www.charlotteshelix.net). All participants were adults and had been diagnosed with AN by their clinician. In addition, all participants completed an online questionnaire based on the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders-IV Module H. The Structured Clinical Interview has been used extensively in epidemiological investigations. The interview was used to capture information on lifetime history of eating disorders including AN, and includes questions on body mass index, age of onset, and experience of eating disorders. DNA from the saliva samples was extracted using standard protocols and was quantified using pico-green. Samples were genotyped on the Infinium HumanExome 12 Beadchip, genotypes were called using GenCall and Zcall algorithms and stringent QC was performed pre- and post-call.

Sequenom). Sample and SNP QC were carried out within each replication data set, using an 80% sample call rate and a 90% SNP call rate threshold, and a Hardy–Weinberg equilibrium threshold of \( 10^{-4} \). Five samples and one SNP were removed using these criteria.

Post-QC, 15 SNPs and 261 de novo cases remained. The de novo replication analysis therefore included 15 SNPs, 261 cases and 1500 controls. Genotypes for 12/16 SNPs were available in the in silico replication cohort, across 1033 in silico cases and 3733 controls.

Expression analysis
Gene expression data were obtained from the Genotype-Tissue Expression (GTEx project) web portal, data release version 6 (dbGap Accession phs000424.v6.p1).\(^{44,46}\)

Power
The sample sizes used in this study are small in the context of other psychiatric phenotypes. Power to identify genome-wide significant signals was calculated using Quanto.\(^{47,48}\) This study is adequately powered to detect low-frequency alleles with large effect sizes and common alleles with substantial effect sizes (80% power to detect common alleles with odds ratio (OR) > 1.5; low-frequency alleles with OR > 2, Supplementary Figure 1).

Data availability
Genotypes of European cases included in this study are publicly available through the European Genome-Phenome Archive (EGA), under accession number EGAS00001000913, data set EGAD000010001043, with the exception of German and Dutch genotypes. Genotypes for cases from the United States of America may be obtained through dbGAP. Summary statistics are available for download from the PGC website (https://www.med.unc.edu/pgc/results-and-downloads).

RESULTS
GWAS and replication meta-analyses
Association testing was performed separately for each of the nine discovery cohorts within this study (2158 cases, 15 485 controls), and the results were meta-analyzed. No inflation was seen in the QQ plot (Figure 2b). Six variants were identified with \( P < 1 \times 10^{-5} \), and nine additional variants with \( P < 1 \times 10^{-4} \) (Figure 2a and Supplementary Table 5). Of these, one variant approached genome-wide significance (exm860538/rs199965409, \( P = 9.97 \times 10^{-5} \)), although this variant is polymorphic only in the Finnish population within these data sets, in the Exome Aggregation Consortium\(^{49}\) and in the 1000 Genomes project panel data.\(^{25}\) Variants with \( P < 1 \times 10^{-4} \) were taken forward for replication.

In total, 16 independent variants were selected for follow-up in one in silico cohort (1033 cases, 3733 controls) and one de novo genotyping cohort (261 cases, 15 000 controls). Of these, five were low frequency (MAF ~ 1%) and 11 were common frequency variants.

Twelve signals passed QC and were polymorphic in the de novo genotyping cohort, of which four were nominally significant (Supplementary Table 6; \( P < 0.05 \), minimum \( P = 0.001 \)). Eight of twelve SNPs had the same direction of effect as in the discovery GWAS, including three of the four nominally significant variants.

Ten of the sixteen variants were present in the in silico cohort, of which six had the same direction of effect as in the discovery cohort, and one of these six was associated with \( P = 0.02 \) (Supplementary Table 7).

Figure 2. Results from discovery-phase meta-analyses. (a) Manhattan plot for meta-analyzed \( P \)-values, across all nine populations. (b) QQ plot (\( \lambda = 0.94 \)).
On the basis of the number of SNPs taken forward for replication, we would not expect to see any variants reaching \( P < 0.05 \) by chance. We also see a higher concordance in direction of effect between discovery and replication cohorts (7/10 in the in silico analysis, 8/12 in the de novo analysis) than might be expected by chance; however, the number of SNPs tested was too small to achieve statistical significance (\( P = 0.17, P = 0.19 \), one-sided binomial test).

Five SNPs had the same direction of effect across the meta-analyzed discovery cohort and both replication cohorts. No SNPs reached genome-wide significance in the final global meta-analysis. Two variants were associated with the same direction of effect across discovery and replication cohorts, and reached \( P < 0.05 \) in at least one replication cohort (Table 3).

rs10791286 was associated with risk for AN across all discovery and replication cohorts (Figure 3a, global \( P = 9.89 \times 10^{-6} \), OR 0.84, 95% confidence interval 0.78–0.91). It resides in intron one of the opioid-binding protein/cell adhesion molecule-like (OPCML) gene. Data from the CommonMind Consortium project indicate that this variant is an eQTL for OPCML in the dorsolateral prefrontal cortex, and is associated with reduced expression (\( P = 0.014 \) after correction for multiple testing).\(^5\) OPCML has a role in opioid-binding and opioid receptor function\(^5\) and is expressed in a range of neuronal tissues, primarily the cerebellum and cerebellar hemispheres.\(^4\)–\(^6\) OPCML has previous associations with body mass index,\(^7\) waist–hip ratio,\(^8\) visceral fat distribution\(^9\) and alcohol dependence,\(^10\) among other phenotypes.

