**Genetic effects influencing risk for major depressive disorder in China and Europe**

Tim B Bigdeli1,\*, Stephan Ripke2-4, Silviu-Alin Bacanu1, Roseann E. Peterson1, Abdel Abdellaoui5, Till F M Andlauer6,7, Aartjan TF Beekman8, Klaus Berger9, Douglas H R Blackwood10, Dorret I Boomsma5, Gerome Breen11,12, Henriette N Buttenschøn13, Enda M Byrne14, Sven Cichon15-18, Toni-Kim Clarke10, Baptiste Couvy-Duchesne14,19,20, Nick Craddock21, EJC de Geus5,22, Franziska Degenhardt18,23, Erin C Dunn24-26, Alexis C. Edwards1, Ayman H Fanous27, Andreas J Forstner18,23, Josef Frank28, Michael Gill29, Scott D Gordon19, Hans J Grabe30, Steven P Hamilton31, Orla Hardiman32, Caroline Hayward33, Andrew C Heath34, Anjali K Henders14, Stefan Herms18,23,35, IB Hickie36, Per Hoffmann18,23,37, Georg Homuth38, Jouke- Jan Hottenga5, Marcus Ising39, Rick Jansen8, Stefan Kloiber39, James A Knowles40, Maren Lang28, Qingqin S Li41, Susanne Lucae39, Donald J MacIntyre10, Pamela AF Madden42, Nicholas G Martin19,43, Patrick J McGrath44, Peter McGuffin12, Andrew M McIntosh10,45, Sarah Elizabeth Medland19, Divya Mehta14, Christel M Middeldorp5, Yuri Milaneschi8, Grant W Montgomery46, Ole Mors47, Bertram Müller-Myhsok6,7,48, Matthias Nauck49, Dale R Nyholt50, Markus M Nöthen18,23, Michael J Owen51, Brenda WJH Penninx8, Michele L Pergadia52, Roy H Perlis53,54, Wouter J Peyrot8, David J Porteous55, James B Potash56, John P Rice57, Marcella Rietschel28, Brien P. Riley1,58, Margarita Rivera59,60, Robert Schoevers61, Thomas G Schulze28,62-65, Jianxin Shi66, Stanley I Shyn67, Johannes H Smit8, Jordan W Smoller24-26, Fabian Streit28, Jana Strohmaier28, Alexander Teumer68, Jens Treutlein28, Sandra Van der Auwera30, Gerard van Grootheest8, Albert M van Hemert69, Henry Völzke68, B. Todd Webb1, Myrna M Weissman70,71, Jürgen Wellmann9, Gonneke Willemsen5, Stephanie H Witt28, Douglas F Levinson72, Cathryn M Lewis12,73, Naomi R Wray14, Jonathan Flint74,75, Patrick F Sullivan76-78 & Kenneth Kendler1,58 on behalf of the CONVERGE consortium and Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium

1Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University School of Medicine; 2Department of Psychiatry, Charite Universitatsmedizin Berlin Campus Benjamin Franklin; 3Medical and Population Genetics, Broad Institute; 4Analytic and Translational Genetics Unit, Massachusetts General Hospital; 5Dept of Biological Psychology, VU University Amsterdam; 6Department of Translational Research in Psychiatry, Max Planck Institute of Psychiatry; 7Munich Cluster for Systems Neurology (SyNergy); 8Department of Psychiatry, VU University Medical Center and GGZ inGeest; 9Institute of Epidemiology and Social Medicine, University of Muenster; 10Division of Psychiatry, University of Edinburgh; 11NIHR BRC for Mental Health, King's College London; 12MRC Social Genetic and Developmental Psychiatry Centre, King's College London; 13Department of Clinical Medicine, Translational Neuropsychiatry Unit, Aarhus University; 14Queensland Brain Institute, The University of Queensland; 15Department of Biomedicine, University of Basel; 16Division of Medical Genetics, University of Basel; 17Institute of Neuroscience and Medicine (INM-1), Research Center Juelich; 18Institute of Human Genetics, University of Bonn; 19Genetics and Computational Biology, QIMR Berghofer Medical Research Institute; 20Centre for Advanced Imaging, The University of Queensland; 21Psychological Medicine, Cardiff University; 22EMGO+ Institute, VU University Medical Center; 23Life&Brain Center, Department of Genomics, University of Bonn; 24Stanley Center for Psychiatric Research, Broad Institute;

25Department of Psychiatry, Massachusetts General Hospital; 26Psychiatric and Neurodevelopmental Genetics Unit (PNGU), Massachusetts General Hospital; 27State University of New York Downstate Medical Center; 28Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University; 29Department of Psychiatry, Trinity College Dublin; 30Department of Psychiatry and Psychotherapy, University Medicine Greifswald; 31Psychiatry, Kaiser-Permanente Northern California; 32Institute of Neuroscience, Trinity College Dublin; 33Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh; 34Psychiatry, Washington University in Saint Louis School of Medicine; 35Division of Medical Genetics and Department of Biomedicine, University of Basel; 36Brain and Mind Research Institute, University of Sydney; 37Human Genomics Research Group, Department of Biomedicine, University of Basel; 38Interfaculty Institute for Genetics and Functional Genomics, Department of Functional Genomics, University Medicine and Ernst Moritz Arndt University Greifswald; 39Max Planck Institute of Psychiatry; 40Psychiatry & The Behavioral Sciences, University of Southern California; 41Neuroscience Therapeutic Area, Janssen Research and Development, LLC; 42Department of Psychiatry, Washington University in Saint Louis School of Medicine; 43School of Psychology, The University of Queensland; 44Psychiatry, New York State Psychiatric Institute & Columbia University College of Physicians and Surgeons; 45Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh; 46Institute for Molecular Biology, University of Queensland; 47Research Department P, Aarhus University Hospital, Risskov; 48University of Liverpool; 49Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald; 50Institute of Health and Biomedical Innovation, Queensland University of Technology; 51MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine; 52Charles E. Schmidt College of Medicine, Florida Atlantic University; 53Psychiatry, Harvard Medical School; 54Psychiatry, Massachusetts General Hospital; 55Medical Genetics Section, CGEM, IGMM, University of Edinburgh; 56Psychiatry, University of Iowa; 57Department of Psychiatry, Washington University in Saint Louis; 58Dept of Human and Molecular Genetics, Virginia Commonwealth University School of Medicine; 59Instituto de Investigación Biosanitaria ibs.Granada and CIBER en Salud Mental (CIBERSAM), University of Granada; 60MRC Social Genetic and Developmental Psychiatry Centre, King´s College London; 61Department of Psychiatry, University of Groningen, University Medical Center Groningen; 62Institute of Psychiatric Phenomics and Genomics (IPPG), Medical Center of the University of Munich, Campus Innenstadt; 63Department of Psychiatry and Psychotherapy, University Medical Center Göttingen; 64Department of Psychiatry and Behavioral Sciences, Johns Hopkins University; 65Human Genetics Branch, NIMH Division of Intramural Research Programs; 66Division of Cancer Epidemiology and Genetics, National Cancer Institute; 67Psychiatry, Group Health; 68Institute for Community Medicine, University Medicine Greifswald; 69Department of Psychiatry, Leiden University Medical Center; 70Division of Epidemiology, New York State Psychiatric Institute; 71Psychiatry, Columbia University College of Physicians and Surgeons; 72Psychiatry & Behavioral Sciences, Stanford University; 73Department of Medical & Molecular Genetics , King's College London; 74Merton College, University of Oxford; 75Wellcome Trust Centre for Human Genetics, University of Oxford; 76Medical Epidemiology and Biostatistics, Karolinska Institutet; 77Genetics, University of North Carolina at Chapel Hill; 78Psychiatry, University of North Carolina at Chapel Hill

