Author’s Accepted Manuscript

An Analysis of Two Genome-Wide Association Meta-Analyses Identifies a New Locus for Broad Depression Phenotype A Susceptibility Locus for Broad Depression Phenotype

To appear in: Biological Psychiatry


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
An Analysis of Two Genome-Wide Association Meta-Analyses Identifies a New Locus for Broad Depression Phenotype

Nese Direk, MD, MSc, Stephanie Williams, ScM, Jennifer A. Smith, PhD, MPH, Stephan Ripke, MD, PhD, Tracy Air, BA(Hons), M.Biostatistics, Azmeraw T. Amare, MPH, MSc, Najaf Amin, PhD, Bernhard T. Baune, MD, PhD, MPH, David A. Bennett, MD, Douglas H.R. Blackwood, MD, PhD, Dorret Boomsma, PhD, Gerome Breen, PhD, Henriette N. Buttenschøn, PhD, Enda M. Byrne, PhD, Anders D. Børorglum, MD, PhD, Enrique Castelao, MSc, Sven Cichon, PhD, Philip L. De Jager, MD, PhD, Ayse Demirkan, PhD, Enrico Domenici, PhD, Cornelia M. van Duijn, PhD, Erin C. Dunn, ScD, MPH, Johan G. Eriksson, MD, DMSc, Tonu Esko, PhD, Michael C. O’Donovan, MD, PhD, Scott D. Gordon, PhD, Hans Jörgen Grabe, MD, Gerard van Grootheest, MSc, Steven P. Hamilton, MD, PhD, Catharina A. Hartman, PhD, Andrew C. Heath, DPhil, Karin Hek, PhD, Albert Hofman, MD, PhD, Georg Homuth, PhD, Carsten Horn, PhD, Jouke Jan Hottenga, PhD, Sharon L.R. Kardia, PhD, Stefan Kloiber, MD, Karestan Koenen, PhD, Zoltán Kutalik, PhD, Karl-Heinz Ladwig, MD, PhD, Jari Lahti, PhD, Douglas F. Levinson, MD, Cathryn M. Lewis, PhD, Glyn Lewis, PhD, Qingqin S Li, PhD, David J. Llewellyn, PhD, Susanne Lucae, MD, PhD, Kathryn L. Lunetta, PhD, Donald J. MacIntyre, MD, Pamela Madden, PhD, Nicholas G. Martin, PhD, Andrew M. McIntosh, MD, Andres Metspalu, MD, PhD, Yuri Milaneschi, PhD, Grant W. Montgomery, PhD, Ole Mors, PhD, Thomas H. Mosley Jr., PhD, Joanne M. Murabito, MD, Bertram Müller-Myshok, MD, Markus M. Nöthen, MD, PhD, Dale R. Nyholt, PhD, Michael C. O’Donovan, MD, PhD, Brenda W. Penninx, PhD, Michele L. Pergadia, PhD, Roy Perlis, MD, MSc, James B. Potash, MD, Martin Preisig, MD, Shaun M. Purcell, PhD, Jorge A. Quirroz, MD, Katri Räikkönen, PhD, John P. Rice, PhD, Marcella Rietschel, MD, PhD, Margarita Rivera, PhD, Thomas G. Schulze, MD, Jianxin Shi, PhD, Stanley Shyn, MD, PhD, Grant C. Sinnamon, PhD, Johannes H. Smit, PhD, Jordan W. Smoller, MD, ScD, Harold Snieder, PhD, Toshiko Tanaka, PhD, Catherine E. Tansey, PhD, Alexander Teumer, PhD, Rudolf Uher, MD, PhD, Daniel Umbricht, MD, Sandra Van der Auwera, Dipl. Biomatematikerin, Erin B. Ware, PhD, David R. Weir, PhD, Myrna M. Weissman, PhD, Gonneke Willemsen, PhD, Jingyun Yang, PhD, Wei Zhao, PhD, Henning Tiemeier, MD, PhD, Patrick F. Sullivan, MD, PhD.
1Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands.

2Department of Psychiatry, Dokuz Eylül University, Izmir, Turkey.

3Department of Genetics, University of North Carolina at Chapel Hill, NC, USA.

4Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA.