| Chr | Pos (Id) | Associated gene | EA | NEA | EAF | OR | OR_95L | OR_95U | P | Het_chisq | N_st (discovery/replication) |
|-----|----------|----------------|----|-----|-----|----|--------|--------|---|----------|-----------------------------|
| 2   | 19503281 | kgp3754622 (rs75245228) | a  | g   | 0.052 | 0.81 | 0.69 | 0.96 | 0.016 | 12.739 | 8 (6/2) |
| 11  | 133096498 | rs10791286 | a  | g   | 0.33 | 0.84 | 0.78 | 0.91 | 4.90 \times 10^{-6} | 4.228 | 8 (6/2) |
| 10  | 53754325 | PRKG1 | a  | g   | 0.20 | 0.87 | 0.79 | 0.95 | 0.0018 | 15.033 | 7 (5/2) |
| 11  | 125655014 | rs536968 | a  | g   | 0.12 | 0.88 | 0.79 | 0.98 | 0.023 | 12.082 | 8 (6/2) |
| 12  | 122695625 | exm860538 (rs199965409) | a  | g   | <0.01 | 10.42 | 4.40 | 24.69 | 9.97 \times 10^{-8} | 0 | 1 (1/0) |
| 4   | 157167891 | rs7700147 (rs75245228) | t  | c   | 0.21 | 1.20 | 1.10 | 1.30 | 2.79 \times 10^{-5} | 9.093 | 8 (6/2) |
| 6   | 34826040 | exm540361 (rs200155060) | a  | g   | <0.01 | 0.18 | 0.08 | 0.37 | 6.47 \times 10^{-6} | 0 | 1 (1/0) |
| 6   | 147840595 | rs669830 | a  | g   | 0.26 | 1.11 | 1.01 | 1.21 | 0.029 | 13.174 | 5 (3/2) |
| 21  | 47963149 | rs11701571 | DIP2A | a  | g   | 0.24 | 1.11 | 1.02 | 1.21 | 0.011 | 12.185 | 6 (4/2) |
| 7   | 49620107 | rs10264162 (rs145290255) | VWC2 | t  | g   | 0.43 | 0.91 | 0.84 | 0.97 | 0.0068 | 23.105 | 8 (6/2) |
| 1   | 197404688 | exm134618 (rs142090517) | CRB1 | a  | g   | <0.01 | 11.97 | 4.24 | 33.81 | 2.76 \times 10^{-6} | 0 | 1 (1/0) |
| 3   | 150748151 | rs1703802 | CLRN1-AS1 | t  | g   | 0.12 | 0.84 | 0.75 | 0.93 | 0.00085 | 7.007 | 8 (6/2) |
| 17  | 31082572 | exm131068 (rs145290255) | MYO1D | t  | c   | 0.001 | 0.02 | 0.00 | 0.10 | 1.74 \times 10^{-6} | 0.276 | 2 (1/1) |
| 4   | 80949892 | exm-rs1333130 | ANTRX2 | t  | c   | 0.38 | 0.89 | 0.83 | 0.94 | 7.14 \times 10^{-5} | 7.526 | 10 (8/2) |
| 4   | 26480221 | rs2854030 | CCKAR | t  | c   | 0.31 | 0.88 | 0.78 | 0.98 | 0.021 | 23.181 | 8 (6/2) |

Abbreviations: CHR, chromosome; EA, effect allele; EAF, effect allele frequency; \( I^2 \), measure of heterogeneity; NEA, non-effect allele; N_st, number of contributing studies; OR, odds ratio; OR_95L, lower 95% confidence interval; OR_95U, upper 95% confidence interval; P, P-value; POS, position in hg18. Gene names given are best-predicted consequence from ensembl,\(^{24,25}\) where none is available; the nearest gene is given instead in bold.

Figure 3. Odds ratios for two notable single-nucleotide polymorphisms (SNPs) across discovery and replication cohorts. (a) rs10791286 and (b) rs7700147.

On the basis of the number of SNPs taken forward for replication, we would not expect to see any variants reaching \( P < 0.05 \) by chance. We also see a higher concordance in direction of effect between discovery and replication cohorts (7/10 in the in silico analysis, 8/12 in the de novo analysis) than might be expected by chance; however, the number of SNPs tested was too small to achieve statistical significance (\( P = 0.17, P = 0.19 \), one-sided binomial test).

Five SNPs had the same direction of effect across the meta-analyzed discovery cohort and both replication cohorts. No SNPs reached genome-wide significance in the final global meta-analysis. Two variants were associated with the same direction of effect across discovery and replication cohorts, and reached \( P < 0.05 \) in at least one replication cohort (Table 3).
The variant itself has no previously reported associations in any phenotype.

rs7700147 was associated with AN across all discovery and replication cohorts (global \(P = 2.93 \times 10^{-5}\), OR 1.2, 95% confidence interval: 1.1, 1.3; Figure 3b). It is an intergenic variant and has no previous associations.

**Burden testing**

Burden testing allows the contribution of multiple low-frequency variants to be aggregated across discrete units (for example, genes). Three genes were identified with \(P < 1 \times 10^{-4}\), although none reached genome-wide significance (Table 4). A further five genes reached \(P < 1 \times 10^{-5}\), but passed inclusion thresholds in one population only (Table 4), and as such are likely to be false-positives.

FAM96A has previously been associated with low-density lipoprotein levels and cholesterol and is primarily expressed in the liver, lymphocytes and adrenal gland. KIF7 has no previous phenotype associations and has generally low expression across a wide range of tissues. C6orf10 has previous associations with visceral fat and childhood obesity, as well as a number of autoimmune disorders. C6orf10 is expressed in testes (see Discussion).

Biological pathways associated with AN

Allowing a 20 kb window for SNP to gene assignment identified two pathways significant at \(q < 0.05\): 'Phospholipase activator' and 'GTP-rho binding' (Table 5).

