Citation for published version (APA):
Bigdeli, T. B., Ripke, S., Peterson, R. E., Trzaskowski, M., Bacanu, S-A., Abdellaoui, A., Andlauer, T. F. M.,
Beekman, A. T. F., Berger, K., Blackwood, D. H. R., Boomsma, D. I., Breen, G., Buttenschøn, H. N., Byrne, E.
M., Cichon, S., Clarke, T-K., Couvy-Duchesne, B., Craddock, N., de Geus, E. J. C., ... Kendler, K. S. (2017).
Genetic effects influencing risk for major depressive disorder in China and Europe. Translational psychiatry,
7(3), e1074. https://doi.org/10.1038/tp.2016.292
GENETIC EFFECTS INFLUENCING RISK FOR MAJOR DEPRESSIVE DISORDER IN CHINA AND EUROPE

TB Bigdeli1, S Ripke3,4, RE Peterson1, M Trzaskowski5,6, S-A Bacanu1, A Abdellaoui7, TFM Andlau8,9, ATF Beekman10, K Berger11, DHR Blackwood12, DI Boomsma7, G Breen13,14, BN Buttenschøn15, EM Byrne6, S Cichon16,17,18,19, T-K Clarke12, B Couvy-Duchesne6,20,21, NJ Craddock22, EJC de Geus7,23, F Degenhardt19,24, EC Dunn25,26,27, AC Edwards1, AH Fanous28, AJ Forstner19,24, J Frank29, M Gill30, DJ MacIntyre12, PAF Madden43, NG Martin20,44, PJ McGrath45, P McGuf

ORIGINAL ARTICLE

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

TB Bigdeli1, S Ripke3,4, RE Peterson1, M Trzaskowski5,6, S-A Bacanu1, A Abdellaoui7, TFM Andlau8,9, ATF Beekman10, K Berger11, DHR Blackwood12, DI Boomsma7, G Breen13,14, BN Buttenschøn15, EM Byrne6, S Cichon16,17,18,19, T-K Clarke12, B Couvy-Duchesne6,20,21, NJ Craddock22, EJC de Geus7,23, F Degenhardt19,24, EC Dunn25,26,27, AC Edwards1, AH Fanous28, AJ Forstner19,24, J Frank29, M Gill30, DJ MacIntyre12, PAF Madden43, NG Martin20,44, PJ McGrath45, P McGuf

Received 19 November 2016; accepted 27 November 2016

Citation: Transl Psychiatry (2017) 7, e1074; doi:10.1038/tp.2016.292

www.nature.com/tp
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 (5282 cases and 5220 controls) and 18 663 European (9447 cases and 9215 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; female-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; log_{10}p = 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.

Translational Psychiatry (2017) 7, e1074; doi:10.1038/tp.2016.292; published online 28 March 2017

INTRODUCTION

Major depressive disorder (MDD) is the most common psychiatric illness and a leading cause of disability worldwide.12 MDD is modestly heritable (30–40%), may be genetically complex and likely heterogeneous, complicating efforts to identify replicable risk loci.3,4 The successful detection and interpretation of genetic associations require both increased sample sizes6 and empirically driven efforts to reduce phenotypic heterogeneity.6

Underpinning the success of genome-wide association studies (GWAS) of numerous traits has been the emergence of large research consortia.7 In addition to facilitating larger sample sizes, many consortia are increasingly ancestrally diverse, enabling identification of novel associations8–10 and independent replication of reported findings,11,12 as well as improving fine mapping of implicated loci.13,14 Consistent associations at replicated loci have been reported for psychiatric disorders15 and non-psychiatric traits,16–18 and shared liabilities are often borne out by genome-wide polygenic analyses.19–21

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 Allelic heterogeneity and population-specific genetic effects have been reported for several complex traits;18,25,26 however, 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 and CONVERGE6 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.

MATERIALS AND METHODS

Ascertainment and genotyping

Sample ascertainment, SNP genotyping and quality-control procedures for PGC and CONVERGE have been described previously.6,22 Individual sites and sample sizes are presented in Table 1.
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 nonoverlapping halves. Within the PGC, we randomly selected 5 of 10 studies (S_EUR), with the remaining five studies taken as a comparison sample (S_EUR2). 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 (S_ASN1), with the remaining 12 batches taken as a comparison sample (S_ASN2). Within-ancestry comparisons were between nonoverlapping subsets (for example, S_EUR versus S_EUR2) 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.