5Stanley Center for Psychiatric Research, The Broad Institute of Harvard and MIT, Cambridge, MA, USA.

6Analytic and Translational Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA.

7Department of Psychiatry and Psychotherapy, Charité, Campus Mitte, Berlin, Germany.

8Discipline of Psychiatry, School of Medicine, University of Adelaide, Adelaide, SA, Australia.

9Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands.

10Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands.

11Rush Alzheimer’s Disease Center & Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA.

12Division of Psychiatry, University of Edinburgh, UK.

13Department of Biological Psychology, VU University, Amsterdam, The Netherlands.

14MRC SGDP Centre, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, UK.

15Translational Neuropsychiatry Unit, Department of Clinical Medicine, Aarhus University, Denmark.

16The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Denmark.

17The University of Queensland, Queensland Brain Institute, St.Lucia, Queensland, Australia.

18Department of Biomedicine and Centre for Integrative Sequencing, iSEQ, Aarhus University, Denmark.

19Department of Psychiatry, Lausanne University Hospital, Switzerland.

20Institute of Human Genetics, University of Bonn, Bonn, Germany.

21Department of Genomics, Life & Brain Center, Bonn, Germany.

22Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich, Germany.

23Division of Medical Genetics, Department of Biomedicine, University of Basel, Basel, Switzerland.

24Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago IL, USA.

25Department of Psychiatry and Psychotherapy, University of Münster, Münster, Germany.
Department of Neurology, Program in Translational NeuroPsychiatric Genomics, Brigham and Women’s Hospital, Harvard Medical School, Boston, USA.

Harvard Medical School, Boston, Massachusetts, USA.

Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA.

Centre for Integrative Biology, University of Trento, Trento, Italy.

Roche Pharmaceutical Research and Early Development, Neuroscience, Ophthalmology and Rare Diseases Discovery & Translational Medicine Area, Roche Innovation Center Basel, F Hoffman-La Roche Ltd., Switzerland.

Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA.

Department of Psychiatry, Harvard Medical School, Boston, MA, USA.

National Institute for Health and Welfare, Department of Chronic Disease Prevention, Helsinki, Finland.

Department of General Practice and Primary Health Care, University of Helsinki, Finland.

Unit of General Practice, Helsinki University Central Hospital, Finland.

Folkhalsan Research Centre, Helsinki, Finland.

Vasa Central Hospital, Vasa, Finland.

Estonian Genome Center, University of Tartu, Tartu, Estonia.

Division of Endocrinology, Boston Children’s Hospital, Cambridge, MA, USA.

Program in Medical and Populational Genetics, Broad Institute, Cambridge, MA, USA.

Department of Genetics, Harvard Medical School, Boston, MA, USA.

Institute for Social Research, University of Michigan, Ann Arbor, MI, USA.

Translational Gerontology Branch, National Institute on Aging, Baltimore, MD, USA.

Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX, USA.

Department of Psychiatry, Trinity Centre for Health Science, Dublin, Ireland.

QIMR Berghofer Medical Research Institute, Brisbane, Australia.

Department of Psychiatry and Psychotherapy, Helios Hospital Stralsund, Germany.

Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Germany.

German Center for Neurodegenerative Diseases (DZNE), Site Rostock/ Greifswald, Germany.

Department of Psychiatry, Neuroscience Campus Amsterdam and EMGO Institute of Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands.

Department of Psychiatry, Kaiser Permanente San Francisco Medical Center, CA, USA.
Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands.

Department of Psychiatry, Washington University St.Louis, Missouri, USA.

Department of Psychiatry, Erasmus University Medical Center, Rotterdam, The Netherlands.

Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Germany.

Max Planck Institute of Psychiatry, Munich, Germany.

Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, USA.

Institute of Social and Preventive Medicine (IUMSP), Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland.

Institute of Epidemiology II, Mental Health Research Unit, Helmholtz Zentrum München, German Research Center for Environmental Health, Germany.

Department of Psychosomatic Medicine and Psychotherapy, Klinikum rechts der Isar, Technische Universität München, Munich, Germany.

Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland.

Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA.

Division of Psychiatry, University College London, London, UK.

Janssen Research & Development, LLC.

University of Exeter Medical School, Exeter, UK.

Boston University School of Public Health, Department of Biostatistics, Boston, MA, USA.