Using the strictest assignment method of SNPs to genes for the full data set, no pathways were significant after multiple-testing correction. The highest ranking pathway was 'Calcium ion import' (\(q\)-value = 0.069).

### DISCUSSION

To our knowledge, this work constitutes the first examination of low frequency (< 1% MAF) and rare exonic variation in AN in the context of a genome-wide scan. No low frequency or rare variant replicating associations were identified, although this study was well-powered to detect low-frequency variants with large effect sizes (Supplementary Figure 1). Although polymorphic only in the Finnish population, rs199965409 approached genome-wide significance. It is a non-synonymous variant with a MAF of 0.5% in the Finnish population. The variant is within the WDR1 gene, which is associated with hypogonadotropic hypogonadism. The clinical features of the disease, such as delayed sexual maturation, suggest that it may be misdiagnosed or comorbid with AN, which may explain its association in the analysis.

Two notable, but common-frequency, signals were identified with consistent direction of effect across discovery and replication cohorts (rs10791286 and rs7700147). These variants had been removed from the first Genetic Consortium for AN, as part of the Wellcome Trust Case Control Consortium 3 (WTCCC3) AN GWAS because of poor cluster plots; therefore, we were not able to compare effect sizes between studies. Burden tests to investigate an aggregation of rare variants within genes rendered three potentially interesting genes, which require further replication.

Studying rare variation presents a range of challenges. The sample sizes required to identify rare variants with modest effect sizes are substantially larger than for common variants. Further, the MAF spectra seen across trans-European populations differ more for rare variants than for common variants, especially when considering genetically distant populations such as Finland and Italy. This can reduce the power to detect a signal and achieve replication. There are also many technical challenges to consider when conducting a rare variant study; for example, the inflation seen in association tests at low minor allele counts and the increased error rate of calling algorithms when applied to rare variants. We mitigated against the latter challenge by comprehensively examining cluster plots of > 10 000 variants that surpassed a \(P\)-value threshold of \(P < 1 \times 10^{-4}\) in any analysis.

Of the genes potentially implicated through the single-point and burden test analyses, three have associations with metabolic and anthropometric phenotypes (OPCML, C6orf10 and FAM96A). OPCML has previously been associated with waist-to-hip ratio, while C6orf10 has associations with childhood obesity. FAM96A has been shown to be associated with metabolic phenotypes such as low-density lipoprotein and cholesterol levels. The associations of these three genes with metabolic and obesity-related phenotypes may indicate some roles for metabolic processes in AN development, although pathway analysis did not corroborate this observation. A growing body of evidence suggests involvement of metabolic processes in AN development, including appetite-satiety pathways, gut motility and gastric-emptying times. For example, application of the LD Score regression method revealed significant negative genetic correlations between AN and HDL cholesterol phenotypes.

Notably, C6orf10 has been previously associated with childhood obesity. This finding is particularly interesting for a number of reasons. First, appetite and satiety dysregulation have been shown to be central to the development of childhood obesity. In particular, reduced satiety responsiveness (experiencing an urge to eat despite internal 'full' signals) and heightened responsiveness to food have a role in increased adiposity. Aberrant responses to satiety signals and reduced responsiveness to food are also operative in AN, suggesting shared biological dysregulation between the two conditions.
The most significant pathway analysis association was with phospholipase activator pathways, which act to catalyze the hydrolysis of glycerophospholipids (GO:0016004 phospholipase activator activity). Phospholipase has a central role in the serotonin-triggered metabolism of arachidonic acid in the brain,\textsuperscript{93–95} which is a common target for antidepressants\textsuperscript{94,95} such as lithium, carbamazepine (Tegretol), valproate and lamotrigine (Lamictal).\textsuperscript{96} These antidepressants have been shown to have varying efficacy in treating AN.\textsuperscript{97–99} Lithium has been used in treatment of AN (with varying success),\textsuperscript{97–99} while carbamazepine and valproate have been successfully used in individuals with complex comorbid eating disorder phenotypes.\textsuperscript{100–104} Finally, lamotrigine has been shown to significantly improve eating disorder and mood symptoms in individuals with binge-eating and purging behaviors.\textsuperscript{105}

The second pathway identified as significantly associated with AN was GTP-rho binding. This pathway has a role in brain development, and is regulated by autism-susceptibility candidate gene 2 (AUTS2).\textsuperscript{106} This finding is consistent with the comorbidity between AN and autism.\textsuperscript{107} Moreover, individuals with AN may be socially withdrawn\textsuperscript{107} and exhibit elevated levels of autistic traits associated with lower social functioning.\textsuperscript{107–109} AUTS2 has also been well studied as a candidate gene for alcohol abuse,\textsuperscript{110} which is commonly comorbid with eating disorders.\textsuperscript{111} There is also a well-established link between GTP-rho activation and cognition.\textsuperscript{112} Mice with altered expression of genes regulating Rho-GTPases have been shown to have altered exploratory and anxiety-related behavior, decreased sociability and memory formation, and decreased body weight, among others.\textsuperscript{112} These findings are in line with some of the comorbidity and intermediate phenotypes noted in AN, for example, the high comorbidity with anxiety-related disorders.\textsuperscript{113}

There is substantial evidence for the involvement of chromatin-modulating genes in the development of autism.\textsuperscript{114–119} schizophrenia\textsuperscript{120–124} and body mass index changes.\textsuperscript{114} Given the comorbidity of these disorders with AN, and the potential overlap with autism indicated in the pathway analysis results, we tested for enrichment of chromatin-modulating genes in these results. We obtained a list of 340 genes involved in modifying chromatin accessibility and/or modifying histone marks from existing literature; of these, 30 reached nominal significance in our burden test, substantially more than expected by chance (binomial test, \(P = 0.0026\)). Moreover, one of the variants identified in the global meta-analysis (exm540361) lays near a gene included in this list (UHRF1BP1). Together, these results may indicate a role for chromatin-modifying genes in AN, although more work will be needed to investigate this further.