\*Correspondence: Tim B. Bigdeli, Ph.D., Virginia Institute for Psychiatric and Behavioral Genetics, 800 E. Leigh St., Biotech One, Suite 101, Richmond, VA 23219-1534, Phone No: (804) 828-5860, Fax No: (804) 828-1471, Email: tim.bigdeli@gmail.com

**Abstract Word Count:** 242

**Article Word count:** 3,224

**Number of Figures:** 1

**Number of Tables:** 2

**Supplemental Information:** 1

**Running Title:** Trans-ancestry genomic analysis of depression

**Keywords:** Major Depressive Disorder; GWAS; trans-ancestry; genetic correlation, polygenic risk

# **Abstract**

Major depressive disorder (MDD) is a common, complex psychiatric disorder and a leading cause of disability worldwide. Despite twin studies indicating its modest heritability (~30-40%), extensive heterogeneity and a complex genetic architecture have complicated efforts to detect associated genetic risk variants. We combined single nucleotide polymorphism (SNP) summary statistics from the CONVERGE and PGC studies of MDD, representing 10 502 Chinese (5 282 cases and 5 220 controls) and 18 663 European (9 447 cases and 9 215 controls) subjects. We determined the fraction of SNPs displaying consistent directions of effect, assessed the significance of polygenic risk scores, and estimated the genetic correlation of MDD across ancestries. Subsequent trans-ancestry meta-analyses combined SNP-level evidence of association. Sign tests and polygenic score profiling weakly support an overlap of SNP effects between East-Asian and European populations. We estimated the trans-ancestry genetic correlation of lifetime MDD as 0.33; females-only and recurrent MDD yielded estimates of 0.40 and 0.41, respectively. Common variants downstream of *GPHN* achieved genome-wide significance by Bayesian trans-ancestry meta-analysis (rs9323497; log10 Bayes Factor=8.08) but failed to replicate in an independent European sample (*P*=0.911). Gene-set enrichment analyses indicate enrichment of genes involved in neuronal development and axonal trafficking. We successfully demonstrate a partially shared polygenic basis of MDD in East-Asian and European populations. Taken together, these findings support a complex etiology for MDD and possible population differences in predisposing genetic factors, with important implications for future genetic studies.

# **Introduction**

Major depressive disorder (MDD) is the most common psychiatric illness and a leading cause of disability worldwide [1,2](https://paperpile.com/c/Qdxtid/nqg3F%2BkkvoK). MDD is modestly heritable (30-40%), may be genetically complex, and likely heterogeneous, complicating efforts to identify replicable risk loci [3,4](https://paperpile.com/c/Qdxtid/Y9kpn%2BmlqsV). The successful detection and interpretation of genetic associations requires both increased sample sizes [5](https://paperpile.com/c/Qdxtid/0lFHn) and empirically-driven efforts to reduce phenotypic heterogeneity [6](https://paperpile.com/c/Qdxtid/5MdWN).

 Underpinning the success of genome-wide association studies (GWAS) of numerous traits has been the emergence of large research consortia [7](https://paperpile.com/c/Qdxtid/Bn8Z0). In addition to facilitating larger sample sizes, many consortia are increasingly ancestrally diverse, enabling identification of novel associations [8–10](https://paperpile.com/c/Qdxtid/cQAh5%2BFzwLM%2BygTv3) and independent replication of reported findings [11,12](https://paperpile.com/c/Qdxtid/YX60P%2B9etuc), as well as improving fine mapping of implicated loci [13,14](https://paperpile.com/c/Qdxtid/uxPhX%2Bq0gDU). Consistent associations at replicated loci have been reported for psychiatric disorders [15](https://paperpile.com/c/Qdxtid/tM2wG) and non-psychiatric traits [8,11,16–18](https://paperpile.com/c/Qdxtid/YX60P%2B7wMKD%2BeJZ6d%2B95KKB%2BcQAh5), and shared liabilities are often borne out by genome-wide polygenic analyses [19–21](https://paperpile.com/c/Qdxtid/4aFEG%2Bqitjh%2BhHxs0).

Whether genetic factors predisposing to MDD are shared across ancestries is not well established, and two replicated genome-wide significant associations for MDD in China had markedly lower allele frequencies in Europeans and thus did not replicate [22–24](https://paperpile.com/c/Qdxtid/L6xO%2BgDp4%2BIbYm). Allelic heterogeneity and population-specific genetic effects have been reported for several complex traits [18,25,26](https://paperpile.com/c/Qdxtid/95KKB%2BbFSjt%2BeSzKV) but the extent of differences across ancestries remains relatively unexplored.

We sought to clarify the extent to which liability to MDD is shared between European and East-Asian populations via collaboration between the Psychiatric Genomics Consortium (PGC) [22](https://paperpile.com/c/Qdxtid/L6xO) and CONVERGE [6](https://paperpile.com/c/Qdxtid/5MdWN) studies of MDD. We asked whether observed directions of allelic effects are consistent across populations, assessed the significance of cross-ancestry polygenic scores, and estimated the trans-ancestry genetic correlation of MDD. We attempted to disentangle population differences from those arising from ascertainment or phenotypic definition through analyses of recurrent MDD and in female subjects. These meta-analyses represent the largest trans-ancestry genetic study of MDD to date.

# **Methods**

## *Ascertainment and genotyping*

Sample ascertainment, SNP genotyping, and quality control procedures for PGC and CONVERGE have been described previously [6,22](https://paperpile.com/c/Qdxtid/5MdWN%2BL6xO). Individual sites and sample sizes are presented in Table 1.