Boston University and NHLBI's Framingham Heart Study, Framingham, MA, USA.

Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia.

Research Department P, Aarhus University Hospital, Risskov, Denmark.

Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA.

Boston University School of Medicine, Department of Medicine, Section of General Internal Medicine, Boston, MA, USA.

University of Liverpool, Institute of Translational Medicine, Liverpool, L69 3BX, UK.

Munich Cluster for Systems Neurology (SyNergy), Munich, Germany.

Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland, Australia.

MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK.
Charles E. Schmidt College of Medicine, Florida Atlantic University, Boca Raton, FL, USA.

Center for Experimental Drugs and Diagnostics, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA.

Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, IA, USA.

Division of Psychiatric Genomics, Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York, USA.

Solid GT, Boston, MA, USA.

Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany.

CIBERSAM-Universidad de Granada, Granada, Spain.

Instituto de Investigación Biosanitaria ibs.GRANADA. Hospitales Universitarios de Granada/Universidad de Granada, Granada, Spain.

Department of Psychiatry and Psychotherapy, University of Göttingen, Göttingen, Germany.

Institute of Psychiatric Phenomics and Genomics, Ludwig-Maximilians-University, Munich, 80336 Munich, Germany.

Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA.

Group Health, Seattle, WA, United States.

Department of Psychiatry and Psychiatric Neuroscience, School of Medicine and Dentistry, James Cook University, Townsville, QLD, Australia.

Unit of Genetic Epidemiology & Bioinformatics, Department of Epidemiology, University Medical Center Groningen, Groningen, the Netherlands.

Institute for Community Medicine, University Medicine Greifswald, Germany.

Dalhousie University, 6299 South St, Halifax, Nova Scotia B3H 4R2, Canada.

College of Physicians and Surgeons and the Mailman School of Public Health, Columbia University and New York State Psychiatric Institute, New York, NY, USA.

Center for Psychiatric Genomics, Department of Genetics, Genomic Medicine, University of North Carolina, Chapel Hill, NC, USA.

Department of Psychiatry, University of North Carolina, Chapel Hill, NC, USA.

Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Sweden.

* Contributed equally

Corresponding author: Henning Tiemeier, MD, PhD. Department of Epidemiology, Erasmus University Medical Centre. PO Box 2040, 3000 CA Rotterdam, the Netherlands. Tel +31 10 7043475. Fax: +31 10 7044657. Email: h.tiemeier@erasusmc.nl
Running Title: A Susceptibility Locus for Broad Depression Phenotype

Author Contributions: ND had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. ND (Department of Epidemiology, Erasmus Medical Center) and SW (Department of Genetics, University of North Carolina at Chapell Hill) were conducted and are responsible for the data analysis.

Conflict of Interest: Dr. Sullivan is a consultant for Pfizer, Inc. The other authors report no conflicts.

Keywords: Depressive symptoms, major depressive disorder, CHARGE consortium, Psychiatric Genomics Consortium, genome-wide association study, FHIT gene
Abstract

**Background:** The genetics of depression has been explored in genome-wide association studies that focused on major depressive disorder or depressive symptoms with mostly negative findings. A broad depression phenotype including both phenotypes has not been tested previously using a genome-wide association approach. We aimed to identify genetic polymorphisms significantly associated with a broad phenotype from depressive symptoms to major depressive disorder.

**Methods:** We analysed two prior studies of 70,017 participants of European ancestry from general and clinical populations in the discovery stage. We performed a replication meta-analysis of 28,328 participants. SNP-based heritability and genetic correlations were calculated using LD score regression. Discovery and replication analyses were performed using a $P$-value based meta-analysis. Lifetime major depressive disorder and depressive symptom scores were used as the outcome measures.

**Results:** The SNP-based heritability of major depressive disorder was 0.21 ($\text{SE}=0.02$), the SNP-based heritability of depressive symptoms was 0.04 ($\text{SE}=0.01$), and their genetic correlation was 1.001 ($\text{SE}=0.2$). We found one genome-wide significant locus related to the broad depression phenotype (rs9825823, chromosome 3: 61,082,153, $P=8.2\times10^{-9}$) located in an intron of the $FHIT$ gene. We replicated this SNP in independent samples ($P=0.02$) and the overall meta analysis of the discovery and replication cohorts ($1.0\times10^{-9}$).