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

| Phase      | Study            | Ancestry | No. of controls | No. of cases | No. of recurrent |
|------------|------------------|----------|-----------------|--------------|-----------------|
| Discovery  | Bonn Mannheim    | EUR      | 1072 (48.3)     | 588 (64.1)   | 408 (66.7)      |
|            | GenRED1          | EUR      | 1097 (42.7)     | 1020 (70.8)  | 1020 (70.8)     |
|            | GSK MIP1         | EUR      | 859 (67.6)      | 861 (67.4)   | 831 (67.3)      |
|            | MP1P MARS 650    | EUR      | 539 (54.5)      | 373 (55.2)   | 128 (59.4)      |
|            | NTR/NESDA        | EUR      | 1727 (62.0)     | 1685 (69.2)  | 834 (71.2)      |
|            | QIMR I317        | EUR      | 960 (55.2)      | 1017 (61.0)  | 524 (65.6)      |
|            | QIMR I617        | EUR      | 748 (61.9)      | 433 (72.3)   | 250 (72.4)      |
|            | RADIANT          | EUR      | 1549 (54.2)     | 1903 (70.5)  | 1441 (70.8)     |
|            | RADIANT—Germany  | EUR      | 217 (53.9)      | 327 (65.7)   | 254 (70.5)      |
|            | STAR*D           | EUR      | 447 (43.8)      | 1240 (58.6)  | 912 (59.1)      |
|            | CONVERGE         | EAS      | 5220            | 5282         | 5282            |
| Replication| Edinburgh        | EUR      | 285 (48.8)      | 372 (59.4)   | 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      | 1067 (50.1)     | 806 (66.9)   | 806 (66.9)      |
|            | Janssen          | EUR      | 1380 (60.5)     | 466 (68.2)   | 466 (68.2)      |
|            | QIMR COEX        | EUR      | 526 (57.0)      | 565 (71.5)   | 565 (71.5)      |
|            | RADIANT—Irish cases | EUR | 340 (52.4) | 109 (66.6) | 109 (66.6) |
|            | 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      | 1087 (44.0)     | 366 (67.8)   | 366 (67.8)      |
|            | SHIP-TRENDO      | EUR      | 484 (44.6)      | 163 (71.8)   | 163 (71.8)      |
| Totals     | Discovery        |         | 14,435 (71.3)   | 14,729 (78.4)| 11,884 (82.2) |
|            | Replication      |         | 7007 (51.6)     | 4504 (72.3)  | 2572 (75.8)     |

Abbreviations: EAS, East Asian; EUR, European. Numbers of female subjects are displayed parenthetically. *All cases and controls were female. See note in Materials and methods section.

SNP-based meta-analyses
Within each study, we tested for association between SNPs and affection status by logistic regression with PLINK,22 using allelic dosages and including ancestry principal components as covariates (plus a site indicator in PGC analyses). Backward-stepwise regression was used to select principal components demonstrating association (P < 0.159) with each diagnosis. We excluded SNPs with minor allele frequency < 0.01 or INFO < 0.5 in either CONVERGE or PGC (overall), or missing in greater than equal to five of nine PGC samples. We analyzed the X chromosome as previously described.22

We performed Bayesian meta-analyses of PGC and CONVERGE studies using MANTRA.33 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. Interstudy genetic distances were calculated from the mean allele frequency differences. We adopted a threshold of log_{10} Bayes factor (log_{10}BF) > 7 for declaring genome-wide significance.

Gene-set enrichment analyses
We applied DEPICT34 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, comprising 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 (Santa Clara, CA, USA).35

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 (that is, 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 4504 cases and 7007 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 A brief description of each study site is given in the Supplementary Material. At the time of writing, neither comparable East Asian GWAS data sets 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 ρ < 0.1 within 500-kb windows based on European 1000 Genomes Project samples), significant autosomal SNPs (log_{10}BF > 5) from the trans-ancestry meta-analyses (10, 7 and 7 MDD, female-only and recurrent MDD, respectively). We tested these SNPs for association using logistic regression and including ancestry principal components 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
Polycenic 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 the MDD status in the other study. Scores based on PGC results were nominally associated with MDD in CONVERGE (Figure 1; Supplementary Tables S1–S3), accounting for ~0.1% of risk (Nagelkerke’s pseudo-$R^2 = 7.46 \times 10^{-4}$; $P = 0.02$). Scores based on results for female-only yielded similar results (Nagelkerke’s pseudo-$R^2 = 7.60 \times 10^{-4}$; $P = 0.0141$), whereas scores for recurrent MDD were most strongly associated overall (Nagelkerke’s pseudo-$R^2 = 0.00201$; $P = 6.56 \times 10^{-5}$). Scores from CONVERGE were nominally associated with MDD status in the PGC data (Nagelkerke’s pseudo-$R^2 = 6.08 \times 10^{-3}$; $P = 6.66 \times 10^{-3}$); these scores yielded similar results when considering female-only (Nagelkerke’s pseudo-$R^2 = 0.00111$; $P = 4.15 \times 10^{-3}$), and recurrent MDD (Nagelkerke’s pseudo-$R^2 = 9.13 \times 10^{-4}$; $P = 2.02 \times 10^{-3}$). However, only the results based on PGC-trained polygenic scores for recurrent MDD remained significant after correction for multiple testing (Supplementary 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 (that is, 50%) using binomial sign tests. Supplementary Table S4 gives 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 $P = 1 \times 10^{-3}$); this finding remains significant after multiple-testing adjustment (Supplementary 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 $P = 3.05 \times 10^{-5}$); the fraction of same-direction effects in the PGC was largest in the female-only analysis, for SNPs significant at $P < 0.1$ in CONVERGE (50.9%; binomial $P = 1.11 \times 10^{-3}$). Although statistically significant after correcting for multiple tests (Supplementary 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 ($\rho_g = 0.332$, 95% confidence interval (CI): (0.270, 0.394)), this was both significantly greater than zero ($P_{\rho_g > 0} = 7.23 \times 10^{-26}$) and significantly less than one ($P_{\rho_g < 1} = 1.40 \times 10^{-99}$), 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