A number of limitations should be borne in mind when evaluating these results. First, the sample size of this study is small. Psychiatric disorders in general require very large sample sizes in order to identify reliable genome-wide significant signals.\textsuperscript{125} The current study was powered to detect common variants with substantial OR, and rare variants conferring substantial increases in disease risk (OR > 2). To our knowledge, this was the first time a study has specifically investigated the role of rare variation in AN, and the lack of low-frequency replicating findings may indicate that little advancements may be made in this particular genomic search space. We did not see any overlap between the pathways identified here and those identified in the recent PGC pathway analysis,\textsuperscript{126} however, this may reflect the relatively small sample size of this study, as well as different pathway analysis methodologies used.

In this study we only examined female AN cases of European origin. It has been suggested that the genetics underlying AN development may be easier to assess in an all-male study,\textsuperscript{3} as there may be a greater genetic risk required to induce trait expression. The higher relative risk in male subjects may also reflect this.\textsuperscript{3} To date, this has not been possible because of the lower prevalence of the disorder in men, resulting in substantially smaller sample sizes. Moreover, if AN is heterogeneous between populations, in order to fully understand the genetic etiology of the disorder, it will be necessary to expand collection to include more diverse samples. Efforts are already underway in a number of Asian populations such as Taiwan, Japan, Korea and China, as well as some South American populations such as Argentina and Brazil.

A caveat to this study is that controls were not screened for AN, and that both male and female controls were used. Given the population prevalence of AN across population of European descent, ~ 80 female and ~ 10 male controls would be expected to have AN diagnoses. Given the low rate of treatment seeking in AN,\textsuperscript{127} it would not be possible to confidently screen population-based or previously existing control cohorts for AN.

The underlying biological etiology of AN is complex and has not been elucidated yet. Here we have identified a number of variants that warrant follow-up in larger sample sizes, and which point to a role for metabolic, appetite-related and obesity-related effects, in line with a growing body of evidence for metabolic involvement in AN development. Substantially increased sample sizes and detailed phenotyping to reduce heterogeneity will be necessary to empower the characterization of the genetic architecture of AN.

CONFLICT OF INTEREST

GB has received grant funding and consultancy fees from Eli Lilly, DD is speaker, consultant or on advisory boards of various pharmaceutical companies, including AstraZeneca, Boehringer, Bristol Myers Squibb, Eli Lilly, 28 Genetics Pharma, GlaxoSmithKline, Janssen, Lundbeck, Organon, Sanofi, UniPharma and Wyeth, and he has unrestricted grants from Lilly and AstraZeneca as director of the Sleep Research Unit of Egnition Hospital (National and Kapodistrian University of Athens, Greece). AK is on the Shire Canada BED Advisory Board. JK is a member of SAB of AssurexHealth Inc (unpaid). ML has received lecture honoraria from Lundbeck, AstraZeneca and Biophasia Sweden, and served as scientific consultant for EPID Research Oy. There exists no other equity ownership, profit-sharing agreements, royalties, or patents. PS is scientific advisor to Pfizer, Inc. JT received an honorarium for speaking at a diabetic conference for Lilly and royalties from a published book. The remaining authors declare no conflicts of interest.

ACKNOWLEDGMENTS

Klarman Family Foundation: This study was supported by a grant from the Klarman Family Foundation.

Wellcome Trust Case Control Consortium: This work was funded by a grant from the WTCCC3 WT0888272/Z/09 entitled ‘A genomewide association study of anorexia nervosa’. Acknowledgments for specific authors: Eleftheria Zeggini is supported by the Wellcome Trust (098051). Lorraine Southam is supported by the Wellcome Trust (098051). Laura Hucks concludes Wellcome Trust (098051) and the MRC (MR/J003925/1). Cynthia Bulik acknowledges funding from the Swedish Research Council (VR Dnr: 538–2013–8864). Esther Walton is supported by the German Research Foundation (Wa 3635/1–1). Zeynep Yilmaz is funded by NIH K01MH109782. Allan Kaplan, Robert Levitan and James Kennedy acknowledge support from Ontario Mental Health Foundation for the collection of samples. Jaakko Kaprio is supported by the Academy of Finland (Grant Ref. 980). LM Huckins acknowledges Wellcome Trust (098051) and the MRC (MR/J003925/1). Cynthia Bulik acknowledges funding from the Swedish Research Council (VR Dnr: 538–2013–8864). Esther Walton is supported by the German Research Foundation (Wa 3635/1–1). Zeynep Yilmaz is funded by NIH K01MH109782. Allan Kaplan, Robert Levitan and James Kennedy acknowledge support from Ontario Mental Health Foundation for the collection of samples. Jaakko Kaprio is supported by the Academy of Finland (263278).

Charlotte’s Helix: We thank all the probands and parents from the patient-led group, Charlotte’s Helix. This charity was set up by Charlotte Bevan, after her daughter was diagnosed with anorexia nervosa. The charity is deeply committed to supporting biological work, particularly genetics, to help understand anorexia and has set up a database of patients’ names and details. Charlotte’s Helix has been collaborating closely with the King’s College London (KCL) team since 2013 by providing the database of probands, some funding and publicizing the scientific projects through regular blogs on its website and social media.

We gratefully acknowledge the participation of NIHR BRC South London and the Maudsley NHS Foundation Trust (SLaM) BioResource volunteers, and thank the NIHR BRC SLaM BioResource centre and staff for their contribution.

