

## Original Article

**Cite this article:** Thorp JG, Marees AT, Ong J-S, An J, MacGregor S, Derks EM (2020). Genetic heterogeneity in self-reported depressive symptoms identified through genetic analyses of the PHQ-9. *Psychological Medicine* **50**, 2385–2396. <https://doi.org/10.1017/S0033291719002526>

Received: 7 April 2019

Revised: 6 August 2019

Accepted: 23 August 2019

First published online: 18 September 2019


**Key words:**

Depression; depressive symptoms; genetic heterogeneity; GWAS; PHQ-9

**Author for correspondence:**

Jackson G. Thorp, E-mail: [Jackson.Thorp@qimrberghofer.edu.au](mailto:Jackson.Thorp@qimrberghofer.edu.au) and Eske M. Derks, E-mail: [Eske.Derks@qimrberghofer.edu.au](mailto:Eske.Derks@qimrberghofer.edu.au)

# Genetic heterogeneity in self-reported depressive symptoms identified through genetic analyses of the PHQ-9

Jackson G. Thorp<sup>1</sup> , Andries T. Marees<sup>1,2</sup>, Jue-Sheng Ong<sup>3</sup>, Jiyuan An<sup>3</sup>, Stuart MacGregor<sup>3</sup> and Eske M. Derks<sup>1,4</sup>

<sup>1</sup>Translational Neurogenomics Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia; <sup>2</sup>Department of Psychiatry, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; <sup>3</sup>Statistical Genetics Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia and <sup>4</sup>School of Biomedical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, Queensland, Australia

**Abstract**

**Background.** Depression is a clinically heterogeneous disorder. Previous large-scale genetic studies of depression have explored genetic risk factors of depression case-control status or aggregated sums of depressive symptoms, ignoring possible clinical or genetic heterogeneity.

**Methods.** We analyse data from 148 752 subjects of white British ancestry in the UK Biobank who completed nine items of a self-rated measure of current depressive symptoms: the Patient Health Questionnaire (PHQ-9). Genome-Wide Association analyses were conducted for nine symptoms and two composite measures. LD Score Regression was used to calculate SNP-based heritability ( $h^2_{\text{SNP}}$ ) and genetic correlations ( $r_g$ ) across symptoms and to investigate genetic correlations with 25 external phenotypes. Genomic structural equation modelling was used to test the genetic factor structure across the nine symptoms.

**Results.** We identified nine genome-wide significant genomic loci (8 novel), with no overlap in loci across symptoms.  $h^2_{\text{SNP}}$  ranged from 6% (concentration problems) to 9% (appetite changes). Genetic correlations ranged from 0.54 to 0.96 (all  $p < 1.39 \times 10^{-3}$ ) with 30 of 36 correlations being significantly smaller than one. A two-factor model provided the best fit to the genetic covariance matrix, with factors representing 'psychological' and 'somatic' symptoms. The genetic correlations with external phenotypes showed large variation across the nine symptoms.

**Conclusions.** Patterns of SNP associations and genetic correlations differ across the nine symptoms, suggesting that current depressive symptoms are genetically heterogeneous. Our study highlights the value of symptom-level analyses in understanding the genetic architecture of a psychiatric trait. Future studies should investigate whether genetic heterogeneity is recapitulated in clinical symptoms of major depression.

**Introduction**

Clinical depression is a markedly complex and debilitating mental disorder characterised by sad, irritable or empty mood, diminished pleasure, and cognitive and somatic impairment (Christian *et al.*, 2016). The heritability of major depression is estimated to be ~37% from twin studies (Sullivan *et al.*, 2000) with common single nucleotide polymorphisms (SNPs) explaining around 9% of the variation in liability (Wray *et al.*, 2018; Howard *et al.*, 2019). Depression has substantial comorbidity with other psychiatric and substance use disorders and is related to a wide range of personality, socioeconomic and human traits (Lee *et al.*, 2013). There is substantial overlap in the genetic risk factors of major depression and other psychiatric disorders (Wray *et al.*, 2018), including significant genetic correlations ( $r_g$ ) with anxiety disorders ( $r_g = 0.80$ ), schizophrenia ( $r_g = 0.34$ ), bipolar disorder ( $r_g = 0.32$ ), autism spectrum disorders ( $r_g = 0.44$ ) and attention-deficit/hyperactivity disorder (ADHD) ( $r_g = 0.42$ ).

Initial efforts to identify genetic variants associated with major depression were unsuccessful, despite successes with other psychiatric diseases and traits. While a Genome-Wide Association Study (GWAS) of schizophrenia (9394 cases), for example, detected seven genome-wide significant associations (Ripke *et al.*, 2011), a mega-analysis of major depression (9240 cases) (Ripke *et al.*, 2013) and a meta-analysis of depressive symptoms ( $N = 34\,549$ ) (Hek *et al.*, 2013) found no significant associations. By 2014, 108 independent genetic loci for schizophrenia had been identified (Ripke *et al.*, 2014), and not a single one for depression. The struggle to identify significant genetic variants was likely related to low statistical power due to the clinical heterogeneity of major depression, although other factors are also involved

(e.g. high polygenicity, high disease prevalence and low heritability of depression) (Levinson *et al.*, 2014).

The Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5) defines major depressive disorder (MDD) based on nine symptoms (American Psychiatric Association, 2013). For a diagnosis of MDD, five or more of these symptoms need to be present during a 2-week period, with at least one symptom being depressed mood or anhedonia. Østergaard *et al.* (2011) highlighted that there are 227 possible combinations of symptoms meeting DSM-5 criteria, indicating MDD is an extremely heterogeneous disorder. Further, individual symptoms have been found to differ substantially in their association with psychosocial impairment, influence from environmental and personality risk factors, and biological correlates (Fried and Nesse, 2015).

Two broad GWAS approaches have been utilised to discover risk loci for depression: (1) maximising sample size by combining different and often broad measures such as questionnaire data, with the view that the increase in sample size overcomes phenotypic heterogeneity; and (2) reducing clinical heterogeneity by analysing homogenous MDD phenotypes (e.g. recurrent, clinically-diagnosed MDD). In the last few years, increasing sample size has proved to be effective with the number of genome-wide significant variants increasing steadily with sample size. Hyde *et al.* (2016) identified 15 genome-wide significant loci associated with self-reported depression ( $N = 307\,354$ ). In 2018, another 17 loci were identified across three broad depression phenotypes ( $N = 322\,580$ ) (Howard *et al.*, 2018), and 44 loci in a GWA meta-analysis of major depression ( $N = 480\,359$ ) (Wray *et al.*, 2018). The largest GWAS of depression to date ( $N = 807\,553$ ) identified 101 significant loci (Howard *et al.*, 2019).

In the present study, we combine these two approaches by conducting genetic analyses on individual depressive symptoms using large-scale population questionnaire data. Previous GWASs of depression have typically focused on MDD case-control status or aggregated sums of depressive symptoms. By combining different symptoms into a single clinical measure, it is implicitly assumed that individual symptoms of depression are genetically similar. However, the extreme heterogeneity of depression and numerous clinical presentations of the disorder suggest that different biological mechanisms could underlie the diverse subtypes of depression. Supporting this notion, symptoms of depression have been found to differ substantially in heritability (twin-based  $h^2$  range, 0–35%); with somatic and cognitive symptoms being most heritable (Jang *et al.*, 2004). Further, the diagnostic criteria of MDD were found to reflect three underlying genetic factors (cognitive/psychomotor symptoms, mood symptoms and neurovegetative symptoms) rather than a single factor of genetic risk in a twin study (Kendler *et al.*, 2013). Nagel *et al.* (2018b) found substantial genetic heterogeneity in neuroticism, a personality trait with extensive phenotypic and genetic overlap with major depression (Hettema *et al.*, 2006), by conducting genetic analyses on the individual items used to measure neuroticism.

To date, the extent to which common genetic risk factors overlap in individual symptoms of depression is not known. The aim of the present study is to examine and assess the extent of genetic heterogeneity in self-rated depressive symptoms as measured by the nine items of the Patient Health Questionnaire (PHQ-9) (Kroenke *et al.*, 2001a), a depression measure which directly maps onto the DSM-5 criteria. Although endorsement of these symptoms may not necessarily be occurring within an episode of major depression, they may possess a similar underlying

genetic basis and therefore provide insight into the genetic architecture of major depression. We conduct genetic analyses in 148 752 participants within the UK Biobank (UKBB). In order to examine genetic heterogeneity, we (a) conduct symptom-level GWA analyses and then compare genetic associations and SNP-based heritability across symptoms; (b) calculate phenotypic and genetic correlations across symptoms and determine their underlying genetic factor structure; and (c) calculate genetic correlations between individual symptoms and a range of psychiatric disorders and human complex traits.

## Methods

### UKBB cohort

UKBB is a major health data resource containing phenotypic information on a wide range of health-related measures and characteristics in over 500 000 participants from the UK general population (Bycroft *et al.*, 2018). Participants were recruited between 2006 and 2010 and provided written informed consent. A total of 157 365 participants completed the PHQ-9, as part of a UKBB mental health follow-up questionnaire administered online in 2016.