*CONVERGE* (**C**hina, **O**xford a**n**d **V**irginia Commonwealth University **E**xperimental **R**esearch on **G**enetic **E**pidemiology): Briefly, all subjects were Han Chinese women and had two or more episodes of MDD meeting DSM-IV criteria. After applying quality controls modeled after the PGC study, 10 502 samples (5 282 cases and 5 220 controls) and 6 242 619 SNPs were retained for analysis.

*PGC MDD:* Samples included here comprise Stage 1 of the PGC MDD study [22](https://paperpile.com/c/Qdxtid/L6xO). Briefly, all subjects were of European ancestry, all cases were assessed using validated methods and met DSM-IV criteria for lifetime MDD, and the majority of controls were screened to exclude lifetime MDD. Available data on number of depressive episodes were used to identify recurrent cases (two or more episodes). Nine studies from the US, Europe, and Australia were genotyped using SNP arrays. Imputation was performed with IMPUTE2 [27](https://paperpile.com/c/Qdxtid/8VljV) using the 1000 Genomes Project data (v3; GRCh37/hg19) [28](https://paperpile.com/c/Qdxtid/6pZTd), resulting in a total of 13 381 627 autosomal and X chromosome SNPs.

## *Polygenic risk score profiling and binomial sign tests*

Each dataset was filtered on the basis of statistical imputation information (INFO) greater than 0.8 and MAF greater than 0.01 in both CONVERGE and PGC overall; linkage disequilibrium (LD) based “clumping” was used to obtain an approximately independent set of SNPs (*r*2<0.1) while preferentially retaining the most significant SNP within 500kb windows. We computed weighted polygenic scores (i.e., log odds ratio of the associated allele), based on varying *P*-value thresholds in the “training set” results (i.e., CONVERGE or PGC); *P*-value thresholds ranged between 10-5 and 0.5. We evaluated the significance of case-control differences using logistic regression and covarying ancestry-based principal components (PCs) and a study indicator variable. The predictive value of these scores is reported in terms of Nagelkerke’s pseudo-*R*2 (fmsb package in R) [29](https://paperpile.com/c/Qdxtid/Dz6V4).

Using the same sets of SNPs and the same *P*-value thresholds, we applied a binomial sign-test to determine whether the number of SNPs demonstrating consistent directions of allelic effects between CONVERGE and PGC was greater than expected by chance (i.e., a one-sided test of whether this fraction is greater than 0.5).

## *Trans-ancestry genetic correlation*

The recently developed *popcorn* software [30](https://paperpile.com/c/Qdxtid/knFMe) allows for estimation of the trans-ancestry genetic effect correlation (*ρg*)using GWAS summary statistics. Cross-ancestry reference scores, representing SNP-wise estimates of the similarity of LD (with neighboring SNPs) between populations, were calculated for East Asian (*N*=286) and European (*N*=379) subjects from the 1000 Genomes Project (v3) [28](https://paperpile.com/c/Qdxtid/6pZTd). For computational efficiency and consistency with previously reported estimates of genetic correlation, these calculations were based on ~1.2M common SNPs present in HapMap3 [31](https://paperpile.com/c/Qdxtid/Z7tpN) following study-wise exclusion of SNPs with INFO<0.9 or MAF<0.01%.

We attempted to address possible heterogeneity by examining estimates of genetic correlation within- and across-ancestries, and for varying phenotypic definitions. Briefly, we divided the PGC and CONVERGE studies into approximate halves, performing association analysis in each subsample as described above, and subsequently estimating the genetic correlations between these non-overlapping halves. Within the PGC, we randomly selected 5 of 10 studies (*SEUR1*), with the remaining 5 studies taken as a comparison sample (*SEUR2*). We selected *N*=30 of a possible 126 paired comparisons for which the sample sizes of each subset were equivalent (~1:1). We followed an analogous procedure in CONVERGE, selecting 12 of 24 sequencing batches (*SASN1*), with the remaining 12 batches taken as a comparison sample (*SASN2*). Within-ancestry comparisons were between non-overlapping subsets (e.g., *SEUR1* versus *SEUR2*) and utilized reference scores based on a single population, calculated as described above. Cross-ancestry estimates were based on comparisons of the full set of CONVERGE results to each of *N*=60 subsets from the within-PGC analysis. We compared cross-ancestry estimates for lifetime MDD, recurrent MDD and females-only by paired Student’s *t*-tests.

## *SNP-based meta-analyses*

Within each study, we tested for association between SNPs and affection status by logistic regression with PLINK [32](https://paperpile.com/c/Qdxtid/9VFhN), using allelic dosages and including ancestry PCs as covariates (plus a site indicator in PGC analyses). Backwards-stepwise regression was used to select PCs demonstrating association (*P*<0.159) with each diagnosis. We excluded SNPs with MAF<0.01 or INFO<0.5 in either CONVERGE or PGC (overall), or missing in ≥5 of 9 PGC samples. We analyzed the X chromosome as previously described [22](https://paperpile.com/c/Qdxtid/L6xO).

We performed Bayesian meta-analyses of PGC and CONVERGE studies using MANTRA [33](https://paperpile.com/c/Qdxtid/5oHQi). By leveraging population differences in local LD structure, MANTRA has greater power to detect genetic effects demonstrating allelic heterogeneity than traditional approaches assuming random effects. When effects are consistent across studies, MANTRA is effectively a Bayesian implementation of fixed-effects meta-analysis. Inter-study genetic distances were calculated from mean allele frequency differences. We adopted a threshold of log10 Bayes factor (log10BF) > 7 for declaring genome-wide significance.

## *Gene-set enrichment analyses*

We applied DEPICT [34](https://paperpile.com/c/Qdxtid/RuVs4), to identify significantly enriched gene-sets and pathways in specific tissues and cell-types. Briefly, genes in the vicinity of associated SNPs are tested for enrichment for “reconstituted” gene sets, comprised of curated sets expanded to include co-regulated loci. Tissue and cell type enrichment analysis is conducted by testing whether genes were highly expressed in any of 209 MeSH annotations based on microarray data for the Affymetrix U133 Plus 2.0 Array platform [35](https://paperpile.com/c/Qdxtid/2t5CC).

Because DEPICT adjusts for potential sources of confounding and multiple testing using precomputed GWAS of randomly distributed phenotypes , we elected to use as input *P*-values from inverse variance weighted (i.e., fixed effects) meta-analysis of PGC and CONVERGE. Recalling that MANTRA is effectively a Bayesian implementation of fixed-effects meta-analysis when allele frequencies are similar between populations, we considered this to be an appropriate strategy, if not somewhat conservative.