**Conclusions:** This large study identified a new locus for depression. Our results support a continuum between depressive symptoms and major depressive disorder. A phenotypically more inclusive approach may help achieve the large sample sizes needed to detect susceptibility loci for depression.
Introduction

The etiology of depression – a worldwide leading cause of disability (1) – is not well understood. As indicated by family, twin and adoption studies, genetic factors mediate part of vulnerability to major depressive disorder (MDD) with a modest heritability of around 40% (2). However, we understand little of the specific genetic architecture of MDD. Multiple genome-wide association studies (GWAS) for MDD have been published (3-10). The largest MDD GWAS was the mega-analysis by the MDD Working Group of the Psychiatric Genomics Consortium (PGC). In that study, over 9,000 MDD cases and 9,500 controls were analyzed but no association with MDD reached genome wide significance (7). Recently, CONVERGE (China, Oxford and VCU Experimental Research on Genetic Epidemiology) consortium identified two genome-wide significant associations in 5,303 Chinese women with severe and recurrent MDD (near the SIRT1 gene, \( P=2.53\times10^{-10} \) and in an intron of the LHPP gene, \( P=6.45\times10^{-12} \)) (11). A GWAS of depressive symptoms (heritability 23%-29%) (12, 13) in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium in approximately 50,000 people from the general population found no genome-wide significant associations (14). Due to the relatively small sample sizes, the previous GWAS of depressive disorders and depressive symptoms were arguably underpowered to detect small genetic effects (15, 16).

Depression can be conceptualized along a continuum of severity from subthreshold or minor depression to MDD of varying severity (e.g., mild, moderate, and severe) (17). Using a continuum approach may augment statistical power as sample size can be increased substantially and patients who fall into the ‘grey area’ can be assessed. Several lines of evidence support a depression continuum. In longitudinal studies, there is an increased risk of MDD in patients with minor depression and subthreshold depression (18, 19). Statistical studies of disorder classification (taxometric) suggested that severity of depression is continuously distributed and there is no discontinuity in the latent structure of depression (19, 20). Family studies report that relatives of probands with milder forms of depression have greater risk of MDD compared to relatives of
probands without any mood disorders (21-24). A higher number of depressive symptoms is related to greater disability, worse quality of life, and a higher mortality risk (18, 25-29). MDD and continuous measures of depression are highly correlated and severity of depressive symptoms along the continuum is linear (30, 31).

The goal of the current study was to combine the results of the largest GWAS using categorical lifetime MDD and continuous measures of depression to identify genetic variants underlying the entire depression continuum.

Methods and Materials

Study Design and Samples

This study was a collaboration between investigators on the PGC MDD and CHARGE genome-wide association meta-analyses (GWAMA). In the discovery phase, we aggregated two GWAMA published in 2013 (7, 14). Basic descriptive features and phenotype definitions of the contributing samples are provided in Table S1. The mega-analysis of MDD consisted of nine studies of 9,240 cases meeting international criteria for lifetime MDD and 9,519 healthy controls. The CHARGE meta-analysis of depressive symptoms included 22 cohorts and comprised 51,258 persons. Each cohort contributing to the GWAMA of the PGC and CHARGE were distinct. In the replication analyses, 16 case-control studies with DSM-IV MDD (6,718 cases and 13,453 controls) were included along with 8,157 subjects from the general population with assessment of depressive symptoms. All subjects were of European ancestry. Institutional review boards approved all studies and all participants provided written informed consent.

Phenotype Characteristics

In the PGC GWAMA, MDD was established with structured clinical interviews (e.g., Clinical Interview Schedule-Revised, Diagnostic Interview for Genetic Studies, and the Structured Clinical Interview for DSM-IV). All clinical evaluations were made by experienced clinicians/interviewers. Most cases were
ascertained from clinical sources. Controls were screened in most of the studies to require the absence of MDD and recruited from the general population. Full details about the PGC samples can be found in the previous publication (7). In the CHARGE GWAMA, depressive symptoms were assessed with validated questionnaires. Measures include the Center for Epidemiological Studies-Depression (CES-D) scale, Geriatric Depression Scale (GDS), Patient Health Questionnaire-9, and the Beck Depression Inventory-II (BDI-II) mostly assessing depressive symptoms in previous weeks rather than lifetime MDD (14). Persons with schizophrenia, bipolar disorder, and dementia were excluded. Persons aged 40 and older and with genotype data and depressive symptom score were included. The 16 MDD case-control replication samples were part of an expanded but unpublished PGC MDD analysis. MDD was diagnosed with interviews. In the depressive symptom replication cohort, the Health and Retirement Study, the 8-item CESD was applied. Respondents were excluded if they were less than 40 years of age or evidence of cognitive impairment.