### Sample selection

First, participants were included in the present study if they were of white British ancestry, identified through self-reported ethnicity and genetic principal components. Participants who self-reported as not white British, but for whom the first two genetic principal components indicated them to be genetically similar to those of white British ancestry were also included in order to maximise sample size (these commonly were participants who reported to be of Irish ancestry). Second, participants were excluded if they were identified with schizophrenia and/or other psychotic disorders, bipolar disorder, cyclothymic disorder or dissociative identity disorder, based on self-reported symptoms or diagnosis, reported prescription of an antipsychotic medication and/or ICD-10 codes from linked hospital admission records. Third, only participants who provided a response for all nine items of the PHQ-9 were included (list-wise deletion represented a <2% reduction in sample size). This resulted in a final sample size of 148 752 (see online Supplementary Fig. S1 for a flow diagram of sample selection).

### PHQ-9

The PHQ-9 is a commonly used self-administered measure of depression containing nine items that map directly onto the nine DSM diagnostic criteria for major depression (Kroenke *et al.*, 2001a). Each PHQ-9 item assesses the frequency of that symptom over the past 2 weeks, rated on a four-point ordinal scale: (0) Not at all, (1) Several days, (2) More than half the days, (3) Nearly every day (see online Supplementary Table S1 for the nine symptoms of major depression, PHQ-9 items and DSM-5 diagnostic criteria).

The PHQ-9 is a psychometrically valid and reliable measure of depression (Kroenke *et al.*, 2001b). Test-retest reliability was high ( $r = 0.84$ , over a span of 48 h) and internal consistency was excellent with Cronbach's  $\alpha$  of 0.89 and 0.86 in primary care and obstetrics-gynaecology samples, respectively. The authors also reported good criterion and construct validity. The PHQ-9 was validated against professional diagnoses of MDD, resulting in

**Table 1.** Sample sizes, prevalence rates and SNP-based heritability estimates of all binary PHQ-9 items

Item	<i>N</i> endorsed	<i>N</i> not endorsed	Prevalence %	$h^2_{\text{SNP}}$ (S.E.)
Anhedonia	27 288	121 464	18.3	0.081 (0.0082)
Depressed mood	32 263	116 489	21.7	0.067 (0.0076)
Sleep problems	72 302	76 450	48.6	0.068 (0.0065)
Fatigue	73 924	74 828	49.7	0.079 (0.0070)
Appetite changes	26 757	121 995	18.0	0.091 (0.0087)
Low self-esteem	28 206	120 546	19.0	0.063 (0.0076)
Concentration problems	26 229	122 523	17.6	0.060 (0.0086)
Psychomotor changes	7914	140 838	5.3	0.066 (0.0154)
Suicidal ideation	6064	142 688	4.1	0.068 (0.0150)

Heritability (on the liability scale) estimated via single-trait LD Score Regression. All estimates are significant after multiple testing correction ( $p < 4.55 \times 10^{-3}$ ).

88% sensitivity and 88% specificity (at a PHQ-9 sum-score of  $\geq 10$ ); and scores correlated highly with similar constructs, such as the 20-item Short-Form General Health Survey (SF-20) (Stewart *et al.*, 1988) mental health scale ( $r = 0.73$ ). Internal consistency of the PHQ-9 in the UKBB sample in the current study was high (Cronbach's  $\alpha = 0.83$ ).

### Depression item phenotypes

Each of the nine PHQ-9 items is considered a separate phenotype in the genetic analyses. The ordinal scale of measurement of these items complicates interpretation of the SNP-based heritability ( $h^2_{\text{SNP}}$ ) estimates. In order to interpret  $h^2_{\text{SNP}}$  of each of the PHQ-9 items, each ordinal phenotype was transformed into a binary phenotype. The nine items were dichotomised such that an item was considered to be endorsed if the item score was one or greater (several days, more than half the days or nearly every day), and not endorsed if the score was zero (not at all). A cut-off score of one was used in order to maximise the number of subjects who endorsed an item and hence statistical power, a strategy that has provided greater benefit in GWASs of depression over ensuring a seamless phenotype (Wray *et al.*, 2012; Hall *et al.*, 2018; Howard *et al.*, 2018; Wray *et al.*, 2018; Howard *et al.*, 2019). In addition to the nine ordinal items and nine binary items, a *sum-score* (sum of all ordinal item scores; ranging from 0 to 27) and *binary sum-score* (number of binary items endorsed; ranging from 0 to 9) were included as phenotypes. We will present the results from the binary items and the two sum-scores while results for ordinal items are provided in online Supplementary material.

### GWA analyses

A total of 20 GWA analyses were conducted (nine ordinal scale depression items, nine binary items, plus the sum-score and binary sum-score phenotypes) using BOLT-LMM (Loh *et al.*, 2015). Associations between SNPs and a phenotype are tested using a linear mixed model in order to correct for population structure and cryptic relatedness. While BOLT-LMM is based on a quantitative trait model, it can be used to analyse binary traits by treating them as continuous and applying a transformation. Ordinal items are treated as continuous. An issue when analysing binary traits in BOLT-LMM is the inflated type 1 error rates for rare SNPs when the number of cases and controls are very unbalanced (Zhou *et al.*, 2018). In practice, all of the traits we consider here

have a case proportion which is large enough ( $>3\%$ ) for this not to be a problem (Loh *et al.*, 2018).

Analyses were limited to autosomal SNPs with high imputation quality score (INFO score  $\geq 0.80$ ) and a minor allele frequency of 1% or higher, resulting in 9 413 637 SNPs being tested for association. Sex, age at baseline and batch were included as covariates. GWAS results were annotated using FUMA GWAS (Watanabe *et al.*, 2017). The conventional genome-wide significance threshold of  $p < 5 \times 10^{-8}$  was applied. Given the exploratory nature of the analyses and identifying causal variants is not the primary interest of this paper, plus the high correlation between the 20 phenotypes, we did not correct for multiple testing of the 20 phenotypes as this would lead to increased type-II error rate.

Significant SNPs were clumped into blocks high in linkage disequilibrium (the non-random association of alleles at a specific locus; LD) using a threshold of  $r^2 < 0.10$  [correlation between allele frequencies of two SNPs; as calculated by PLINK (Purcell *et al.*, 2007)]. Independent significant SNPs were defined as the SNP with the lowest  $p$  value within an LD block. Genomic risk loci (distinct, fixed positions on a chromosome) were identified by merging independent SNPs if  $r^2 \geq 0.10$  and their LD blocks are physically close to each other at a distance of 1000 kb.

### LDSC analyses

Estimates of the variance in each phenotype attributable to the additive effects of all SNPs (SNP-based heritability;  $h^2_{\text{SNP}}$ ) were calculated via single-trait LD Score Regression using GWAS summary statistics from our analyses (Bulik-Sullivan *et al.*, 2015b) (see online Supplementary methods). To interpret  $h^2_{\text{SNP}}$  for binary items, estimates were converted to the liability scale, where the population prevalence of PHQ-9 items was estimated from our UKBB sample (population prevalence = sample prevalence; see Table 1). We applied a Bonferroni-corrected significance threshold for the 11  $h^2_{\text{SNP}}$  estimates ( $p < 4.55 \times 10^{-3}$ ).

Cross-trait LD Score Regression was used to estimate genetic correlations ( $r_g$ ) between each of the nine binary items. These estimates are not biased by sample overlap (even with complete sample overlap) (Bulik-Sullivan *et al.*, 2015a). We applied a Bonferroni-corrected significance threshold for these 36  $r_g$  tests ( $p < 1.39 \times 10^{-3}$ ). Additionally, we also calculated pairwise genetic correlations between our phenotypes (nine items and sum-scores) and 25 other psychiatric, substance use, socioeconomic and human traits with publicly available GWAS summary

**Table 2.** GWAS Results for binary PHQ-9 items and sum-score phenotypes

Gen. locus	SNP id	Chr	BP	A1	A2	Freq A1	Phenotype	$\beta$	s.e.	$p$ value	nSNPs	Nearest gene	eQTL genes
1	rs2279681	1	201861016	C	G	0.657	Sum-score	0.093	0.014	$5.60 \times 10^{-11}$	25	<i>SHISA4</i>	<i>SHISA4, LMOD1</i>
							Binary sum-score	0.023	0.004	$8.60 \times 10^{-09}$	6	<i>SHISA4</i>	
2	rs62158169	2	114081827	C	T	0.784	Sleep problems	0.015	0.002	$1.40 \times 10^{-10}$	22	<i>PAX8</i>	<i>FOXD4L1, CBWD2</i>
3	rs137997194	3	48824937	A	G	0.955	Depressed mood	-0.023	0.004	$7.70 \times 10^{-10}$	15	<i>PRKAR2A</i>	<i>AMT, KLHDC8B</i>
	rs143756010	3	49312248	C	T	0.956	Binary sum-score	-0.053	0.009	$7.40 \times 10^{-9}$	15	<i>C3orf62</i>	<i>AMT, NICN1, RNF123</i>
4	rs12492113	3	50521402	G	A	0.875	Binary sum-score	-0.037	0.006	$1.20 \times 10^{-10}$	114	<i>CACNA2D2</i>	<i>C3orf18, CACNA2D2, CYB561D2, DOCK3, HEMK1, HYAL1, LSMEM2, RBM6, MANF, MAPKAPK3</i>
							Sum-score	-0.124	0.020	$1.30 \times 10^{-09}$	76	<i>CACNA2D2</i>	
							Appetite changes	-0.013	0.002	$8.80 \times 10^{-09}$	28	<i>CACNA2D2</i>	
5	rs13127129	4	30501860	A	G	0.499	Sleep problems	0.010	0.002	$4.90 \times 10^{-08}$	1	<i>PCDH7</i>	-
6	rs7073667	10	107809043	T	C	0.452	Sleep problems	0.010	0.002	$4.30 \times 10^{-08}$	12	<i>RP11-298H24.1</i>	-
7	rs140920627	11	41577268	C	A	0.988	Anhedonia	-0.037	0.007	$3.70 \times 10^{-08}$	1	<i>RP11-124G5.3</i>	-
8	rs840161	12	57323523	A	G	0.366	Binary sum-score	-0.021	0.004	$3.90 \times 10^{-08}$	1	<i>SDR9C7</i>	-
9	rs2335859	19	7871837	T	G	0.602	Anhedonia	-0.009	0.001	$9.80 \times 10^{-09}$	10	<i>EXOSC3P2</i>	-