## *Replication analyses*

A total of 4 504 cases and 7 007 controls from 10 independent, European-ancestry cohorts were available for replication (Table 1). These studies represent recent additions to the PGC that were not included in the previously published analysis [22](https://paperpile.com/c/Qdxtid/L6xO). A brief description of each study site is given in the Supplemental Material. At the time of writing, neither comparable East-Asian GWAS datasets nor subject-level data on the number of depressive episodes were readily available. For analyses of recurrent illness, we included those replication studies that specifically ascertained recurrent cases.

For each phenotype definition, we identified independent (pairwise *r*2<0.1 within 500kb windows based on European 1000 Genomes Project samples), significant autosomal SNPs (log10BF>5) from the trans-ancestry meta-analyses (10, 7, and 7 for MDD, females-only, and recurrent MDD, respectively). We tested these SNPs for association using logistic regression and including ancestry PCs as covariates. We performed inverse-variance weighted meta-analyses of the replication samples using METAL. We also performed binomial sign tests comparing the directions of allelic effects across discovery and replication stages.

# **Results**

## *Polygenic risk score profiling and binomial sign tests*

We employed polygenic risk score profiling to determine whether findings from CONVERGE or the PGC are, in aggregate, significantly associated with MDD status in the other study. Scores based on PGC results were nominally associated with MDD in CONVERGE (Figure 1; Supplemental Tables S1-S3), accounting for ~0.1% of risk (Nagelkerke’s pseudo-*R*2=7.46×10-4; *P*=0.02). Scores based on results for females-only yielded similar results (Nagelkerke’s pseudo-*R*2=7.60×10-4; *P*=0.0141), while scores for recurrent MDD were most strongly associated overall (Nagelkerke’s pseudo-*R*2=0.00201; *P*=6.56×10-5). Scores from CONVERGE were nominally associated with MDD status in the PGC data (Nagelkerke’s pseudo-*R*2=6.08×10-4; *P*=6.66×10-3); these scores yielded similar results when considering females-only (Nagelkerke’s pseudo-*R*2=0.00111; *P*=4.15×10-3), and recurrent MDD (Nagelkerke’s pseudo-*R*2=9.13×10-4; *P*=2.02×10-3). However, only the results based on PGC-trained polygenic scores for recurrent MDD remained significant after correction for multiple testing (Supplemental Table S2).

We evaluated whether the observed fraction of results displaying the same direction of allelic effects across studies was significantly greater than expected by chance (i.e., 50%) using binomial sign tests. Supplemental Table S4 give the number of LD-independent SNPs considered, fraction of these SNPs displaying the same direction of effect in the other study, and a one-sided binomial test *P*-value. For lifetime MDD, we observed the largest excess of same-direction effects in PGC for SNPs significant at *P*<0.2 in CONVERGE (50.7%; binomial *P*=1×10-3); this finding remains significant after multiple-testing adjustment (Supplemental Table S4). For the reverse comparison, the largest excess of same-direction effects was observed for SNPS significant at *P*<0.2 in PGC (50.5%; binomial *P*=0.016).

Overall, the greatest excess of same-direction effects in CONVERGE was observed for SNPs significant at *P*<0.1 in the PGC recurrent MDD analysis (51.1%; binomial *P*=3.05×10-5); the fraction of same-direction effects in the PGC was largest in the females-only analysis, for SNPs significant at *P*<0.1 in CONVERGE (50.9%; binomial *P*=1.11×10-3). While statistically significant after correcting for multiple tests (Supplemental Table S4), the observed excess of same-direction effects represents only a very small deviation from expectation under the null hypothesis.

## *Trans-ancestry genetic correlation*

 Table 2 displays the results of the trans-ancestry genetic correlation between East-Asian and European populations. For lifetime MDD (*ρg*=0.332, 95% CI:[0.270,0.394]), this was both significantly greater than zero (*Pρg*>0=7.23×10-26) and significantly less than one (*Pρg*<1=1.40×10-99), indicating a partially shared polygenic basis of MDD risk between East-Asians and Europeans. These findings remain significant after correction for multiple tests.

By comparison, recurrent MDD and females-only yielded slightly higher estimates of genetic correlation (Table 2). We compared these estimates by assuming an approximately normal distribution for *ρg* and obtaining a *Z*-score for the difference in values; these differences were found to be nominally significant for both recurrent MDD (*P1-sided*=0.023) and females-only (*P1-sided*=0.044). We followed-up these results by calculating genetic effect correlation estimates based on comparisons of CONVERGE to *N*=60 random subsets of the PGC data (Supplemental Figure S1). Compared to lifetime MDD (*ρg=*0.309; 95% CI:[0.290 0.327]), estimates of *ρg* were significantly higher for females-only (*ρg=*0.372, 95% CI:[0.344,0.401]; *t*(59)=7.41, *P*1-sided=2.69×10-10) and recurrent MDD (*ρg=*0.375, 95% CI:[0.362,0.389]; *t*(59)=15.29, *P*1-sided=1.74×10-22).

To aid our interpretation of the cross-ancestry results, we derived analogous within-ancestry estimates for East-Asians (*ρg=*0.926, 95% CI:[0.967,0.967]) and Europeans (*ρg=*0.807, 95% CI:[0.856,0.856]) (Supplemental Figure S2). Notably, within-ancestry analysis of East-Asians yielded significantly greater estimates of *ρg* (*t*(56.45)=3.70, *P*2-sided=0.0005). However, as CONVERGE represents a single study, actual population differences are confounded here with those arising from ascertainment or heterogeneity in assessment methods and instruments among participating PGC studies.

## *SNP-based meta-analyses*

We observed the strongest overall evidence of association experiment-wide between SNPs upstream of gephyrin (*GPHN*) at 14q23.3 (rs9323497; log10BF=8.08) and lifetime MDD (Supplemental Figures S3 and S4). Associated SNPs show marked differences in allele frequencies between East-Asian and European populations and opposing directions of allelic effect in CONVERGE and PGC (Supplemental Figure 4). This locus encodes a neuronal assembly protein that anchors glycine and GABAA receptors to the postsynaptic density in inhibitory neurons [36](https://paperpile.com/c/Qdxtid/FSXY). Intriguingly, the gephryin region exhibits an unusual ‘yin-yang’ haplotype structure reflecting strong positive selection related to recent, rapid human evolution [37](https://paperpile.com/c/Qdxtid/SgmH), and has previously yielded suggestive evidence of association with depressive symptoms in the general population [38](https://paperpile.com/c/Qdxtid/L5kQ).