**Genotyping and Imputation**

In the PGC samples, (Table S1), individual genotypes were assembled, processed through a central quality control pipeline and imputed using the CEU and TSI HapMap3 reference panels. Quality control procedures were extensive (7). In the CHARGE cohorts, genotype quality control and imputation were conducted in each study separately. The imputation reference was the HapMap2 Central EUrOpe (CEU) panel (14). In the MDD replication cohorts (Table S3), imputation was performed using IMPUTE2 / SHAPEIT (chunk size of 3 Mb and default parameters). The imputation reference set consisted of 2,186 phased haplotypes from the 1000 Genomes Project. In the Health and Retirement Study, imputation was performed using the HapMap2 CEU reference panel.

**Statistical Analyses**

LD score regression was used to compute the SNP-based heritability and the genetic correlation using the 1000 Genomes CEU reference panel (32).
In the PGC GWAMA, a logistic regression analysis was used to test the association between MDD and imputed SNP dosages under an additive model and adjusting for study indicators and five principal components (7). In the CHARGE GWAMA, a linear regression analysis was applied to test the association of depressive symptom score on imputed SNP dosages in the contributing studies adjusting for age and sex. Analyses were adjusted for principal components for most but not all cohorts in the CHARGE GWAMA. A P value based meta-analysis was applied in the CHARGE GWAMA (14). Effect size estimates were based on a dichotomous outcome in the PGC and on a continuous outcome in the CHARGE GWAMA. To combine these effect estimates, a P value based meta-analysis weighted by sample size with METAL was used. This method allows different weights for each study and takes into account the direction of effect at each SNP (33). To specify the direction of the effect, the PGC used the logistic regression coefficient beta and the CHARGE used z-scores. Weights were based on the number of the MDD cases in the PGC study (n=9,240) and the number of individuals in the CHARGE with clinically significant depressive symptoms (n=5,976) using population specific cut-off scores of the questionnaires were considered for weighting. To test whether the results are affected by different sample size weightings, equal weights per study, or no weight as suggested by Stouffer (34), we carried out a series of sensitivity analyses.

We selected the genome wide significant SNPs in two loci from the discovery stage for replication. After analyzing these data, we performed a P value based meta-analysis combining all replication samples. Further, we analyzed the results of the discovery and all replication samples weighting for number of cases.

**Results**

In the discovery stage, we performed a GWAMA in 70,017 participants of European ancestry by combining the PGC MDD (7) and CHARGE GWAMA (14). We applied an LD score regression to the summary statistics from each study to compute the SNP-based heritabilities and the genetic correlation. As reported previously (35), the SNP-based liability scale heritability of MDD was 0.2
(standard error 0.02) for 20% of prevalence. The lambda was 1.1 and the regression intercept was 1.0 (standard error 0.01). The SNP-based heritability of depressive symptoms was 0.04 (standard error 0.01). The lambda was 1.1 and the regression intercept was 1.0 (standard error 0.01). The SNP-based heritability of the broad depression phenotype was 0.3 (standard error 0.04). MDD and depressive symptoms showed significant co-heritability (1.001, standard error 0.2, Z-score= 4.6, \( P=4.6x10^{-6} \)). This result supports the contention of a continuum between depressive symptoms and MDD. However, the genetic correlation should be interpreted carefully as LD regression is quite sensitive to environmental confounding and like twin studies often lacks precision. Also, different evaluation methods of the depression phenotypes might cause different genetic correlation estimates that cannot easily be compared.