This paper represents independent research partly funded by the National Institute for Health Research (NIHR) Biomedical Research Centre and Dementia Unit at South London and Maudsley NHS Foundation Trust and King’s College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. We gratefully acknowledge capital equipment funding from the Maudsley Charity (Grant Ref. 980) and Guy’s and St Thomas’s Charity (Grant Ref. STR130505).

Molecular Psychiatry (2018), 1169 – 1180
Low-frequency and rare variation in anorexia nervosa
LM Huikins et al

EATING DISORDERS WORKING GROUP OF THE PSYCHIATRIC GENOMICS CONSORTIUM

RAH Adam15,16, L Alfredsson17, T Ando18, OA Andreassen19,20, H Aschauer21, Lara Mangravite, Mette Peters (Sage Bionetworks), Thomas Lehner and Barbara (University of Pennsylvania), Keisuke Hirai, Hiroyoshi Toyoshiba (Takeda Pharmaceuticals Company Limited, F Hoffman-La Roche and NIH grants R01MH057881, R01MH097276, R01-MH-075916, R01-MH060639, R50MH080405, R01MH097276, R01-MH075916, R05M509689, R01MH080455, R37MH075881 and R37MH075881, HHSN271201300031C, AG02219, AG05138 and MH06692. Brain tissue for the study was obtained from the following brain bank collections: the Alzheimer tissue for the study was obtained from the following brain bank collections: the Alzheimer Disease Core Center, the University of Pittsburgh NeuroBioBank and Brain and Tissue Repositories and the NIHIM Human Brain Collection Center. CMC Leadership: Pamela Sklar, Joseph Buxbaum (Icahn School of Medicine at Mount Sinai), Bernie Devlin, David Lewis (University of Pittsburgh), Rachel Gur, Chang-Gyu Hahn (University of Pennsylvania), Keeske Hsiai, Hiroyoshi Toyoshiba (Takeda Pharmaceuticals Company Limited), Enrico Domenici, Laurent Essouss (F. Hoffman-La Roche Ltd), Lara Mangravite, Mette Peters (Sage Bionetworks), Thomas Lehner and Barbara Lipska (NIHIM)
REFERENCES

1. Strober M, Freeman R, Lampert C, Diamond J, Kaye W. Males with anorexia nervosa: a controlled study of eating disorders in first-degree relatives. Int J Eat Disord 2001; 29: 263–269.

2. Lilenfeld LR, Kaye WH, Greeno CG, Merikangas KR, Plotnikov K, Pollice C et al. A controlled family study of anorexia nervosa and bulimia nervosa: psychiatric disorders in first-degree relatives and effects of proband comorbidity. Arch Gen Psychiatry 1998; 55: 69–70.

3. Bulik CM, Sif-Opt Landt MCT, van Furth E, Sullivan PF. The genetics of anorexia nervosa. Annu Rev Nutr 2007; 27: 263–275.

4. Bulik CM, Sullivan PF, Tozzi F, Furberg H, Lichtenstein P, Pedersen NL. Prevalence, heritability, and prospective risk factors for anorexia nervosa. Arch Gen Psychiatry 2006; 63: 305–312.

5. Wang K, Zhang H, Bloss CS, Dussved V, Kaye W, Schork NJ et al. A genome-wide association study of common SNPs and rare CNVs in anorexia nervosa. Mol Psychiatry 2011; 16: 949–959.

6. Boraska V, Franklin CS, Floyd JAB, Thornton LM, Hawkins LM, Southam L et al. A genome-wide association study of anorexia nervosa. Mol Psychiatry 2014; 19: 1085–1094.

7. Duncan L, Yilmaz Z, Gaspar H, Walters R, Goldstein J, Antilla V et al. Significant locus and metabolic genetic correlation revealed in genome-wide association study of anorexia nervosa. Am J Psychiatry 2017 (in press).

8. Poyato Pinheiro A, Thornton LM, Plotnikov KH, Tozzi F, Klump KL, Berrettini WH et al. Patterns of menstrual disturbance in eating disorders. Int J Eat Disord 2007; 40: 424–434.

9. American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders. 5th edn. American Psychiatric Association: Arlington, 2013.

10. American Psychiatric Association in Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association, 2013; doi:10.1176/appi.books.97808.

11. Preti A, de Girolamo G, Vilagut G, Alonso J, de Graaf R, Bruffaerts R et al. The epidemiology of eating disorders in six European countries: results of the ESEMeD-WMH project. J Psychiatr Res 2009; 43: 1125–1132.

12. Keski-Rahkonen A, Hoen HW, Susser ES, Linnis MS, Sivola E, Raevuo A et al. Epidemiology and course of anorexia nervosa in the community. Am J Psychiatry 2007; 164: 1259–1265.

13. Rathner G, Messner K. Detection of eating disorders in a small rural town: an epidemiological study. Psychol Med 1993; 2: 175–184.

14. Bijl RV, Ravelli A, van Zessen G. Prevalence of psychiatric disorder in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). Soc Psychiatry Psychiatr Epidemiol 1998; 33: 587–595.

15. Kringlein E, Torgersen S, Cramer V. A Norwegian Psychiatric Epidemiological Study. Am J Psychiatry; 2001: 158: 1091–1098.

16. PriceWaterhouseCoopers & BEAT. The costs of eating disorders Social, health and economic impacts. 2015; https://www.b-eat.co.uk/assets/000/000/302/The_costs_of_eatdisorders_Final_original.pdf.

17. Hudson J, Birui E, Pope HG, Kessler RC. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. Biol Psychiatry 2007; 61: 348–358.