A total of 10 independent associated SNPs (log10BF>5) were prioritized for replication (Supplemental Table S5; Supplemental Figure S4); of these, 3 were in or near *GPHN*, 2 represent previously reported associations in CONVERGE that did not replicate in PGC [6](https://paperpile.com/c/Qdxtid/5MdWN), and 1 was the strongest reported association in the original PGC study [22](https://paperpile.com/c/Qdxtid/L6xO). No single SNP in either the females-only or recurrent MDD analyses attained genome-wide significance (Supplemental Figure S3). From each of these analyses, 7 independent associated SNPs were taken forward to the replication stage (Supplemental Tables S6 and S7).

We attempted to replicate these single-SNP associations in a collection of independent replication samples (4 504 MDD cases and 7 007 controls). For lifetime MDD, no single SNP yielded nominally significant evidence of association (*P*<0.05) in fixed-effects meta-analysis of these replication samples (Supplemental Table S5). Replication analyses also failed to generate replication support for SNP associations identified for females-only or recurrent MDD (Supplemental Tables S6 and S7). Regional association and forest plots for these SNPs are provided in the Supplemental Material (Supplemental Figures S4-S6).

For selected SNPs from the trans-ancestry meta-analyses, we assessed the significance of the observed fraction of SNPs showing the same direction of effect across discovery and replication phases; these fractions were 0.30 (*P*=0.9453), 0.286 (*P*=0.9375), and 0.571 (*P*=0.5) for lifetime, females-only, and recurrent MDD, respectively.

## *Gene-set enrichment analyses*

We used DEPICT to investigate whether particular pathways or gene-sets were enriched for associations with any of the phenotypic definitions considered. For SNPs significant at *P*<10-5 in fixed-effects meta-analyses of lifetime, females-only, and recurrent MDD, (29, 24, and 27 independent loci, respectively), no single pathway or gene-set was significantly enriched, or contained more significant genes than expected by chance, after correction for multiple testing (*q*≥0.20).

When we considered a more inclusive threshold (*P*<10-4), there were 167, 161, and 161 independent loci for lifetime, females-only, and recurrent MDD. Following correction for multiple testing, only central nervous system neuron differentiation (GO:0021953) and axon cargo transport pathways (GO:0008088) were found to be significantly enriched (*q*<0.05) in the analysis of lifetime MDD. An additional eleven (11) gene-sets were suggestively enriched (*q*<0.20) and included several ontology terms related to neurodevelopmental processes (Supplemental Table S8). Finally, no tissue or cell-types were enriched for associations with any definition of MDD (*q*≥0.20), irrespective of the significance threshold applied.

# **Discussion**

We have conducted a large, trans-ancestry meta-analysis representing, to our knowledge, the first systematic effort to analyze European and Han Chinese studies of MDD. As expected, we identified a shared, common polygenic basis of MDD between these population, as exemplified by an excess of same-direction allelic effects, significant polygenic risk score profiling results, and modest estimates of genetic correlation.

We initially considered the simple hypothesis that disease-relevant SNP effects would have similar sizes and directions of effect across European and Han Chinese studies [39](https://paperpile.com/c/Qdxtid/p3tQj), without explicit consideration of population differences arising from genetic drift or divergent genetic architectures. Scores constructed from either PGC or CONVERGE results were significantly associated with lifetime MDD in the other study, albeit explaining a diminutive fraction of risk. However, it is commonly observed that polygenic prediction is generally poorer when ‘training’ and ‘testing’ datasets do not originate from a single ancestral population, likely attributable to differences in allele frequencies and patterns of LD [20,21](https://paperpile.com/c/Qdxtid/qitjh%2BhHxs0).

Next, we applied the recently developed *popcorn* method [30](https://paperpile.com/c/Qdxtid/knFMe) to obtain estimates of the genetic effect correlation between these populations. Briefly*,* the genetic correlation is the correlation coefficient of per-allele SNP effect sizes across populations. We found that the genetic correlation of lifetime MDD was significantly different from both zero and one, suggesting that there is substantial but incomplete overlap in common SNP effects predisposing to MDD in Europe and China. Of particular interest, comparisons based on females-only or recurrent MDD, which better recapitulated the ascertainment strategy in CONVERGE, yielded significantly higher estimates of genetic correlation despite an attendant reduction in sample size.

Given the extensive heterogeneity of MDD, and an expected and demonstrable loss of power arising from between-study differences in ancestry and ascertainment, our limited success in identifying novel, replicable evidence of genome-wide significant association is perhaps unsurprising. It is well understood that a trait’s heritability—and by extension, a shared polygenic liability—is a less important determinant of successful identification of relevant associations than its underlying genetic architecture. Considering the relatively low genetic correlations reported here, we might expect an attenuation of statistical power to detect individual variants, i.e., as compared to a similarly sized studies of the same ancestry. A concomitant, statistically significant enrichment of biologically relevant gene sets is taken as additional support for this interpretation.

*Limitations*

Firstly, the absence of replicable associations with MDD in ancestrally diverse populations precluded more pointed comparisons of specific genetic effects.

Our attempts to reduce the heterogeneity of MDD, namely by focusing on two particular subtypes of illness, should be regarded as preliminary. Furthermore, questions pertaining to both screening and ascertainment of controls were not addressed in the current study, and could have reduced our power to detect relevant variation. We expect that with larger sample sizes, future studies will be sufficiently powered to address these issues.

Finally, by having conducted multiple separate analyses for females-only and recurrent MDD, we increased the multiple-testing burden. As these do not represent completely independent analyses, we have not corrected exhaustively for the total number of tests performed.

*Conclusions*

We have demonstrated a common polygenic basis of MDD that is partially shared between European and Han Chinese populations. Importantly, our findings appear to reinforce the idea that subtyping of MDD may yield additional insight into its etiology [40](https://paperpile.com/c/Qdxtid/tOb6). Striking an advantageous balance between phenotypically more homogeneous definitions of illness and sample size represents an ongoing and nuanced challenge for genetic studies of MDD.

# **Acknowledgements**

The CONVERGE study was funded by the Wellcome Trust (WT090532/Z/09/Z, WT083573/Z/07/Z, WT089269/Z/09/Z) and by NIH grant MH100549. All authors are part of the CONVERGE consortium (China, Oxford and VCU Experimental Research on Genetic Epidemiology) and gratefully acknowledge the support of all partners in hospitals across China.

The PGC is supported by NIH grant U01 MH094421.

Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003 PI: Posthuma) along with a supplement from the Dutch Brain Foundation and the VU University Amsterdam.