We conducted a meta-analysis of the PGC MDD and the CHARGE depressive symptoms GWAMA using a weighted, \( P \)-value based meta-analysis. The results are summarized in Figure 1 and Figures S1-S3. The combined meta-analysis was conducted for 918,921 SNPs. Two loci were genome-wide significant: a SNP in an intron of \( FHIT \) (rs9825823, chr3: 61,082,153, \( P =8.2x10^{-9} \)) and a SNP in an intron of \( PLEK2 \) (rs9323497, chr14: 67,873,128, \( P=3.3x10^{-8} \)) (Table 1). All SNPs with a \( P \) value of association <5x10\(^{-5}\) are presented in Table S2. Using different weights or Stouffer’s unweighted method had only slight effects on the results (data not shown). Figures S4 and S5 shows forest plots for two SNPs shown in Table 1.

Table 2 presents the replication analyses and the meta-analysis of discovery and replication results. One of the genome-wide significant variants within the \( FHIT \) gene (rs9825823) was associated with depression continuum in the replication cohorts (\( z \)-score=2.4, \( P=0.02 \)). The result of the final meta-analysis of discovery and replication samples also indicated a positive replication as indexed by a lower \( p \)-value (\( z \)-score=6.1, \( P=1.0x10^{-9} \)). This SNP had a positive association with depressive symptoms in the CHARGE study (\( P=5.5 x 10^{-4} \)) and a similar pattern was observed in the PGC study (\( P=4.1 x 10^{-5} \)). The SNP in an intron of \( PLEK2 \) (rs9323497) was not related to depression continuum significantly (\( z \)-score=0.2, \( P=0.9 \)).
We performed an additional replication analysis of our two genome-wide significant SNPs using the publicly available data of the recently published GWAMA of depressive disorders in a sample of Chinese women (the CONVERGE study) (11). In CONVERGE, rs9825823 (odds ratio=1.01, P=0.12) and rs9323497 (odds ratio=0.97, P=0.0002, with a different direction of association than in our discovery sample) were not related to depression at the genome-wide significance level although the latter reached nominal significance. However, in the joint meta-analysis of the HRS, PGC MDD and the CONVERGE studies, we found that the association between the rs9825823 and the depression continuum (z-score=2.85, P=0.004) was slightly stronger than our initial replication analysis. When these replication and discovery samples were combined, the association with our top hit also became stronger (analyses without the CONVERGE data: z-score= 6.1, $P = 1 \times 10^{-9}$; with the CONVERGE data z-score= 6.2, $P = 6.8 \times 10^{-10}$). Results of additional replication analyses are given in the Table S4.

Discussion

We report the results of a combined GWAMA of depression continuum including MDD (18,759 cases and controls) and depressive symptoms (51,258 participants). In the discovery stage, we found genome-wide significant associations in the FHIT and PLEK2 genes. One SNP in the intron of the FHIT gene showed a significant association in the combined analysis of discovery and replication samples of MDD and depressive symptoms samples, and exceeded a genome-wide significance threshold. The significant locus (rs9825823, chr3: 61,082,153) maps to the intronic region of the fragile histidine triad (FHIT) gene, a tumor suppressor protein implicated in several cancers (36). FHIT is expressed in multiple brain regions (amygdala, anterior cingulate cortex, caudate nucleus, prefrontal cortex, hippocampus, and hypothalamus, http://www.gtexportal.org/home/gene/FHIT, accessed 10.07.2016). It plays an important role in oxidative stress and level of DNA damage (37), biological processes implicated in MDD (38, 39). FHIT is a circadian clock modifier gene (40) and has been related to daytime sleepiness (41), which may be salient to the etiology of depression.
In a GWAS of recurrent, early-onset MDD, three SNPs located in the FHIT gene were among the strongest associations in the overall and sex-stratified analyses (8) although none was genome-wide significant. Genetic variants located in FHIT have been reported in genetic studies of anxiety, (42) autism (43), mental stress (44), comorbid depressive syndromes and alcohol dependence (45), citalopram-induced side effects (46) and in a latent class analysis of MDD symptoms (7), but none met genome-wide significance.