Table displays SNPs significant at  $p < 5 \times 10^{-8}$  and independent at  $r^2 < 0.10$ . Genomic risk loci (Gen. locus) are defined by  $r^2 < 0.10$ , window size 1000 kb. Chromosome (Chr), location in base pairs (BP) on Hg19, effect allele (A1), allele 2 (A2), frequency of effect allele (Freq A1), effect size  $\beta$ , standard error of  $\beta$  (s.e.),  $p$  value and number of SNPs clumped under lead SNP (nSNPs) are shown. Proximity (nearest gene) and eQTL (eQTL genes) mapping results are given. eQTL mapping limited to significant (FDR < 0.05) cis-eQTLs from GTEx v7 (Lonsdale et al., 2013) and the CommonMind Consortium (CMC) (Fromer et al., 2016). SNP rs137997194 and rs143756010 are tagging the same signal, but was the SNP with lowest  $p$  value in 'depressed mood' and 'binary sum-score', respectively.

statistics. Multiple testing was corrected for by adjusting *p* values based on false discovery rate (FDR) across all tests.

We calculated  $I^2$  heterogeneity statistics for each external trait (across the nine symptoms) to quantify the amount of variation in genetic correlations that is due to heterogeneity and not chance (Higgins and Thompson, 2002).  $I^2$  values range from 0% to 100% with higher values indicating a greater amount of variation is attributable to heterogeneity.

### Confirmatory factor analyses

Confirmatory factor analyses (CFAs) were conducted using genomic structural equation modelling (genomic SEM) (Grotzinger *et al.*, 2019) with a weighted least squares (WLS) estimator, in order to assess the genetic factor structure of the PHQ-9. The fit of three different factor structures commonly identified in phenotypic factor analyses was compared, including a one-factor model, a two-factor model containing ‘psychological’ and ‘somatic’ factors (Elhai *et al.*, 2012; Petersen *et al.*, 2015; Beard *et al.*, 2016), and a two-factor model containing ‘psychological/cognitive’ and ‘neurovegetative’ factors (Krause *et al.*, 2008; Krause *et al.*, 2010).

Model fit was evaluated with the following fit indices (and their commonly used thresholds for acceptable model fit): CFI ( $\geq 0.95$ ) and SRMR ( $\leq 0.06$ ) (Kline, 2005). Models were compared using AIC indices, which take into account both model fit and complexity. The most parsimonious model is the model with the lowest AIC value.

## Results

### Descriptive statistics

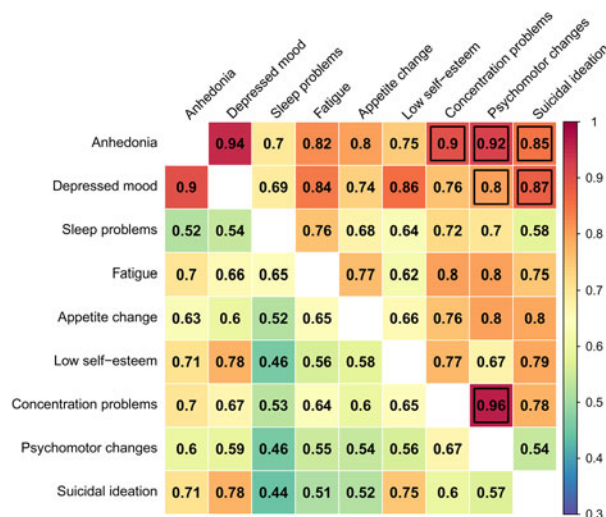
The final sample ( $N = 148\,752$ ) was 56% female, ranging in age from 46 to 80 years old ( $M = 63.88$ ,  $s.d. = 7.72$ ). The distribution of responses to all PHQ-9 items (on the ordinal scale) is displayed in online Supplementary Table S2. The distribution of item scores varied considerably across items; sleep problems and fatigue had the highest endorsement rates while suicidal ideation and psychomotor changes had the lowest rates. Sum-scores ranged from 0 to 27, with a mean of 2.71 ( $s.d. = 3.61$ ). Endorsement rates of binary items are shown in Table 1. The number of symptoms endorsed ranged from zero to nine, with a mean of 2.02 ( $s.d. = 2.20$ ).

### GWA analyses

GWA analyses of the nine binary items plus sum-score phenotypes identified a total of 326 genome-wide significant SNPs ( $p < 5 \times 10^{-8}$ ), tagged by 13 independent SNPs. Two lead SNPs were significant in more than one phenotype, such that across all phenotypes there are 11 unique, independent genome-wide significant SNPs. These SNPs mapped onto nine genomic risk loci (see Table 2 for results, online Supplementary Figs S2–S11 for QQ plots and Manhattan plots of all phenotypes; and online Supplementary Table S3 for the ordinal item GWAS results).

### Heritability estimates

The amount of variance explained by common SNPs (SNP-based heritability;  $h^2_{\text{SNP}}$ ) ranged from 6% of the variance in concentration problems up to 9% of the variance in appetite changes (mean  $h^2_{\text{SNP}}$  across the nine items was 7%; see Table 1).  $h^2_{\text{SNP}}$  estimates for the sum-score and binary sum-score phenotypes were 6% and



**Fig. 1.** Inter-item genetic and phenotypic correlations. Note: Genetic correlations ( $r_g$ ) above diagonal and phenotypic correlations ( $r_{net}$ ) below diagonal. Genetic correlations estimated using cross-trait LD Score Regression. All correlations are significantly different from zero at  $p < 1.39 \times 10^{-3}$ . Black squares indicate correlation is not significantly different from one (95% CI includes 1).

7%, respectively. All estimates were significant after Bonferroni correction ( $p < 4.55 \times 10^{-3}$ ; see online Supplementary Table S4).

### Inter-item phenotypic and genetic correlations

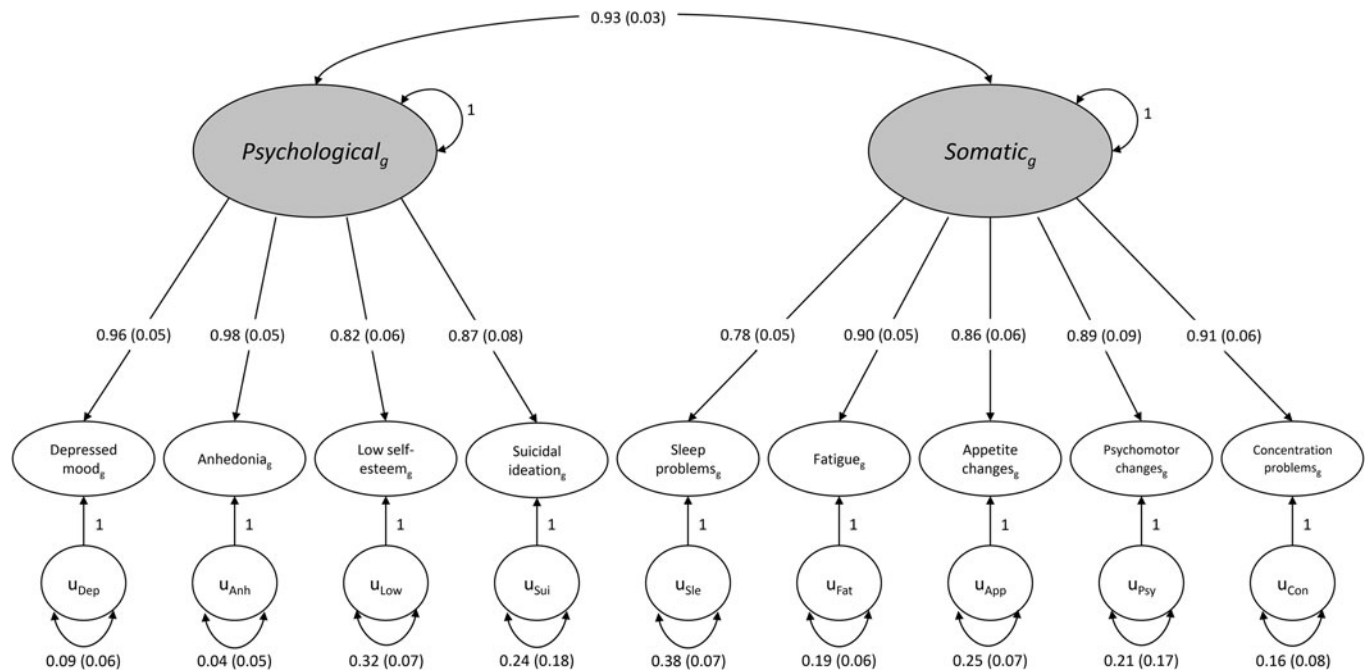
Tetrachoric correlations between all pairs of PHQ-9 binary items showed that all items were significantly and positively correlated with each other phenotypically. Coefficients ranged from 0.44 ( $s.e. = 0.007$ ) to 0.90 ( $s.e. = 0.002$ ), with the strongest association between anhedonia and depressed mood, the two core depressive symptoms (see Fig. 1).