We thank T.H. Pers for support for the DEPICT software.

Generation Scotland (GS:SFHS) was supported by a Wellcome Trust Strategic Award “Stratifying Resilience and Depression Longitudinally” (STRaDL) (Reference 104036/Z/14/Z), and received core support from the Chief Scientist Office (CSO) of the Scottish Government Health Directorates [grant number CZD/16/6] and the Scottish Funding Council [HR03006]. AMM is supported by a Scottish Funding Council Senior Clinical Fellowship and by the Dame Theresa and Mortimer Sackler Foundation.

The Bonn/Manheim (BoMa) study was supported by the German Federal Ministry of Education and Research(BMBF) through the Integrated Genome Research Network (IG) MooDS (Systematic Investigation of the Molecular Causes of Major Mood Disorders and Schizophrenia; grant 01GS08144 to Markus M Nöthen and Sven Cichon, grant 01GS08147 to Marcella Rietschel), under the auspices of the National Genome Research Network plus (NGFNplus), and through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders, grant BMBF01ZX1314G), under the auspices of the e:Med Programme.

The GenRED GWAS project was supported by NIMH R01 Grants MH061686 (DF Levinson), MH059542 (WH Coryell), MH075131 (WB Lawson), MH059552 (JB Potash), MH059541 (WA Scheftner) and MH060912 (MM Weissman). We acknowledge the contributions of Dr George S Zubenko and Dr Wendy N Zubenko, Department of Psychiatry, University of Pittsburgh School of Medicine, to the GenRED I project. The NIMH Cell Repository at Rutgers University and the NIMH Center for Collaborative Genetic Studies on Mental Disorders made essential contributions to this project. Genotyping was carried out by the Broad Institute Center for Genotyping and Analysis with support from Grant U54 RR020278 (which partially subsidized the genotyping of the GenRED cases). Collection and quality control analyses of the control data set were supported by grants from NIMH and the National Alliance for Research on Schizophrenia and Depression. We are grateful to Knowledge Networks (Menlo Park, CA, USA) for assistance in collecting the control data set. We express our profound appreciation to the families who participated in this project, and to the many clinicians who facilitated the referral of participants to the study.

The RADIANT studies were funded by: a joint grant from the UK Medical Research Council and GlaxoSmithKline (G0701420); the National Institute for Health Research (NIHR) Specialist Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and King’s College London; and the UK Medical Research Council (G0000647).

The GENDEP study was funded by a European Commission Framework 6 grant, EC Contract Ref.: LSHB-CT-2003-503428.

Max Planck Institute of Psychiatry MARS study was supported by the BMBF Program Molecular Diagnostics: Validation of Biomarkers for Diagnosis and Outcome in Major Depression (01ES0811). Genotyping was supported by the Bavarian Ministry of Commerce, and the Federal Ministry of Education and Research (BMBF) in the framework of the National Genome Research Network (NGFN2 and NGFN-Plus, FKZ 01GS0481 and 01GS08145).

The Netherlands Study of Depression and Anxiety (NESDA) and the Netherlands Twin Register (NTR) contributed to GAIN-MDD and to MDD2000. Funding was from: the Netherlands Organization for Scientific Research (MagW/ZonMW Grants 904-61-090, 985-10- 002, 904-61-193, 480-04-004, 400-05-717, 912-100-20; Spinozapremie 56-464-14192; Geestkracht program Grant 10-000-1002); the Center for Medical Systems Biology (NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure, VU University’s EMGO Institute for Health and Care Research and Neuroscience Campus Amsterdam, NBIC/BioAssist/ RK (2008.024); the European Science Foundation (EU/QLRT-2001-01254); the European Community’s Seventh Framework Program (FP7/2007-2013); ENGAGE (HEALTH-F4-2007-201413); and the European Science Council (ERC, 230374). Genotyping was funded in part by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health, and analysis was supported by grants from GAIN and the NIMH (MH081802).

Funding for the QIMR samples was provided by the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496675, 496739, 552485, 552498, 613602, 613608, 613674, 619667), the Australian Research Council (FT0991360, FT0991022), the FP-5 GenomEUtwin Project (QLG2-CT- 2002-01254) and the US National Institutes of Health (AA07535, AA10248, AA13320, AA13321, AA13326, AA14041, MH66206, DA12854, DA019951), and the Center for Inherited Disease Research (Baltimore, MD, USA). We thank the twins and their families registered at the Australian Twin Registry for their participation in the many studies that have contributed to this research.

Genotyping of STAR\*D was supported by an NIMH Grant MH072802 (SP Hamilton). STAR\*D was funded by the National Institute of Mental Health (contract N01MH90003) to the University of Texas Southwestern Medical Center at Dallas (AJ Rush, principal investigator).

SHIP 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 data of SHIP have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the Caché Campus program of the InterSystems GmbH. SHIP LEGEND was funded by the German Research Foundation (grant GR1912/5-1). The Harvard i2b2 project was supported by Award # U54LM008748 from the National Library of Medicine (to ISK) and R01MH086026 and R01MH085542 from the National Institute of Mental Health (to R.H.P. and J.W.S., respectively).

**Conflict of Interest**

PFS is a scientific advisor to Pfizer, Inc. Hans J Grabe reports receiving funding from: German Research Foundation; Federal Ministry of Education and Research Germany; speakers honoraria from, Eli Lilly, and Servier. Bertram Muller-Myhsok consulted for Affectis Pharmaceuticals. Roy Perlis has received consulting fees from Proteus Biomedical, Concordant Rater Systems, Genomind, and RID Ventures. No other author reports a conflict of interest.

# **References**

1 [Ustun TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJ. Global burden of depressive disorders in the year 2000. *Br J Psychiatry* 2004; **184**: 386–392.](http://paperpile.com/b/Qdxtid/nqg3F)

2 [Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H *et al.* Disability and quality of life impact of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl* 2004; : 38–46.](http://paperpile.com/b/Qdxtid/kkvoK)

3 [Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. *Am J Psychiatry* 2000; **157**: 1552–1562.](http://paperpile.com/b/Qdxtid/Y9kpn)

4 [Flint J, Kendler KS. The genetics of major depression. *Neuron* 2014; **81**: 484–503.](http://paperpile.com/b/Qdxtid/mlqsV)

5 [Wray NR, Pergadia ML, Blackwood DH, Penninx BW, Gordon SD, Nyholt DR *et al.* Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Mol Psychiatry* 2012; **17**: 36–48.](http://paperpile.com/b/Qdxtid/0lFHn)

6 [CONVERGE consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* 2015; **523**: 588–591.](http://paperpile.com/b/Qdxtid/5MdWN)