18. Illumina. Infinium II Human Exome-12 v1.2 BeadChip. Data Sheet: Genotyping 2015. Available at: http://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet%7B_%7Dhummexome%7B_%7Dbeadchips.pdf.

19. Illumina. Infinium CoreExome-24 v1.1 BeadChip. Data Sheet: Genotyping 2015; pp 1–2. Available at: http://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/datasheet%7B_%7Dhummexome%7B_%7Dbeadchips.pdf.

20. Gurdasani D, Cantensen T, Tzekola-Ayale F, Pagani L, Tschasmodiou I, Hatzikotoulas K et al. An integrated map of genetic variation from 1,092 human genomes. Nature 2014; 517: 327–332.

21. Ritchie ME, Liu R, Carvalho BS, Irizarry RA. Comparing genotyping algorithms for Illumina’s Infinium whole-genome SNP BeadChips. BMC Bioinformatics 2011; 12: 68.

22. Goldstein JI, Crenshaw A, Carey J, Grant GB, Maguire J, Fromer M et al. Zell: a rare variant caller for array-based genotyping: genetics and population analysis. Bioinformatics 2012; 28: 2542–2545.

23. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Plink: a tool set for whole-genome association and population-based linkage analyses. Am J Human Genet 2006; 79: 313–323.

24. ESEMeD-WMH project. The costs of eating disorders Social, health and economic impacts. 2015; https://www.b-eat.co.uk/assets/000/000/302/The_costs_of_eatdisorders_Final_original.pdf.

25. Strouf SA, The American Soldier (Vol. 1): Adjustment During Army Life – The American Soldier (Vol. 1): Adjustment During Army Life – The American Soldier (Vol. 1): Adjustment During Army Life – The American Soldier (Vol. 1): Adjustment During Army Life. Princeton University Press: Princeton, NJ, 1949.

26. Koenig H, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet 2010; 11: 499–511.

27. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genome-wide association scans. Bioinformatics 2010; 26: 2190–2191.

28. Rivas D, McLaren WM, Chen Y, Birney E, Stubenau A, Flicek P et al. A database and API for variation, dense genotyping and resequencing data. BMC Bioinformatics 2010; 11: 238.

29. Bronwen LA, Sarah A, Daniel B, Laura C, Valery C, Susan F et al. The Ensemble gene annotation system. Database (Oxford) 2016; 2016: baw093; doi: 10.1093/database/baw093.

30. Rayner NW, Borsboom, McCarty M. ScatterShot: A Java program for creating cluster plots from Affymetrix and Illumina genotype data. American Society of Human Genetics Meeting (Poster Presentation). American Society of Human Genetics Meeting. Boston Convention & Exhibition Center (BCC); 22–26 October 2013.

31. Morris AP, Zeggini E. An evaluation of statistical approaches to rare variant analysis in genetic association studies. Genet Epidemiol 2010; 34: 188–193.

32. Stouffer SA. The American Soldier (Vol. 1): Adjustment During Army Life. Princeton University Press: Princeton, NJ, 1949.