Summary statistics from the GWASs of the nine binary items were used to calculate genetic correlations ( $r_g$ ) between items. All correlations were significant after correcting for multiple testing ( $p < 1.39 \times 10^{-3}$ ) (see Fig. 1). Estimated  $r_g$ 's ranged from 0.54 (suicidal ideation/psychomotor changes;  $s.e. = 0.15$ ) to 0.96 (psychomotor changes/concentration problems;  $s.e. = 0.11$ ), with a mean  $r_g$  of 0.77. Thirty out of the 36 genetic correlations were significantly less than one (95% CI did not include one), indicating genetic heterogeneity across the PHQ-9 items (see Fig. 1 and online Supplementary Table S5). Some of the genetic correlations that were not significantly different from 1 were relatively low, but have large standard errors which explain their overlap with 1.

A very similar pattern of genetic correlations emerged for the ordinal items (Pearson correlation between the set of binary item  $r_g$ 's and ordinal item  $r_g$ 's of  $r = 0.90$ ,  $p < 0.001$ ; see online Supplementary Fig. S12). The Pearson correlation between the genetic correlations and phenotypic correlations was moderate,  $r = 0.68$ ,  $p < 0.001$  (see online Supplementary Fig. S13).

### Confirmatory factor analyses

CFAs of the genetic factor structure found that all three models provided an adequate fit to the data (see online Supplementary Table S6). Comparison of models based on AIC values revealed that the two-factor model containing a ‘psychological’ factor and a ‘somatic’ factor provided the best fit (see Fig. 2).



**Fig. 2.** Factor loadings for a genetic two-factor model of the PHQ-9. *Note:* Standardised loadings estimated in Genomic SEM (using a WLS estimator) are presented for the best-fitting model.

### Genetic correlations with external traits

Genetic correlations of the nine PHQ-9 items, sum-score and binary sum-score with 25 other psychiatric, substance use, socio-economic and human traits are displayed in Fig. 3 (and online Supplementary Table S7). Individual items correlated as expected with closely related traits, supporting the validity of the individual symptom phenotypes in the present study. For example, appetite changes had a substantially stronger positive genetic correlation with body mass index ( $r_g = 0.61$ , *S.E.* = 0.03) than the other eight depression symptoms ( $r_g$ 's range between 0.10 and 0.29); and sleep problems had a strong, positive correlation with insomnia ( $r_g = 0.71$ , *S.E.* = 0.06). All symptoms were negatively correlated with subjective well-being ( $r_g$  range =  $-0.54$  to  $-0.91$ ), with suicidal ideation having the strongest association. Furthermore, all items positively correlated (and showed a similar pattern) with the other major depression and overall depression phenotypes.

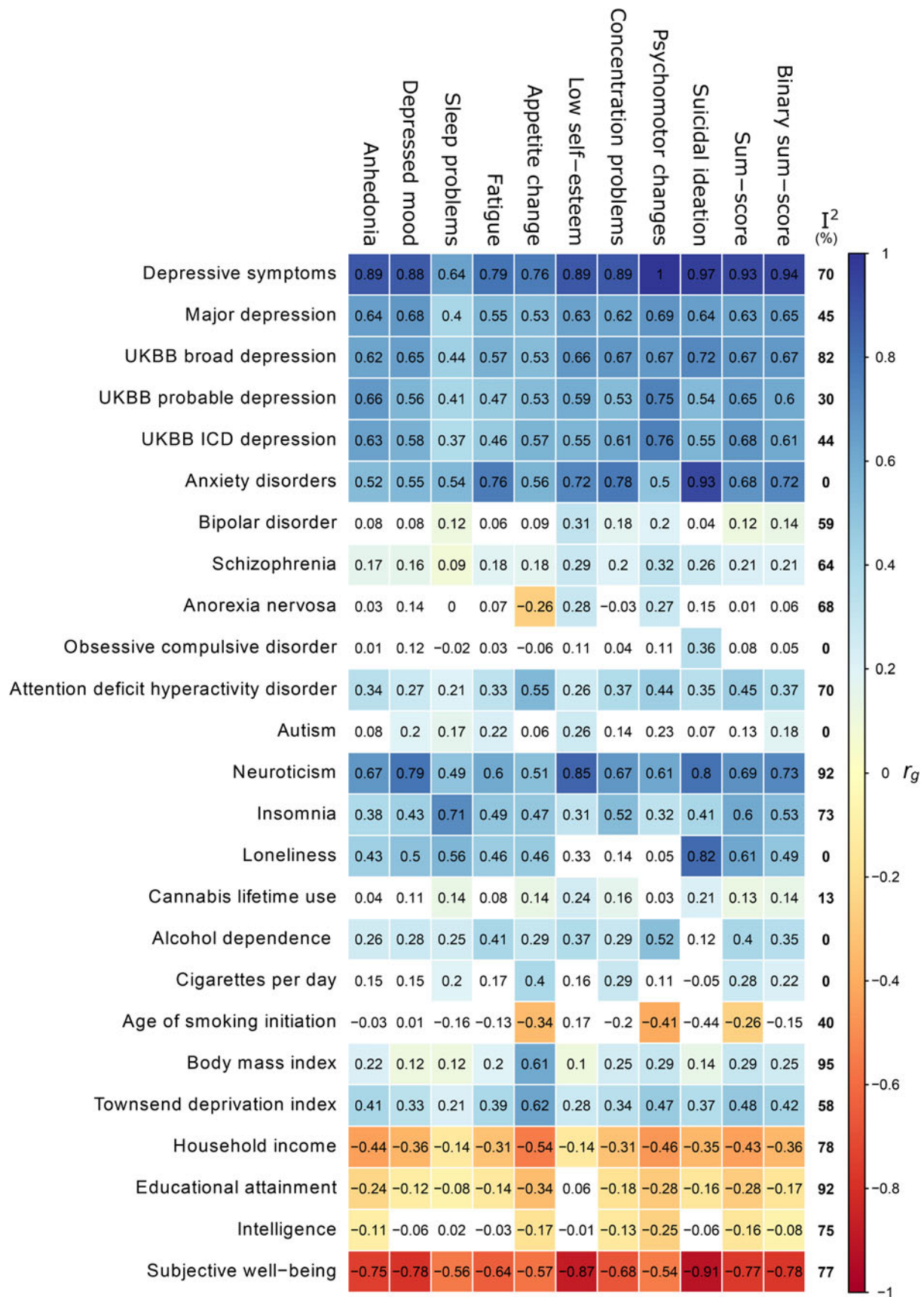
The pattern of genetic correlations with other psychiatric disorders and traits showed substantial variation across symptoms, such as with neuroticism ( $r_g$  range = 0.49–0.85;  $I^2 = 92\%$ ), schizophrenia ( $r_g$  range = 0.09–0.32,  $I^2 = 64\%$ ) and insomnia ( $r_g$  range = 0.31–0.71,  $I^2 = 72\%$ ).  $I^2$  values indicate this variation is largely due to heterogeneity and not error. Bipolar disorder was significantly correlated with only four out of nine depression items (sleep problems, low self-esteem, concentration problems and psychomotor changes;  $I^2 = 59\%$ ). Anorexia nervosa significantly overlapped with just three items ( $I^2 = 68\%$ ), with genetic correlations even being in different directions (low self-esteem  $r_g = 0.28$ , *S.E.* = 0.09; psychomotor changes  $r_g = 0.27$ , *S.E.* = 0.13; appetite change  $r_g = -0.26$ , *S.E.* = 0.08).

### Discussion

In the present study, we investigated genetic heterogeneity in depressive symptoms by conducting genetic analyses on individual items of the PHQ-9 in 148 752 participants from the UKBB.

We identified nine genomic risk loci across the nine depressive symptoms and sum-score phenotypes. One locus (locus 3, tagged by lead SNP rs143756010) was identified in a recent GWAS of depression (Howard *et al.*, 2019). The other eight loci have not been associated with depression in previous GWASs (Kohli *et al.*, 2011; Hek *et al.*, 2013; Cai *et al.*, 2015; Hyde *et al.*, 2016; Okbay *et al.*, 2016; Direk *et al.*, 2017; Power *et al.*, 2017; Hall *et al.*, 2018; Howard *et al.*, 2018; Li *et al.*, 2018; Wray *et al.*, 2018; Howard *et al.*, 2019), illustrating the importance of exploring genetic associations for specific symptoms. Our results revealed genetic heterogeneity in depressive symptoms with no overlap in significant loci across PHQ items. Though we acknowledge that the lack of overlap may be due to low statistical power to detect all true associations, we highlight some notable examples where a specific depressive symptom is linked to a gene that was previously found to be associated with a strongly related phenotype. For the item 'sleep problems', we found SNPs that implicate *PAX8* (based on proximity), a transcription factor related to thyroid follicular cell development and expression of thyroid-specific genes, replicating previous studies linking this gene to sleep duration (Gottlieb *et al.*, 2015; Jones *et al.*, 2016; Lane *et al.*, 2017). In addition, SNPs associated with 'depressed mood' influenced the expression of *KLHDC8B* (protein coding gene involved in cytokinesis). This gene has been previously linked to depressed affect, a sub-cluster of neuroticism that is strongly related to depression (Nagel *et al.*, 2018a).