7 [Sullivan P, Psychiatric Genetics, Investigators. Don’t give up on GWAS. *Mol Psychiatry* 2012; **17**: 2–3.](http://paperpile.com/b/Qdxtid/Bn8Z0)

8 [Wang H, Burnett T, Kono S, Haiman CA, Iwasaki M, Wilkens LR *et al.* Trans-ethnic genome-wide association study of colorectal cancer identifies a new susceptibility locus in VTI1A. *Nat Commun* 2014; **5**: 4613.](http://paperpile.com/b/Qdxtid/cQAh5)

9 [EArly\_Genetics\_Lifecourse\_Epidemiology\_(EAGLE)\_Eczema\_Consortium, Australian Asthma Genetics, Consortium, Australian Asthma Genetics Consortium, Aagc. Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nat Genet* 2015; **47**: 1449–1456.](http://paperpile.com/b/Qdxtid/FzwLM)

10 [Coffee\_and\_Caffeine\_Genetics\_Consortium, Cornelis MC, Byrne EM, Esko T, Nalls MA, Ganna A *et al.* Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. *Mol Psychiatry* 2015; **20**: 647–656.](http://paperpile.com/b/Qdxtid/ygTv3)

11 [Keller MF, Reiner AP, Okada Y, van Rooij FJ, Johnson AD, Chen MH *et al.* Trans-ethnic meta-analysis of white blood cell phenotypes. *Hum Mol Genet* 2014; **23**: 6944–6960.](http://paperpile.com/b/Qdxtid/YX60P)

12 [Ng MC, Shriner D, Chen BH, Li J, Chen WM, Guo X *et al.* Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. *PLoS Genet* 2014; **10**: e1004517.](http://paperpile.com/b/Qdxtid/9etuc)

13 [Replication DIG, Meta-analysis, Consortium, Asian Genetic Epidemiology Network Type 2 Diabetes, Consortium, South Asian Type 2 Diabetes, Consortium, Mexican American Type 2 Diabetes, Consortium, Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in muylti-Ethnic Samples, Consortium *et al.* Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet* 2014; **46**: 234–244.](http://paperpile.com/b/Qdxtid/uxPhX)

14 [Liu CT, Buchkovich ML, Winkler TW, Heid IM, African Ancestry Anthropometry Genetics, Consortium, Consortium, Giant *et al.* Multi-ethnic fine-mapping of 14 central adiposity loci. *Hum Mol Genet* 2014; **23**: 4738–4744.](http://paperpile.com/b/Qdxtid/q0gDU)

15 [Schizophrenia Working Group of the Psychiatric Genomics, Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–427.](http://paperpile.com/b/Qdxtid/tM2wG)

16 [McKenzie CA, Abecasis GR, Keavney B, Forrester T, Ratcliffe PJ, Julier C *et al.* Trans-ethnic fine mapping of a quantitative trait locus for circulating angiotensin I-converting enzyme (ACE). *Hum Mol Genet* 2001; **10**: 1077–1084.](http://paperpile.com/b/Qdxtid/7wMKD)

17 [Nishida N, Sawai H, Kashiwase K, Minami M, Sugiyama M, Seto WK *et al.* New susceptibility and resistance HLA-DP alleles to HBV-related diseases identified by a trans-ethnic association study in Asia. *PLoS One* 2014; **9**: e86449.](http://paperpile.com/b/Qdxtid/eJZ6d)

18 [Wu Y, Waite LL, Jackson AU, Sheu WH, Buyske S, Absher D *et al.* Trans-ethnic fine-mapping of lipid loci identifies population-specific signals and allelic heterogeneity that increases the trait variance explained. *PLoS Genet* 2013; **9**: e1003379.](http://paperpile.com/b/Qdxtid/95KKB)

19 [de Candia TR, Lee SH, Yang J, Browning BL, Gejman PV, Levinson DF *et al.* Additive genetic variation in schizophrenia risk is shared by populations of African and European descent. *Am J Hum Genet* 2013; **93**: 463–470.](http://paperpile.com/b/Qdxtid/4aFEG)

20 [Ikeda M, Aleksic B, Kinoshita Y, Okochi T, Kawashima K, Kushima I *et al.* Genome-wide association study of schizophrenia in a Japanese population. *Biol Psychiatry* 2011; **69**: 472–478.](http://paperpile.com/b/Qdxtid/qitjh)

21 [International Schizophrenia, Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O’Donovan MC *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.](http://paperpile.com/b/Qdxtid/hHxs0)

22 [Major Depressive Disorder Working Group of the Psychiatric, Gwas Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* 2013; **18**: 497–511.](http://paperpile.com/b/Qdxtid/L6xO)

23 [consortium, Converge. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* 2015; **523**: 588–591.](http://paperpile.com/b/Qdxtid/gDp4)

24 [Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR *et al.* Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. *Nat Genet* 2016. doi:](http://paperpile.com/b/Qdxtid/IbYm)[10.1038/ng.3623](http://dx.doi.org/10.1038/ng.3623)[.](http://paperpile.com/b/Qdxtid/IbYm)

25 [Yin X, Low HQ, Wang L, Li Y, Ellinghaus E, Han J *et al.* Genome-wide meta-analysis identifies multiple novel associations and ethnic heterogeneity of psoriasis susceptibility. *Nat Commun* 2015; **6**: 6916.](http://paperpile.com/b/Qdxtid/bFSjt)

26 [Seyerle AA, Young AM, Jeff JM, Melton PE, Jorgensen NW, Lin Y *et al.* Evidence of heterogeneity by race/ethnicity in genetic determinants of QT interval. *Epidemiology* 2014; **25**: 790–798.](http://paperpile.com/b/Qdxtid/eSzKV)

27 [Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009; **5**: e1000529.](http://paperpile.com/b/Qdxtid/8VljV)

28 [Genomes Project, Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**: 56–65.](http://paperpile.com/b/Qdxtid/6pZTd)

29 [Nakazawa M. *Practices of Medical and Health Data Analysis using R*. Pearson Education: Japan, 2007.](http://paperpile.com/b/Qdxtid/Dz6V4)

30 [Brown BC, Asian Genetic Epidemiology Network Type 2 Diabetes, Consortium, Ye CJ, Price AL, Zaitlen N. Transethnic Genetic-Correlation Estimates from Summary Statistics. *Am J Hum Genet* 2016; **99**: 76–88.](http://paperpile.com/b/Qdxtid/knFMe)

31 [International HapMap Consortium. A haplotype map of the human genome. *Nature* 2005; **437**: 1299–1320.](http://paperpile.com/b/Qdxtid/Z7tpN)