33. Kaneshia M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res 2014; 42(Database issue): D199–D205.
35 Kaneshia M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000; 28: 27–30.
36 Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G et al. The reactome pathway knowledgebase. Nucleic Acids Res 2014; 42(Database issue): D472–D477.
37 Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res 2013; 41(Database issue): D234–D238.
38 Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R et al. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 2003; 13: 2129–2141.
39 The Gene Ontology Consortium. Gene Ontology Consortium: going forward. Nucleic Acids Res 2014; 43(D1): D1049–D1056.
40 Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM et al. Gene ontology tool for the unbiased definition of biology. The Gene Ontology Consortium. Nat Genet 2000; 25: 25–29.
41 de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 2015; 11: e1004219.
42 Benjamin Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc 1995; 57: 289–300.
43 First MB, Spitzer RL, Gibbon M, Williams JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-Patient Edition (SCID-I/NP) Bio- metrics Research. New York State Psychiatric Institute: New York, NY, USA, 2002.
44 Ardile KG, Deluca DS, Segre AV, Sullivan TJ, Young TR, Gelfand ET et al. The genotype-tissue expression (GTEX) pilot study: multitissue gene regulation in humans. Science 2015; 348: 646–660.
45 The GTEX Consortium. The Genotype-Tissue Expression (GTex) project. Nature Genetics 2013; 45: 580–585.
46 Mele M, Ferreira PG, Reverter F, DeLuca DS, Butler H, Cherry JM et al. Gene ontology tool for the unbiased definition of biology. The Gene Ontology Consortium. Nat Genet 2000; 25: 25–29.
47 Gauderman WJ. Sample size requirements for matched case-control studies of gene-environment interaction. Stat Med 2002; 21: 35–50.
48 Gauderman WJ, Morrison JM. QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies. 2006. Available at: http://biostats.usc.edu/software.
49 Exome Aggregation Consortium et al. Analysis of protein-coding genetic variation in 60,706 humans. bioRxiv (Cold Spring Harbor Labs Journals, 2015). doi:10.1101/030338.
50 Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM et al. A genome-wide association study of body fat distribution in African ancestry populations supports an ancestral ACE polymorphism. Hum Mol Genet 2013; 22(Suppl 1): S18.
51 Liu CT, Monda KL, Taylor KC, Lange L, Demerath EW, Palmas W et al. Genome-wide association of body fat distribution in African ancestry populations supports new loci. PLoS Genet 2013; 9: e1003681.
52 Fox CS, Heard-Costa N, Culppeers LA, Dupuis J, Vasan RS, Atwood LD. Genome-wide association to body mass index and waist circumference: the Framingham Heart Study 100K project. BMC Med Genet 2007; 8(Suppl 1): S18.
53 Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S et al. Congenital hypogonadotropic hypogonadism and kallmann syndrome: past, present, and future. Endocr Metab Clin 2015; 30: 456–466.
54 Benficial P. Delayed puberty. Endocr Dev 2012; 22: 138–159.
55 Shah TS, Liu JZ, Floyd JAB, Morris JA, Wirth N, Barrett JC et al. optCall: a robust genotype-calling algorithm for rare, low-frequency and common variants. Bioinformatics 2012; 28: 1598–1603.
56 McCallum RW, Grill BB, Lange R, Planky M, Glass EE, Greenfield DG. Definition of a gastric emptying abnormality in patients with anorexia nervosa. Dig Dis Sci 1985; 30: 713–722.
57 Szmukler GI, Young GP, Lichtenstein M, Andrews JT. A serial study of gastric emptying in anorexia nervosa and bulimia. Aust N Z J Med 1990; 20: 220–225.
58 Benini L, Todesco T, Dalle Grave R, Deiorio F, Salandini L, Vantini I. Gastric emptying in patients with restricting and binge/purging subtypes of anorexia nervosa. Am J Gastroenterol 2004; 99: 1448–1454.
59 Humphries L, Shih WJ. Gastric emptying time in anorexia nervosa and bulimia. Arch Surg 1988; 123: 783.
60 Kamal N, Chami T, Andersen A, Rosell FA, Schuster MM, Whitehead WE. Delayed gastrointestinal transit times in anorexia nervosa and bulimia nervosa. Gastroentology 1991; 101: 1320–1324.
61 Stacher G, Kiss A, Wiesnagrotzki S, Bergmann H, Hörbat J, Schneider C. Oesophageal and gastric motility disorders in patients categorised as having primary anorexia nervosa. Gut 1986; 27: 1120–1126.
62 García Aroca J, Alonso Calderón JL, García Redondo C, Rollán Villamarín V. Anorexia nervosa or somatic disease. Cir Pediatr 2001; 14: 98–102.
63 Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Consortium R et al. An atlas of genetic correlations across human diseases and traits. Nat Genet 2015; 47: 1236–1241.
64 Llewellyn CH, Trzaskowski M, van Jaarsveld CHM, Plomin R, Wardle J. Delayed puberty. Endocr Dev 2012; 22: 138–159.
65 Fairburn CG, Cooper Z, Doll HA, Davies BA. Identifying dieters who will develop an eating disorder: a prospective, population-based study. Am J Psychiatry 2005; 162: 2249–2255.
66 Patton GC, Selzer R, Coffey C, Carlin JB, Wolfe R. Onset of adolescent eating disorders: population based cohort study over 3 years. BMJ 1999; 318: 765–768.
67 Santonastaso P, Frederici S, Favaro A. Full and partial syndromes in eating disorders: a 1-year prospective study of risk factors among female students. Psychopathology 2012; 50: 56–50.
68 Rohde P, Stice E, Marti CN. Development and predictive effects of eating disorder risk factors during adolescence: Implications for prevention efforts. Int J Eat Disord 2015; 48: 187–198.
69 Gardner RM, Stark K, Friedman BN, Jackson NA. Predictors of eating disorder scores in children ages 6 through 14: a longitudinal study. J Psychosom Res 2000; 49: 199–205.
91 Killen JD, Taylor CB, Hayward C, Haydel KF, Wilson DM, Hammer L et al. Weight concerns influence the development of eating disorders: a 4-year prospective study. J Consult Clin Psychol 1996; 64: 936–940.
92 Copeland WE, Bulik CM, Zucker N, Wolke D, Lereya ST, Costello EJ. Does child-
hood bullying predict eating disorder symptoms? A prospective, longitudinal analysis. Int J Eat Disord 2015; 48: 1141–1149.
93 Felder CC, Kanterman RY, Ma AL, Axelrod J. Serotonin stimulates phospholipase A2 and the release of arachidonic acid in hippocampal neurons by a type 2 serotonin receptor that is independent of inositolphospholipid hydrolysis. Proc Natl Acad Sci USA 1990; 87: 2187–2191.