Genetic correlations between depressive symptoms ranged from moderate ( $r_g < 0.60$ ) to high ( $r_g > 0.90$ ), suggesting that while some symptoms have high genetic overlap, a considerable amount of genetic variation is not shared between symptoms. This suggests genetic heterogeneity in symptoms of depression, in line with the finding that depression represents multiple dimensions of genetic risk (Kendler *et al.*, 2013) and previous associations between individual symptoms and specific polymorphisms (Myung *et al.*, 2012).



**Fig. 3.** Genetic correlations between PHQ-9 items and a range of other complex traits (psychiatric, substance use and socioeconomic phenotypes). Note: White squares indicate correlations do not meet significance after correcting for FDR.  $I^2$  values represent the amount of variation in genetic correlations across the nine symptoms attributable to heterogeneity and not chance.

The underlying genetic structure between symptoms was best explained by two genetic factors. While these factors were highly correlated, this suggests there are risk factors specific to clusters which could indicate underlying biology specific to either ‘psychological’ or ‘somatic’ symptoms of depression. This is consistent with symptoms differing in their biological correlates, with somatic symptoms such as weight gain, increased appetite and sleep problems being associated with higher levels of inflammation markers (Motivala *et al.*, 2005; Lamers *et al.*, 2013). These clusters were not in full agreement with the three genetic factors found by Kendler *et al.* (2013) based on an analysis of twin data. As an example, in Kendler *et al.*'s study, suicidal ideation loaded onto the same factor as psychomotor changes and concentration problems, while we find that suicidal ideation clusters with symptoms of depressed mood, anhedonia and low self-esteem that together form a ‘psychological symptoms’ factor. However, results are not easily comparable given that they derived factors from a twin study (and therefore captured rare genetic variants as well as common SNPs), used a subset of eight symptoms (appetite changes did not load onto any factor) and symptom phenotypes came from structured clinical interview of lifetime major depression rather than a self-report measure such as the PHQ-9.

Results from genetic correlations between items and a range of external traits lead us to note three general observations. First, genetic correlations with external traits differed substantially between symptoms (with  $I^2$  statistics suggesting this is largely due to heterogeneity), providing evidence for genetic heterogeneity in depressive symptoms. In agreement with previous findings for major depression (Howard *et al.*, 2018; Wray *et al.*, 2018), all symptoms overlapped with anxiety, schizophrenia, ADHD, insomnia, neuroticism and subjective well-being; however, the proportion of overlap varied considerably across symptoms. Second, some traits (such as bipolar disorder, cannabis lifetime use and intelligence) were significantly genetically correlated with a subset of items only. Bipolar disorder, for example, was significantly correlated with only four items (low self-esteem, concentration, psychomotor changes and sleep problems), suggesting that the moderate genetic overlap between bipolar and depression (Wray *et al.*, 2018) may be predominately driven by these selected symptoms. This highlights how insight into the genetic architecture of complex traits can be gained from conducting symptom-level analyses. Third, we found traits that were genetically correlated with individual items, but not with the sum-score phenotypes. Anorexia nervosa did not overlap with aggregate measures of depression symptoms as operationalised in the sum-score phenotypes, in agreement with Howard *et al.* (2018) who similarly found no genetic overlap between anorexia and their three overall depression phenotypes. Yet, anorexia nervosa was genetically correlated with appetite change, low self-esteem and psychomotor changes. Interestingly, the overlap with appetite change was in the opposite direction than the other two items, again emphasising the importance of analysing individual symptoms of a disorder, as important information is ignored by relying on sum-scores or overall phenotypes.

### Limitations

The findings and conclusions of this study should be interpreted in view of some key limitations. First, despite having the largest sample available to date, the current study is still underpowered to detect significant SNPs. Given the relatively high prevalence of depression, much larger sample sizes are needed compared

to other psychiatric disorders (Wray *et al.*, 2012). To not reduce power further we did not correct for multiple testing (of 11 GWA analyses) and hence our GWAS results require independent replication. Second, depression items were analysed in isolation, regardless of the overall MDD status of the participant. For example, a participant could strongly endorse the symptom fatigue, yet have no other signs of depression and hence the endorsement of fatigue is unrelated to major depression. Nevertheless, it is possible that fatigue, regardless of the context it occurs in, possesses the same underlying genetic basis. Third, we used a PHQ-9 cut-off score of 1 to dichotomise items in order to maximise the number of cases and improve statistical power. A PHQ-9 item score of one does not meet the diagnostic criteria for endorsement, hence the phenotypes may represent a predisposition rather than full endorsement of the particular symptom. Fourth, our results may be affected by ascertainment bias due to healthy volunteerism within the UKBB. As such our sample could represent a truncated version of the population's genetic distribution for symptoms (people on the far end of liability scale may be less likely to participate), hence resulting in reduced number of cases for some symptoms or reduced variation between cases and controls.

### Implications

The recent success in the discovery of genetic variants associated with depression has been driven by ever increasing sample sizes, an approach that has been favoured over reducing phenotypic heterogeneity. Consequently, GWASs have been conducted on a diverse range of depression-related phenotypes that often include a small subset of symptoms, generally with the view that the increase in sample size can overcome the lack of clinical precision. While this has indeed been proven to be effective at increasing the number of significant variants identified, our finding of symptom-level genetic heterogeneity raises questions about this approach. Using broad diagnostic phenotypes ignores the unique genetic factors associated with specific symptoms of depression that would likely provide useful information to further unravel the genetic architecture of depression. Further, our finding of genetic heterogeneity across depressive symptoms implicates that individuals with depression may show variation in disease pathogenesis. This variation may be linked to response to clinical interventions, such that patients presenting with specific symptom patterns (e.g. characterised primarily by somatic symptoms) may be expected to respond differently.

### Conclusion

Our results provide evidence that current self-reported depressive symptoms are genetically heterogeneous, and highlight the utility of analysing the genetics of individual items or symptoms. Insights into the genetic aetiology and underlying biology of depression will be maximised by combining large-scale genetic studies of broad clinical definitions with follow-up studies of more refined phenotypic measures of specific diagnostic subtypes. Future studies should investigate to what extent genetic heterogeneity in depressive symptoms is recapitulated in clinical symptoms of depression.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/S0033291719002526>



**Acknowledgements.** This work was conducted using the UK Biobank Resource (application number 25331). The UK Biobank was established by the Wellcome Trust medical charity, Medical Research Council (UK), Department of Health (UK), Scottish Government and Northwest Regional Development Agency. It also had funding from the Welsh Assembly Government, British Heart Foundation and Diabetes UK.

**Financial support.** S.M. is supported by a National Health and Medical Research Council (NHMRC) Fellowship. A.T.M. is supported by the Foundation Volksbond Rotterdam.