Several methodological aspects should be discussed. First, we evaluated depression continuum by combining cases from clinical populations diagnosed with MDD and participants from the general population who had been assessed for depressive symptoms. Such an inclusive approach may increase heterogeneity of the phenotype especially because lifetime MDD was evaluated whereas depressive symptoms indicate past weeks only. If anything, such approach would cause an underestimation of the effects as less information on depressive symptoms were obtained. However, the advantages of a large sample can outweigh the disadvantages of a less precisely defined phenotype. This has been observed in the GWAS of educational attainment which was successfully used as a proxy for intelligence (47). Our additional replication analysis showed that increasing the sample size yielded a stronger association of the top hit with depression continuum. It is complex to calculate statistical power of the current analysis as quantitative and qualitative measures were combined. In the current study, a genetic association with the depression continuum may reflect an effect on broad depressive phenotypes but could also be accounted for by an association with low levels of general well-being (12-18% heritability) that co-occur with depressive symptoms (48).

Second, we used a P-value based meta-analysis, as effect estimates could not be directly evaluated in a straightforward manner. Third, the heterogeneity of the imputation methods used in the PGC and CHARGE discovery samples might reduce the statistical power. However, different imputation references did not change the results in the published PGC MDD study (7).
In conclusion, in this large GWAMA of a broad depression phenotype, we detected a locus associated with depression in clinical and general population samples. Our results suggest the importance of a broader depression phenotype to identify genetic variants underlying depression. Large samples with different depression phenotypes may also help disentangle the genetic background of different forms of depression.

Acknowledgements

The authors acknowledge the essential role of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium in development and support of this manuscript. CHARGE members include the Netherland’s Rotterdam Study (RS), the NHLBI’s Framingham Heart Study (FHS), Cardiovascular Health Study (CHS), the NHLBI’s Atherosclerosis Risk in Communities (ARIC) Study, and the NIA’s Iceland Age, Gene/Environment Susceptibility (AGES) Study. Core funding for the Psychiatric Genomics Consortium is from the US National Institute of Mental Health (U01MH094421). The PGC was supported by NIMH R01 MH094421 and NIMH R01 MH094421. HRS is supported by the National Institute on Aging (NIA U01AG009740). The genotyping was funded separately by the National Institute on Aging (RC2 AG036495, RC4 AG039029). Our genotyping was conducted by the NIH Center for Inherited Disease Research (CIDR) at Johns Hopkins University. Genotyping quality control and final preparation of the data were performed by the Genetics Coordinating Center at the University of Washington. The CoLaus|PsyCoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 320080–105993, 320080-118308, 33CSCO-122661, 33CS30-139468 and 33CS30-148401). The University of Edinburgh is a charitable body, registered in Scotland, with registration number SC005336. DJM is an NRS Research Fellow, supported by the CSO. AMM is supported by The Wellcome Trust (104036/Z/14/Z), The Dr Mortimer and Theresa Sackler Foundation and by a SFC Senior Clinical Fellowship (SCD/12). GenPod was funded by the Medical Research Council (MRC, UK G0200243) and we acknowledge the other co-investigators, Cowen, Nutt, Peters and Sharp. This study makes use of data generated by the Wellcome Trust Case-Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113, 085475 and 090355. NEWMEDS: The research leading to these results has received support from the Innovative Medicine Initiative Joint Undertaking (IMI-JU) under grant agreement no. 115008 of which resources are composed of European Union and the European Federation of Pharmaceutical Industries and Associations (EFPIA) in-kind contribution and financial contribution from the European Union’s Seventh Framework Programme (FP7/2007-2013). Funding for the NESDA and NTR studies were obtained from the Netherlands Organization for Scientific Research (NWO), Center for Medical Systems Biology (CSMB, NWO Genomics), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007). VU University’s Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA); the European Science Foundation (ESF, EU/QLRT-2001-
01254), the European Community's Seventh Framework Program (FP7/2007-2013), ENGAGE (HEALTH-F4-2007-201413); the European Science Council (ERC Advanced, 230374), Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06), the Avera Institute, Sioux Falls, South Dakota (USA) and the National Institutes of Health (NIH, R0100042157-01A, MH081802, Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health. Computing was supported by BiG Grid, the Dutch e-Science Grid, which is financially supported by NWO. The RADIANT studies present independent research part funded by a joint grant from the Medical Research Council, UK and GlaxoSmithKline (G0701420), and by financial support from the National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The GENDEP project was funded by the European Commission Framework 6 grant, EC Contract Ref.: LSHB-CT-2003-503428. Lundbeck provided nortriptyline and escitalopram for the GENDEP study. GlaxoSmithKline and the UK NIHR of the Department of Health contributed to the funding of the sample collection at the Institute of Psychiatry, London. GENDEP was funded by the European Commission Framework 6 grant, EC Contract Ref.: LSHB-CT-2003-503428. Lundbeck provided nortriptyline and escitalopram for the GENDEP study. GlaxoSmithKline and the UK National Institute for Health Research of the Department of Health contributed to the funding of the sample collection at the Institute of Psychiatry, London. The collection of samples and genotyping of the Danish controls was supported by grants from The Danish Strategic Research Council, The Stanley Research Foundation and H. Lundbeck A/S and we thank David M Hougaard, Section of Neonatal Screening and Hormones, Statens Serum Institute, Copenhagen, Denmark, Preben Bo Mortensen, National Centre for Register-based Research, Aarhus University, Denmark and The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Denmark. SHIP-TREND is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network ‘Greifswald Approach to Individualized Medicine (GANI_MED) funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide genotyping in SHIP-TREND-0 was supported by the Federal Ministry of Education and Research (grant no. 03ZIK012). This work was also funded by the German Research Foundation (DFG: GR 1912/5-1). The Cognitive Function and Mood Study (CoFaMS) is supported by the Faculty of Health Science, University of Adelaide, Adelaide, Australia. We are grateful to the Janssen study volunteers for participating in the clinical studies, participating clinical investigators and the staff for enabling patient recruitment and blood sample collection. We thank the staff in Janssen Neuroscience for sample banking, DNA extraction, quality control, plating, and clinical data anonymization, and the staff at Illumina for genotyping of the Janssen DNA samples.