94 Lee H-J, Rao JS, Chang L, Rapoport SI. Chronic imprompine but not bupropion increases arachidonic acid signaling in rat brain: is this related to ‘switching’ in bipolar disorder? Mol Psychiatry 2010; 15: 602–614.
95 Bazinet RP. Is the brain arachidonic acid cascade a common target of drugs used to manage bipolar disorder? Biochem Soc Trans 2009; 37(Pt 5): 1104–1109.
96 Rao JS, Rapoport SI. Mood-stabilizers target the brain arachidonic acid cascade. Curr Mol Pharmacol 2009; 2: 207–214.
97 Brewerton TD. Antipsychotic agents in the treatment of anorexia nervosa: neuropsychopharmacologic rationale and evidence from controlled trials. Curr Psychiatry Rep 2012; 14: 398–405.
98 Suárez-Pinilla P, Peña-Pérez C, Abtaizar-Barrenechea B, Crespo-Facorro B, Del Barrio JAG, Treasure J et al. Inpatient treatment for anorexia nervosa: a systematic review of randomized controlled trials. J Psychiatr Pract 2015; 21: 49–59.
99 Milano W, De Rosa M, Milano L, Riccio A, Sanseverino B, Capasso A. The pharmacological options in the treatment of eating disorders. ISRN Pharmacol 2013; 2013: 352865.
100 McClory SL, Kotwal R, Keck PE, Akisal HS. Comorbidity of bipolar and eating disorders: distinct or related disorders with shared dysregulations? J Affect Disord 2005; 86: 107–127.
101 McClory SL, Guerdjikova AI, Martens B, Keck PE, Pope HG, Hudson JI. Role of antiepileptic drugs in the management of eating disorders. CNS Drugs 2009; 23: 139–156.
102 Cordás TA, Tavares H, Calderoni DM, Stump GV, Ribeiro RB. Oxcarbazepine for self-mutilating bulimic patients. Int J Neuropsychopharmacol 2006; 9: 769–771.
103 Tachibana N, Sugita Y, Teshima Y, Hishikawa Y. A case of anorexia nervosa associated with epileptic seizures showing favorable responses to sodium valproate and clonazepam. Jpn J Psychiatry Neurol 1989; 43: 77–84.
104 Tor PC, Lee EL. Treatment emergent mania responding to valproate in a Chinese female adolescent population with eating disorders: a case series. Eur Eat Disord Rev 2008; 16: 421–426.
105 Trunko ME, Schwartz TA, Marzola E, Klein AS, Kaye WH. Lamotrigine use in patients with binge eating and purging, significant affect dysregulation, and poor impulse control. Int J Eat Disord 2014; 47: 329–334.
106 Horii K, Nagai T, Shan W, Sakamoto A, Taya S, Hashimoto R et al. Cytoskeletal regulation by AUTS2 in neuronal migration and neurogenesis. Cell Rep 2014; 9: 2166–2179.
107 Zucker NL, Losh M, Bulik CM, LaBar KS, Piven J, Pelphrey KA. Anorexia nervosa and autism spectrum disorders: guided investigation of social cognitive endo-
phenotypes. Psychol Bull 2007; 133: 976–1006.
108 Baron-Cohen S, Jaffa T, Davies S, Ayuengu B, Allison C, Wheelwright S. Do girls with anorexia nervosa have elevated autistic traits? Mol Autism 2013; 4: 24.
109 Koch SV, Larsen JT, Mouridsen SE, Bentz M, Petersen L, Bulik C et al. Autism spectrum disorder in individuals with anorexia nervosa and in their first- and second-degree relatives: Danish nationwide register-based cohort-study. Br J Psychiatry 2015; 206: 401–407.
110 Schumann G, Com LJ, Lourdusamy A, Charoen P, Berger KH, Stacey D et al. Genome-wide association and genetic functional studies identify autism suscept-
ibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. Proc Natl Acad Sci USA 2011; 108: 7119–7124.
111 Bulik CM, Klump KL, Thornton L, Kaplan AS, Devlin B, Fichter MM et al. Alcohol use disorder comorbidity in eating disorders: a multicenter study. J Clin Psychiatry 2004; 65: 1000–1006.
112 De Filippis B, Romano E, Laviola G. Aberrant Rho GTPases signaling and cognitive dysfunction: in vivo evidence for a compelling molecular relationship. Neurosci Biobehav Rev 2014; 46(Pt 2): 285–301.
113 Kaye WH, Bulik CM, Thornton L, Barbarich N, Masters K. Comorbidity of anxiety disorders with anorexia and bulimia nervosa. Am J Psychiatry 2004; 161: 2215–2221.
114 Loviglio MN, Leleu M, Männik K, Passergi G, Giannuzzi G, van der Werf I et al. Chromosomal contacts connect loci associated with autism, BMI and head circumference phenotypes. Mol Psychiatry 2016.
115 Ciernia AV, LaSalle J. The landscape of DNA methylation amid a perfect storm of autism aetologies. Nat Rev Neurosci 2016; 17: 411–423.
116 Barnard RA, Pomaville MB, O’Roak BJ. Mutations and modeling of the chromatin remodeler CHD8 define an emerging autism etiology. Front Neurosci 2015; 9: 477.
117 Shang L, Henderson LB, Cho MT, Petrey DS, Fong C-T, Haude KM et al. De novo missense variants in PPP2R5D are associated with intellectual disability, macrocephaly, hypotonia, and autism. Neurogenetics 2016; 17: 43–49.
118 Grayson DR, Guidotti A. Merging data from genetic and epigenetic approaches to better understand autistic spectrum disorder. Epigenomics 2016; 8: 85–104.
119 Sanders SJ. First glimpses of the neurobiology of autism spectrum disorder. Curr Opin Genet Dev 2015; 33: 80–92.
120 Eckart N, Song Q, Yang R, Wang R, Zhu H, McCallion AS et al. Functional char-
acterization of schizophrenia-associated variation in CACNA1C. PLoS ONE 2016; 11:e0157086.
121 Fryland T, Christensen JH, Pallesen J, Matthesen M, Palmfeldt J, Bak M et al. Identification of the BRD1 interaction network and its impact on mental disorder risk. Genome Med 2016; 8: 53.
122 González-Peñas J, Amigo J, Santomé L, Sobrino B, Brenilla J, Agra S et al. Targeted resequencing of regulatory regions at schizophrenia risk loci: role of rare func-
tional variants at chromatin repressive states. Schizophr Res 2016; 174: 10–16.
123 Maschietto M, Tahira AC, Puga R, Lima L, Mariani D, Paulsen B et al. Co-expression network of neural-differentiation genes shows specific pattern in schizophrenia. BMC Med Genomics 2015; 8: 23.
124 Bharadwaj R, Peter CJ, Jiang Y, Roussos P, Vogel-Ciernia A, Shen Y et al. Conserved higher-order chromatin regulates NMDA receptor gene expression and cognition. Neuron 2014; 84: 997–1008.
125 McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008; 9: 356–369.
126 Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat Neurosci 2015; 18: 199–209.
127 Keski-Rahkonen A, Mustelin L. Epidemiology of eating disorders in Europe. Curr Opin Psychiatry 2016; 29: 340–345.

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)