**Conflict of interest.** None.

## References

- American Psychiatric Association** (2013) *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. Arlington, VA: American Psychiatric Publishing.
- Beard C, Hsu KJ, Rifkin LS, Busch AB and Björgvinsson T** (2016) Validation of the PHQ-9 in a psychiatric sample. *Journal of Affective Disorders* **193**, 267–273.
- Bulik-Sullivan BK, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, Reprogen C, Psychiatric Genomics C, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control C, Duncan L, Perry JRB, Patterson N, Robinson EB, Daly MJ, Price AL and Neale BM** (2015a) An atlas of genetic correlations across human diseases and traits. *Nature Genetics* **47**, 1236.
- Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Patterson N, Daly MJ, Price AL and Neale BM** (2015b) LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics* **47**, 291.
- Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, Mcvean G, Leslie S, Allen N, Donnelly P and Marchini J** (2018) The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209.
- Cai N, Bigdeli TB, Kretzschmar W, Li Y, Liang J, Song L, Hu J, Li Q, Jin W and Hu Z** (2015) Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588.
- Christian O, Stefan MG, Brenda WP, Carmine MP, Amit E, Maurizio F, David CM and Alan FS** (2016) Major depressive disorder. *Nature Reviews Disease Primers* **2**.
- Direk N, Williams S, Smith JA, Ripke S, Air T, Amare AT, Amin N, Baune BT, Bennett DA, Blackwood DHR, Boomsma D, Breen G, Buttenschøn HN, Byrne EM, Børglum AD, Castelao E, Cichon S, Clarke T-K, Cornelis MC, Dannlowski U, De Jager PL, Demirkan A, Domenici E, Van Duijn CM, Dunn EC, Eriksson JG, Esko T, Faul J. D, Ferrucci L, Fornage M, De Geus E, Gill M, Gordon SD, Grabe H. J, Van Grootheest G, Hamilton SP, Hartman CA, Heath AC, Hek K, Hofman A, Homuth G, Horn C, Jan Hottenga J, Kardia SLR, Kloiber S, Koenen K, Kutalik Z, Ladwig K-H, Lahti J, Levinson DF, Lewis CM, Lewis G, Li QS, Llewellyn DJ, Lucae S, Lunetta KL, Macintyre DJ, Madden P, Martin NG, Mcintosh AM, Metspalu A, Milaneschi Y, Montgomery GW, Mors O, Mosley TH, Murabito J M, Müller-Myhsok B, Nöthen MM, Nyholt DR, O'donovan MC, Penninx BW, Pergadia ML, Perlis R, Potash JB, Preisig M, Purcell S M, Quiroz JA, Rääkkönen K, Rice JP, Rietschel M, Rivera M, Schulze TG, Shi J, Shyn S, Sinnamón GC, Smit JH, Smoller JW, Snieper H, Tanaka T, Tansey KE, Teumer A, Uher R, Umbricht D, Van Der Auwera S, Ware EB, Weir DR, Weissman MM, Willemsen G, Yang J, Zhao W, Tiemeier H and Sullivan PF** (2017) An analysis of two genome-wide association meta-analyses identifies a new locus for broad depression phenotype. *Biological Psychiatry* **82**, 322–329.
- Ehali JD, Contractor AA, Tamburrino M, Fine TH, Prescott MR, Shirley E, Chan PK, Slembariski R, Liberzon I, Galea S and Calabrese JR** (2012) The factor structure of major depression symptoms: a test of four competing models using the Patient Health Questionnaire-9. *Psychiatry Research* **199**, 169–173.
- Fried EI and Nesse RM** (2015) Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR\*D study. *Journal of Affective Disorders* **172**, 96–102.
- Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, Ruderfer DM, Oh EC, Topol A, Shah HR, Klei LL, Kramer R, Pinto D, Gümüş ZH, Cicek AE, Dang KK, Browne A, Lu C, Xie L, Readhead B, Stahl EA, Xiao J, Parvizi M, Hamamsy T, Fullard JF, Wang Y-C, Mahajan MC, Derry JMJ, Dudley JT, Hemby SE, Logsdon BA, Talbot K, Raj T, Bennett DA, De Jager PL, Zhu J, Zhang B, Sullivan PF, Chess A, Purcell SM, Shinobu LA, Mangravite LM, Toyoshiba H, Gur RE, Hahn C-G, Lewis DA, Haroutunian V, Peters MA, Lipska BK, Buxbaum JD, Schadt EE, Hirai K, Roeder K, Brennand KJ, Katsanis N, Domenici E, Devlin B and Sklar P** (2016) Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nature Neuroscience* **19**, 1442.
- Gottlieb DJ, Hek K, Chen TH, Watson NF, Eiriksdottir G, Byrne EM, Cornelis M, Warby SC, Bandinelli S, Cherkas L, Evans DS, Grabe HJ, Lahti J, Li M, Lehtimäki T, Lumley T, Marciani KD, Perusse L, Psaty BM, Robbins J, Tranah GJ, Vink JM, Wilk JB, Stafford JM, Bellis C, Biffar R, Bouchard C, Cade B, Curhan GC, Eriksson JG, Ewert R, Ferrucci L, Fulop T, Gehrman PR, Goodloe R, Harris TB, Heath AC, Hernandez D, Hofman A, Hottenga JJ, Hunter DJ, Jensen MK, Johnson AD, Kahonen M, Kao L, Kraft P, Larkin EK, Lauderdale DS, Luik AI, Medici M, Montgomery GW, Palotie A, Patel SR, Pistis G, Porcu E, Quaye L, Raitakari O, Redline S, Rimm EB, Rotter JJ, Smith AV, Spector TD, Teumer A, Uitterlinden AG, Vohl MC, Widen E, Willemsen G, Young T, Zhang X, Liu Y, Blangero J, Boomsma DI, Gudnason V, Hu F, Mangino M, Martin NG, O'connor GT, Stone KL, Tanaka T, Viikari J, Gharib SA, Punjabi NM, Raikonen K, Volzke H, Mignot E and Tiemeier H** (2015) Novel loci associated with usual sleep duration: the CHARGE Consortium Genome-Wide Association Study. *Molecular Psychiatry* **20**, 1232–1239.
- Grotzinger AD, Rhemtulla M, De Vlaming R, Ritchie SJ, Mallard TT, Hill WD, Ip HF, Marioni RE, Mcintosh AM, Deary IJ, Koellinger PD, Harden KP, Nivard MG and Tucker-Drob EM** (2019) Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour* **3**, 513–525.
- Hall LS, Adams MJ, Arnau-Soler A, Clarke T-K, Howard DM, Zeng Y, Davies G, Hagenaars SP, Maria Fernandez-Pujals A, Gibson J, Wigmore EM, Boutin TS, Hayward C, Scotland G, Major Depressive Disorder Working Group of the Psychiatric Genomics C, Porteous DJ, Deary IJ, Thomson PA, Haley CS and Mcintosh AM** (2018) Genome-wide meta-analyses of stratified depression in Generation Scotland and UK Biobank. *Translational Psychiatry* **8**, 9.
- Hek K, Demirkan A, Lahti J, Terracciano A, Teumer A, Cornelis MC, Amin N, Bakshis E, Baumert J, Ding J, Liu Y, Marciani K, Meirelles O, Nalls MA, Sun YV, Vogelzangs N, Yu L, Bandinelli S, Benjamin EJ, Bennett DA, Boomsma D, Cannas A, Coker LH, De Geus E, De Jager PL, Diez-Roux AV, Purcell S, Hu FB, Rimma EB, Hunter DJ, Jensen MK, Curhan G, Rice K, Penman AD, Rotter JJ, Sotoodehnia N, Emeny R, Eriksson JG, Evans DA, Ferrucci L, Fornage M, Gudnason V, Hofman A, Illig T, Kardia S, Kelly-Hayes M, Koenen K, Kraft P, Kuningas M, Massaro JM, Melzer D, Mulas A, Mulder CL, Murray A, Oostra BA, Palotie A, Penninx B, Petersmann A, Pilling LC, Psaty B, Rawal R, Reiman EM, Schulz A, Shulman JM, Singleton AB, Smith AV, Sutin AR, Uitterlinden AG, Volzke H, Widen E, Yaffe K, Zonderman AB, Cucca F, Harris T, Ladwig KH, Llewellyn DJ, Raikonen K, Tanaka T, Van Duijn CM, Grabe HJ, Launer LJ, Lunetta KL, Mosley Jr TH, Newman AB, Tiemeier H and Murabito J** (2013) A genome-wide association study of depressive symptoms. *Biological Psychiatry* **73**, 667–678.
- Hettema JM, Neale MC, Myers JM, Prescott CA and Kendler KS** (2006) A population-based twin study of the relationship between neuroticism and internalizing disorders. *American Journal of Psychiatry* **163**, 857–864.
- Higgins JP and Thompson SG** (2002) Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* **21**, 1539–1558.
- Howard DM, Adams MJ, Shirali M, Clarke T-K, Marioni RE, Davies G, Coleman JRI, Alloza C, Shen X, Barbu MC, Wigmore EM, Gibson J,**