The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

ND had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. ND (Department of Epidemiology, Erasmus Medical Center) and
SW (Department of Genetics, University of North Carolina at Chapel Hill) were conducted and are responsible for the data analysis.

**Conflict of Interest:** Dr. Sullivan is a consultant for Pfizer, Inc. All other authors report no biomedical financial interests of potential conflicts of interest.

**Figure legends**

**Figure 1: Manhattan Plot**

X-axis represents the chromosomal position for each SNP, and y-axis the -log10 P value for association with depression.
Table 1. Meta-analysis results of the PGC MDD GWAMA and the CHARGE depressive symptoms GWAMA (discovery)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>BP</th>
<th>Closest Gene</th>
<th>Location</th>
<th>Allele</th>
<th>MAF</th>
<th>Direction</th>
<th>z-score</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs9825823</td>
<td>3</td>
<td>61,082,153</td>
<td>FHIT</td>
<td>intron</td>
<td>T/C</td>
<td>0.46</td>
<td>++</td>
<td>5.8</td>
<td>8.2e-09</td>
</tr>
<tr>
<td>rs9323497</td>
<td>14</td>
<td>67,873,128</td>
<td>PLEK2</td>
<td>intron</td>
<td>T/C</td>
<td>0.05</td>
<td>++</td>
<td>5.5</td>
<td>3.3e-08</td>
</tr>
</tbody>
</table>

Chr: chromosome, BP: base pair, MAF: minor allele frequency (GRCh37/hg19). Allele: Minor/Major on the + strand. Direction of effect: CHARGE (N=51,258; continuous outcome analysis), PGC (N=18,759; of which 9,240 were MDD cases). * This analysis was weighted for number of participants with clinically significant depressive symptoms in the CHARGE study (n=5,976) and number of cases in the PGC study (n=9,240).

Table 2. Replication analyses and final meta-analysis of discovery and replication stages

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr:BP</th>
<th>Allele</th>
<th>PGC replication (N=20,171)</th>
<th>HRS replication (N=8,157)</th>
<th>Overall replication (N=28,328)</th>
<th>Meta-analysis of discovery &amp; two replication samples (N=98,345)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs98258</td>
<td>3:61,082,153</td>
<td>T/C</td>
<td>+++-</td>
<td>+</td>
<td>0.0</td>
<td>0.4</td>
</tr>
<tr>
<td>rs93234</td>
<td>14:67,873,128</td>
<td>T/C</td>
<td>+++-</td>
<td>++</td>
<td>0.8</td>
<td>0.0</td>
</tr>
</tbody>
</table>

References