32 [Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.](http://paperpile.com/b/Qdxtid/9VFhN)

33 [Morris AP. Transethnic meta-analysis of genomewide association studies. *Genet Epidemiol* 2011; **35**: 809–822.](http://paperpile.com/b/Qdxtid/5oHQi)

34 [Pers TH, Karjalainen JM, Chan Y, Westra HJ, Wood AR, Yang J *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* 2015; **6**: 5890.](http://paperpile.com/b/Qdxtid/RuVs4)

35 [Fehrmann RS, Karjalainen JM, Krajewska M, Westra HJ, Maloney D, Simeonov A *et al.* Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat Genet* 2015; **47**: 115–125.](http://paperpile.com/b/Qdxtid/2t5CC)

36 [Tyagarajan SK, Fritschy JM. Gephyrin: a master regulator of neuronal function? *Nat Rev Neurosci* 2014; **15**: 141–156.](http://paperpile.com/b/Qdxtid/FSXY)

37 [Climer S, Templeton AR, Zhang W. Human gephyrin is encompassed within giant functional noncoding yin-yang sequences. *Nat Commun* 2015; **6**: 6534.](http://paperpile.com/b/Qdxtid/SgmH)

38 [Hek K, Demirkan A, Lahti J, Terracciano A, Teumer A, Cornelis MC *et al.* A genome-wide association study of depressive symptoms. *Biol Psychiatry* 2013; **73**: 667–678.](http://paperpile.com/b/Qdxtid/L5kQ)

39 [Ioannidis J, Ntzani EE, Trikalinos TA. ‘Racial’differences in genetic effects for complex diseases. *Nat Genet* 2004.](http://paperpile.com/b/Qdxtid/p3tQj)<http://www.nature.com/ng/journal/v36/n12/abs/ng1474.html>[.](http://paperpile.com/b/Qdxtid/p3tQj)

40 [Power RA, Tansey KE, Buttenschøn HN, Cohen-Woods S, Bigdeli T, Hall LS *et al.* Genome-wide Association for Major Depression Through Age at Onset Stratification: Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. *Biol Psychiatry* 2016. doi:](http://paperpile.com/b/Qdxtid/tOb6)[10.1016/j.biopsych.2016.05.010](http://dx.doi.org/10.1016/j.biopsych.2016.05.010)[.](http://paperpile.com/b/Qdxtid/tOb6)

**Tables**

## Table 1. Sample sizes by participating study site in discovery and replication phases.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Phase** | **Study** | **Ancestry** | **No. controls** | **No. cases** | **No. recurrent** |
| *Discovery* | Bonn Mannheim | EUR | 1 072 (48.3) | 588 (64.1) | 408 (66.7) |
|  | GenRED1 | EUR | 1 097 (42.7) | 1 020 (70.8) | 1 020 (70.8) |
|  | GSK MPIP | EUR | 859 (67.6) | 861 (67.4) | 831 (67.3) |
|  | MPIP MARS 650 | EUR | 539 (54.5) | 373 (55.2) | 128 (59.4) |
|  | NTR/NESDA | EUR | 1 727 (62.0) | 1 685 (69.2) | 834 (71.2) |
|  | QIMR I317 | EUR | 960 (55.2) | 1 017 (61.0) | 524 (65.6) |
|  | QIMR I610 | EUR | 748 (61.9) | 433 (72.3) | 250 (72.4) |
|  | RADIANT | EUR | 1 549 (54.2) | 1 903 (70.5) | 1 441 (70.8) |
|  | RADIANT - Germany | EUR | 217 (53.9) | 327 (65.7) | 254 (70.5) |
|  | STAR\*D | EUR | 447 (43.8) | 1 240 (58.6) | 912 (59.1) |
|  | CONVERGE† | EAS | 5 220 | 5 282 | 5 282 |
| *Replication*‡ | Edinburgh | EUR | 285 (48.8) | 372 (59.4) | . |
|  | DepGenesNetwork | EUR | 470 (62.6) | 471 (77.5) | 471 (77.5) |
|  | GenRED2 | EUR | 474 (48.9) | 830 (82.8) | 830 (82.8) |
|  | Harvard i2b2 | EUR | 1 067 (50.1) | 806 (66.9) | 806 (66.9) |
|  | Janssen | EUR | 1 380 (60.5) | 466 (68.2) | 466 (68.2) |
|  | QIMR COEX | EUR | 526 (57.6) | 565 (71.5) | . |
|  | RADIANT - Irish cases | EUR | 340 (52.4) | 109 (82.6) | 109 (82.6) |
|  | RADIANT - US cases | EUR | 378 (51.9) | 223 (78.5) | 223 (78.5) |
|  | RADIANT - Denmark cases | EUR | 516 (40.7) | 133 (69.9) | 133 (69.9) |
|  | SHIP-LEGEND | EUR | 1 087 (44.0) | 366 (67.8) | . |
|  | SHIP-TREND-0 | EUR | 484 (44.6) | 163 (71.8) | . |
| *Totals* | Discovery |  | 14 435 (71.3) | 14 729 (78.4) | 11 884 (82.2) |
|  | Replication |  | 7 007 (51.6) | 4 504(72.3) | 2 572 (75.8) |

Numbers of female subjects are displayed parenthetically. Ancestry abbreviations are as follows: European (EUR), East-Asian (EAS). †All cases and controls were female. ‡See note in Methods section*.*

## Table 2. Trans-ancestry genetic correlations between East-Asian and European MDD subtypes.

|  |  |  |  |
| --- | --- | --- | --- |
| **Trait**† | ***ρg\**** | ***Pρg>0*** | ***Pρg<1*** |
| Lifetime MDD | 0.332 [0.270,0.394] | 7.23×10-26 | 1.40×10-99 |
| Females-only MDD | 0.402 [0.326,0.477] | 2.04×10-25 | 2.59×10-54 |
| Recurrent MDD | 0.410 [0.343,0.477] | 5.40×10-33 | 2.23×10-66 |

\*Estimates of *ρg* are displayed with corresponding 95% confidence intervals.†European prevalences of lifetime, females-only, and recurrent MDD were assumed to be 0.15, 0.20, and 0.105, respectively; prevalence of recurrent MDD among Chinese women was assumed to be 0.08.

# **Figure legends**

**Figure 1. Trans-ancestry Association of Polygenic Risk Scores with Major Depressive Disorder.** For scores based on results from PGC or CONVERGE, the variance in risk explained in the other study is shown on the *y*-axis in terms of Nagelkerke’s pseudo-*R*2; scores based on various *P*-value inclusion thresholds are displayed as shaded bars.