- Hagenaars SP, Lewis CM, Ward J, Smith DJ, Sullivan PF, Haley CS, Breen G, Deary IJ and McIntosh AM (2018) Genome-wide association study of depression phenotypes in UK Biobank identifies variants in excitatory synaptic pathways. *Nature Communications* 9, 1470.
- Howard DM, Adams MJ, Clarke T-K, Hafferty JD, Gibson J, Shirali M, Coleman JRI, Hagenaars SP, Ward J, Wigmore EM, Alloza C, Shen X, Barbu MC, Xu EY, Whalley HC, Marioni RE, Porteous DJ, Davies G, Deary IJ, Hemani G, Berger K, Teismann H, Rawal R, Arold V, Baune BT, Dannlowski U, Domschke K, Tian C, Hinds DA, Trzaskowski M, Byrne EM, Ripke S, Smith DJ, Sullivan PF, Wray NR, Breen G, Lewis CM, McIntosh AM, 23andme Research Team and Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium (2019) Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience* 22, 343–352.
- Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, Tung JY, Hinds DA, Perlis RH and Winslow AR (2016) Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nature Genetics* 48, 1031.
- Jang KL, Livesley WJ, Taylor S, Stein MB and Moon EC (2004) Heritability of individual depressive symptoms. *Journal of Affective Disorders* 80, 125–133.
- Jones SE, Tyrrell J, Wood AR, Beaumont RN, Ruth KS, Tuke MA, Yaghootkar H, Hu Y, Teder-Laving M, Hayward C, Roenneberg T, Wilson JF, Del Greco F, Hicks AA, Shin C, Yun C-H, Lee SK, Metspalu A, Byrne EM, Gehrman PR, Tiemeier H, Allebrandt KV, Freathy RM, Murray A, Hinds DA, Frayling TM and Weedon MN (2016) Genome-Wide Association Analyses in 128,266 Individuals Identifies New Morningness and Sleep Duration Loci. *PLOS Genetics* 12, e1006125.
- Kendler KS, Aggen SH and Neale MC (2013) Evidence for multiple genetic factors underlying dsm-iv criteria for major depression. *JAMA Psychiatry* 70, 599–607.
- Kline RB (2005) *Principles and Practice of Structural Equation Modeling*, 2nd ed. New York, NY, US: Guilford Press.
- Kohli MA, Lucae S, Saemann PG, Schmidt MV, Demirkan A, Hek K, Czamara D, Alexander M, Salyakina D, Ripke S, Hoehn D, Specht M, Menke A, Hennings J, Heck A, Wolf C, Ising M, Schreiber S, Czisch M, Müller MB, Uhr M, Bettecken T, Becker A, Schramm J, Rietschel M, Maier W, Bradley B, Ressler KJ, Nöthen MM, Cichon S, Craig IW, Breen G, Lewis CM, Hofman A, Tiemeier H, Van duijn CM, Holsboer F, Müller-Myhsok B and Binder EB (2011) The neuronal transporter gene SLC6A15 confers risk to major depression. *Neuron* 70, 252–265.
- Krause JS, Bombardier C and Carter RE (2008) Assessment of depressive symptoms during inpatient rehabilitation for spinal cord injury: is there an underlying somatic factor when using the PHQ? *Rehabilitation Psychology* 53, 513.
- Krause JS, Reed KS and Mcardle JJ (2010) Factor structure and predictive validity of somatic and nonsomatic symptoms from the Patient Health Questionnaire-9: a longitudinal study after spinal cord injury. *Archives of Physical Medicine and Rehabilitation* 91, 1218–1224.
- Kroenke K, Spitzer RL and Williams JBW (2001a) The PHQ-9. *Journal of General Internal Medicine* 16, 606–613.
- Kroenke K, Spitzer RL and Williams JBW (2001b) The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine* 16, 606–613.
- Lamers F, Vogelzangs N, Merikangas KR, De Jonge P, Beekman AT and Penninx BW (2013) Evidence for a differential role of HPA-axis function, inflammation and metabolic syndrome in melancholic versus atypical depression. *Molecular Psychiatry* 18, 692–699.
- Lane JM, Liang J, Vlasac I, Anderson SG, Bechtold DA, Bowden J, Emsley R, Gill S, Little MA, Luik AI, Loudon A, Scheer FAJL, Purcell SM, Kyle SD, Lawlor DA, Zhu X, Redline S, Ray DW, Rutter MK and Saxena R (2017) Genome-wide association analyses of sleep disturbance traits identify new loci and highlight shared genetics with neuropsychiatric and metabolic traits. *Nature Genetics* 49, 274–281.
- Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH, Mowry BJ, Thapar A, Goddard ME and Witte JS (2013) Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics* 45, 984.
- Levinson DF, Mostafavi S, Milaneschi Y, Rivera M, Ripke S, Wray NR and Sullivan PF (2014) Genetic studies of major depressive disorder: why are there no genome-wide association study findings and what can we do about it? *Biological Psychiatry* 76, 510–512.
- Li X, Luo Z, Gu C, Hall LS, McIntosh AM, Zeng Y, Porteous DJ, Hayward C, Li M, Yao Y-G, Zhang C, Luo X-J and The 23andme Research Team (2018) Common variants on 6q16.2, 12q24.31 and 16p13.3 are associated with major depressive disorder. *Neuropsychopharmacology* 43, 2146–2153.
- Loh P-R, Tucker G, Bulik-Sullivan BK, Vilhjálmsson BJ, Finucane HK, Salem RM, Chasman DI, Ridker PM, Neale BM, Berger B, Patterson N and Price AL (2015) Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nature Genetics* 47, 284.
- Loh P-R, Kichaev G, Gazal S, Schoech AP and Price AL (2018) Mixed-model association for biobank-scale datasets. *Nature Genetics* 50, 906–908.
- Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, Hasz R, Walters G, Garcia F, Young N, Foster B, Moser M, Karasik E, Gillard B, Ramsey K, Sullivan S, Bridge J, Magazine H, Syron J, Fleming J, Siminoff L, Traino H, Mosavel M, Barker L, Jewell S, Rohrer D, Maxim D, Filkins D, Harbach P, Cortadillo E, Berghuis B, Turner L, Hudson E, Feenstra K, Sobin L, Robb J, Branton P, Korzeniewski G, Shive C, Tabor D, Qi L, Groch K, Nampally S, Buia S, Zimmerman A, Smith A, Burges R, Robinson K, Valentino K, Bradbury D, Cosentino M, Diaz-Mayoral N, Kennedy M, Engel T, Williams P, Erickson K, Ardlie K, Winckler W, Getz G, Deluca D, Macarthur D, Kellis M, Thomson A, Young T, Gelfand E, Donovan M, Meng Y, Grant G, Mash D, Marcus Y, Basile M, Liu J, Zhu J, Tu Z, Cox NJ, Nicolae DL, Gamazon ER, Im HK, Konkashbaev A, Pritchard J, Stevens M, Fluttre T, Wen X, Dermitzakis ET, Lappalainen T, Guigo R, Monlong J, Sammeth M, Koller D, Battle A, Mostafavi S, Mccarthy M, Rivas M, Maller J, Rusyn I, Nobel A, Wright F, Shabalin A, Feolo M, Sharopova N, Sturcke A, Paschal J, Anderson JM, Wilder EL, Derr LK, Green ED, Struewing JP, Temple G, Volpi S, Boyer JT, Thomson EJ, Guyer MS, Ng C, Abdallah A, Colantuoni D, Insel TR, Koester SE, Little AR, Bender PK, Lehner T, Yao Y, Compton CC, Vaught JB, Sawyer S, Lockhart NC, Demchok J and Moore HF (2013) The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* 45, 580.
- Motivala SJ, Sarfatti A, Olmos L and Irwin MR (2005) Inflammatory markers and sleep disturbance in major depression. *Psychosomatic Medicine* 67, 187–194.
- Myung W, Song J, Lim S-W, Won H-H, Kim S, Lee Y, Kang HS, Lee H, Kim J-W, Carroll BJ and Kim DK (2012) Genetic association study of individual symptoms in depression. *Psychiatry Research* 198, 400–406.
- Nagel M, Jansen PR, Stringer S, Watanabe K, De Leeuw CA, Bryois J, Savage JE, Hammerschlag AR, Skene NG, Muñoz-Manchado AB, Agee M, Alipanahi B, Auton A, Bell RK, Bryc K, Elson SL, Fontanillas P, Furlotte NA, Hinds DA, Hromatka BS, Huber KE, Kleinman A, Litterman NK, McIntyre MH, Mountain JL, Noblin ES, Northover C a M, Pitts SJ, Sathirapongsasuti JF, Sazonova OV, Shelton JF, Shringarpure S, Tian C, Tung JY, Vacic V, Wilson CH, White T, Tiemeier H, Linnarsson S, Hjerling-Leffler J, Polderman TJC, Sullivan PF, Van Der Sluis S, Posthuma D and 23andme Research Team (2018a) Meta-analysis of genome-wide association studies for neuroticism in 449, 484 individuals identifies novel genetic loci and pathways. *Nature Genetics* 50, 920–927.
- Nagel M, Watanabe K, Stringer S, Posthuma D and Van Der Sluis S (2018b) Item-level analyses reveal genetic heterogeneity in neuroticism. *Nature Communications* 9, 905.
- Okbay A, Baselmans BML, De Neve JE, Turley P, Nivard MG, Fontana MA, Meddens SFW, Linnér RK, Rietveld CA, Derringer J, Gratten J, Lee JJ, Liu JZ, De Vlaming R, Ahluwalia TS, Buchwald J, Cavadin A, Frazier-Wood AC, Furlotte NA, Garfield V, Geisel MH, Gonzalez JR, Haitjema S, Karlsson R, Van Der Laan SW, Ladwig K-H, Lahti J, Van Der Lee SJ, Lind PA, Liu T, Matteson L, Mihailov E, Miller MB, Minica CC, Nolte IM, Mook-Kanamori D, Van Der Most PJ,

- Oldmeadow C, Qian Y, Raitakari O, Rawal R, Realo A, Rueedi R, Schmidt B, Smith AV, Stergiakouli E, Tanaka T, Taylor K, Thorleifsson G, Wedenoja J, Wellmann J, Westra H-J, Willems SM, Zhao W, Lifelines Cohort S, Amin N, Bakshi A, Bergmann S, Bjornsdottir G, Boyle PA, Cherney S, Cox SR, Davies G, Davis OSP, Ding J, Direk N, Eibich P, Emery RT, Fatemifar G, Faul JD, Ferrucci L, Forstner AJ, Gieger C, Gupta R, Harris TB, Harris JM, Holliday EG, Hottenga J-J, De Jager PL, Kaakinen MA, Kajantie E, Karhunen V, Kolcic I, Kumari M, Launer LJ, Franke L, Li-Gao R, Liewald DC, Koini M, Loukola A, Marques-Vidal P, Montgomery GW, Mosing MA, Paternoster L, Pattie A, Petrovic KE, Pulikki-Räback L, Quaye L, Rääkkönen K, Rudan I, Scott RJ, Smith JA, Sutin AR, Trzaskowski M, Vinkhuyzen A E, Yu L, Zabaneh D, Attia JR, Bennett DA, Berger K, Bertram L, Boomsma DI, Snieder H, Chang S-C, Cucca F, Deary IJ, Van Duijn CM, Eriksson JG, Bültmann U, De Geus EJC, Groenen PJJ, Gudnason V, Hansen T, Hartman CA, Haworth CMA, Hayward C, Heath AC, Hinds DA, Hyytiäinen E, Iacono WG, Järvelin M-R, Jöckel K-H, Kaprio J, Kardina SLR, Keltikangas-Järvinen L, Kraft P, Kubzansky LD, Lehtimäki T, Magnusson PKE, Martin NG, McGue M, Metspalu A, Mills M, De Mutsert R, Oldehinkel AJ, Pasterkamp G, Pedersen NL, Plomin R, Polasek O, Power C, Rich SS, Rosendaal FR, Den Ruijter HM, Schlessinger D, Schmidt H, Svento R, Schmidt R, Alizadeh BZ, Sørensen TIA, Spector TD, Starr JM, Stefansson K, Steptoe A, Terracciano A, Thorsteinsdottir U, hurik AR, Timpson NJ, Tiemeier H, Uitterlinden AG, Vollenweider P, Wagner GG, Weir DR, Yang J, Conley DC, Smith GD, Hofman A, Johannesson M, Laibson DI, Medland SE, Meyer MN, Pickrell JK, Esko T, Krueger RF, Beauchamp JP, Koellinger PD, Benjamin DJ, Bartels M and Cesarini D (2016) Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics* 48, 624.
- Østergaard SD, Jensen SOW and Bech P (2011) The heterogeneity of the depressive syndrome: when numbers get serious. *Acta Psychiatrica Scandinavica* 124, 495–496.
- Petersen JJ, Paulitsch MA, Hartig J, Mergenthal K, Gerlach FM and Gensichen J (2015) Factor structure and measurement invariance of the Patient Health Questionnaire-9 for female and male primary care patients with major depression in Germany. *Journal of Affective Disorders* 170, 138–142.
- Power RA, Tansey KE, Buttenschön HN, Cohen-Woods S, Bigdeli T, Hall LS, Katalik Z, Lee SH, Ripke S, Steinberg S, Teumer A, Viktorin A, Wray NR, Arolt V, Baune BT, Boomsma DI, Borglum AD, Byrne EM, Castelao E, Craddock N, Craig IW, Dannowski U, Deary IJ, Degenhardt F, Forstner AJ, Gordon SD, Grabe HJ, Grove J, Hamilton SP, Hayward C, Heath AC, Hocking LJ, Homuth G, Hottenga JJ, Kloiber S, Krogh J, Landén M, Lang M, Levinson DF, Lichtenstein P, Lucae S, Macintyre DJ, Madden P, Magnusson PKE, Martin NG, Mcintosh AM, Middeldorp CM, Milaneschi Y, Montgomery GW, Mors O, Müller-Myhsok B, Nyholt DR, Oskarsson H, Owen MJ, Padmanabhan S, Penninx BWJH, Pergadia ML, Porteous DJ, Potash JB, Preisig M, Rivera M, Shi J, Shyn SI, Sigurdsson E, Smit JH, Smith BH, Stefansson H, Stefansson K, Strohmaier J, Sullivan PF, Thomsson P, Thorgeirsson TE, Van Der Auwera S, Weissman MM, Converge Consortium CCGC, Breen G and Lewis CM (2017) Genome-wide association for major depression through age at onset stratification: major depressive disorder working group of the psychiatric genomics consortium(). *Biological Psychiatry* 81, 325–335.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ and Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* 81, 559–575.
- Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA, Lin D-Y, Duan J, Ophoff RA and Andreassen OA (2011) Genome-wide association study identifies five new schizophrenia loci. *Nature Genetics* 43, 969.
- Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G, Byrne EM, Blackwood DH, Boomsma DI, Cichon S, Heath AC, Holsboer F, Lucae S, Madden PA, Martin NG, McGuffin P, Muglia P, Nothen MM, Penninx BP, Pergadia ML, Potash JB, Rietschel M, Lin D, Muller-Myhsok B, Shi J, Steinberg S, Grabe HJ, Lichtenstein P, Magnusson P, Perlis RH, Preisig M, Smoller JW, Stefansson K, Uher R, Katalik Z, Tansey KE, Teumer A, Viktorin A, Barnes MR, Bettecken T, Binder EB, Breuer R, Castro VM, Churchill SE, Coryell WH, Craddock N, Craig IW, Czamara D, De Geus EJ, Degenhardt F, Farmer AE, Fava M, Frank J, Gainer VS, Gallagher PJ, Gordon SD, Goryachev S, Gross M, Guipponi M, Henders AK, Herms S, Hickie IB, Hoefels S, Hoogendijk W, Hottenga JJ, Iosifescu DV, Ising M, Jones I, Jones L, Jung-Ying T, Knowles JA, Kohane IS, Kohli MA, Korszun A, Landen M, Lawson WB, Lewis G, Macintyre D, Maier W, Mattheisen M, McGrath PJ, Mcintosh A, Mclean A, Middeldorp CM, Middleton L, Montgomery GM, Murphy SN, Nauck M, Nolen WA, Nyholt DR, O'donovan M, Oskarsson H, Pedersen N, Scheftner WA, Schulz A, Schulze TG, Shyn SI, Sigurdsson E, Slager SL, Smit JH, Stefansson H, Steffens M, Thorgeirsson T, Tozzi F, Treutlein J, Uhr M, Van Den Oord EJ, Van Grootheest G, Volzke H, Weibull JB, Willemsen G, Zitman FG, Neale B, Daly M, Levinson DF and Sullivan PF (2013) A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular Psychiatry*, 18, 497–511.
- Ripke S, Neale BM, Corvin A, Walters JT, Farh K-H, Holmans PA, Lee P, Bulik-Sullivan B, Collier DA and Huang H (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511, 421.
- Stewart AL, Hays RD and Ware Jr JE (1988) The MOS short-form general health survey. Reliability and validity in a patient population. *Medical Care* 26, 724–735.
- Sullivan PF, Neale MC and Kendler KS (2000) Genetic epidemiology of major depression: review and meta-analysis. *American Journal of Psychiatry* 157, 1552–1562.
- Watanabe K, Taskesen E, Bochoven A and Posthuma D (2017) Functional mapping and annotation of genetic associations with FUMA. *Nature Communications* 8, 1826.
- Wray NR, Pergadia ML, Blackwood DH, Penninx BW, Gordon SD, Nyholt DR, Ripke S, Macintyre DJ, McGhee KA, Maclean AW, Smit JH, Hottenga JJ, Willemsen G, Middeldorp CM, De Geus EJ, Lewis CM, McGuffin P, Hickie IB, Van Den Oord EJ, Liu JZ, Macgregor S, Mcevoy BP, Byrne EM, Medland SE, Statham DJ, Henders AK, Heath AC, Montgomery GW, Martin NG, Boomsma DI, Madden PA and Sullivan PF (2012) Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Molecular Psychiatry* 17, 36–48.
- Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, Adams MJ, Agerbo E, Air TM, Andlauer TMF, Bacanu S-A, Bækvad-Hansen M, Beekman AFT, Bigdeli TB, Binder EB, Blackwood DRH, Bryois J, Buttenschön HN, Bybjerg-Grauholm J, Cai N, Castelao E, Christensen E, Clarke T-K, Coleman JIR, Colodro-Conde L, Couvy-Duchesne B, Craddock N, Crawford GE, Crowley CA, Dashti HS, Davies G, Deary IJ, Degenhardt F, Derks EM, Direk N, Dolan CV, Dunn EC, Eley TC, Eriksson N, Escott-Price V, Kiadeh FHF, Finucane HK, Forstner AJ, Frank J, Gaspar HA, Gill M, Giusti-Rodríguez P, Goes FS, Gordon SD, Grove J, Hall LS, Hannon E, Hansen CS, Hansen TF, Herms S, Hickie IB, Hoffmann P, Homuth G, Horn C, Hottenga J-J, Hougaard DM, Hu M, Hyde CL, Ising M, Jansen R, Jin F, Jorgenson E, Knowles JA, Kohane IS, Kraft J, Kretschmar WW, Krogh J, Katalik Z, Lane JM, Li Y, Li Y, Lind PA, Liu X, Lu L, Macintyre DJ, Mackinnon DF, Maier RM, Maier W, Marchini J, Mbarek H, McGrath P, McGuffin P, Medland SE, Mehta D, Middeldorp CM, Mihailov E, Milaneschi Y, Milani L, Mill J, Mondimore FM, Montgomery GW, Mostafavi S, Mullins N, Nauck M, Ng B, Nivard MG, Nyholt DR, O'reilly PF, Oskarsson H, Owen MJ, Painter JN, Pedersen CB, Pedersen MG, Peterson RE, Pettersson E, Peyrot WJ, Pistic G, Posthuma D, Purcell SM, Quiroz JA, Qvist P, Rice JP, Riley BP, Rivera M, Saeed Mirza S, Saxena R, Schoevers R, Schulte EC, Shen L, Shi J, Shyn SI, Sigurdsson E, Sinnamoni GBC, Smit JH, Smith DJ, Stefansson H, Steinberg S, Stockmeier CA,

- Streit F, Strohmaier J, Tansey KE, Teismann H, Teumer A, Thompson W, Thomson PA, Thorgeirsson TE, Tian C, Traylor M, Treutlein J, Trubetsky V, Uitterlinden AG, Umbricht D, Van Der Auwera S, Van Hemert AM, Viktorin A, Visscher PM, Wang Y, Webb BT, Weinsheimer SM, Wellmann J, Willemsen G, Witt SH, Wu Y, Xi HS, Yang J, Zhang F, Arolt V, Baune BT, Berger K, Boomsma DI, Cichon S, Dannlowski U, De Geus ECJ, Depaulo JR, Domenici E, Domschke K, Esko T, Grabe HJ, Hamilton SP, Hayward C, Heath AC, Hinds DA, Kendler KS, Kloiber S, Lewis G, Li QS, Lucae S, Madden PFA, Magnusson PK, Martin NG, McIntosh AM, Metspalu A, Mors O, Mortensen PB, Müller-Myhsok B, Nordentoft M, Nöthen MM, O'donovan MC, Paciga SA, Pedersen NL, Penninx BWJH, Perlis RH, Porteous DJ, Potash JB, Preisig M, Rietschel M, Schaefer C, Schulze TG, Smoller JW, Stefansson K, Tiemeier H, Uher R, Völzke H, Weissman MM, Werge T, Winslow AR, Lewis CM, Levinson DF, Breen G, Børglum AD, Sullivan PF, Eqtllgen, 23andme and The Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium (2018) Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics* 50, 668–681.
- Zhou W, Nielsen JB, Fritsche LG, Dey R, Gabrielsen ME, Wolford BN, Lefaive J, Vandelaar P, Gagliano SA, Gifford A, Bastarache LA, Wei W-Q, Denny JC, Lin M, Hveem K, Kang HM, Abecasis GR, Willer CJ and Lee S (2018) Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nature Genetics* 50, 1335–1341.