![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. CONVERGE, China, Oxford and Virginia Commonwealth University Experimental Research on Genetic Epidemiology; MDD, major depressive disorder; PGC, Psychiatric Genomics Consortium.](image_url)
these estimates by assuming an approximately normal distribution for $\rho_g$ and obtaining a $Z$-score for the difference in values; these differences were found to be nominally significant for both recurrent MDD ($P_{\text{one-sided}} = 0.023$) and females-only ($P_{\text{one-sided}} = 0.044$). 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 (Supplementary Figure S1). Compared with lifetime MDD ($\rho_g = 0.309$; 95% CI: (0.290, 0.327)), estimates of $\rho_g$ were significantly higher for females-only ($\rho_g = 0.372$, 95% CI: (0.344,0.401); $t(59) = 7.41$, $P_{\text{one-sided}} = 2.69 \times 10^{-10}$) and recurrent MDD ($\rho_g = 0.375$, 95% CI: (0.362, 0.389); $t(59) = 15.29$, $P_{\text{one-sided}} = 1.74 \times 10^{-22}$).

To aid our interpretation of the cross-ancestry results, we derived analogous within-ancestry estimates for East Asians ($\rho_g = 0.926$, 95% CI: (0.967,0.967)) and Europeans ($\rho_g = 0.807$, 95% CI: (0.856,0.856); Supplementary Figure S2). Notably, within-ancestry analysis of East Asians yielded significantly greater estimates of $\rho_g$ ($t(56.45) = 3.70$, $P_{\text{two-sided}} = 0.0005$). However, as CONVERGE represents a single study, actual population differences are conflated 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; $\log_{10}BF = 8.08$) and lifetime MDD (Supplementary 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 (Supplementary Figure 4). This locus encodes a neuronal assembly protein that anchors glycine and GABAA receptors to the postsynaptic density in inhibitory neurons. Intriguingly, the gephyrin region exhibits an unusual ‘yin-yang’ haplotype structure reflecting strong positive selection related to recent, rapid human evolution, and has previously yielded suggestive evidence of association with depressive symptoms in the general population.

A total of 10 independent associated SNPs ($\log_{10}BF > 5$) were prioritized for replication (Supplementary Table S5; Supplementary Figure S4); of these, three were in or near GPHN, two represent previously reported associations in CONVERGE that did not replicate in PGC and one was the strongest reported association in the original GPHC study. No single SNP in either the females-only or recurrent MDD analyses attained genome-wide significance (Supplementary Figure S3). From each of these analyses, seven independent associated SNPs were taken forward to the replication stage (Supplementary Tables S6 and S7).

We attempted to replicate these single-SNP associations in a collection of independent replication samples (4504 MDD cases and 7007 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 (Supplementary Table S5). Replication analyses also failed to generate replication support for SNP associations identified for females-only or recurrent MDD (Supplementary Tables S6 and S7). Regional association and forest plots for these SNPs are provided in the Supplementary 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 $P < 10^{-4}$ in 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 ($q \geq 0.20$).

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, respectively. 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 11 gene sets were suggestively enriched ($q < 0.20$) and included several ontology terms related to neurodevelopmental processes (Supplementary Table S8). Finally, no tissue or cell types were enriched for associations with any definition of MDD ($q \geq 0.20$), 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 populations, 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, 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’ data sets do not originate from a single ancestral population, likely attributable to differences in allele frequencies and patterns of LD.

Next, we applied the recently developed popcorn method 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, that is, as compared with a similarly sized studies of the same ancestry. A concomitant, statistically
significant enrichment of biologically relevant gene sets is taken as an additional support for this interpretation.

Limitations
First, 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. 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.

CONFLICT OF INTEREST
PFS is a scientific advisor to Pfizer. HUG reports receiving funding from the following: German Research Foundation; Federal Ministry of Education and Research Germany; speakers honoraria from Eli Lilly and Servier. BM-M consulted for Affectis Pharmaceuticals. RP has received consulting fees from Proteus Biomedical, Concordant Rater Systems, Genomind and RID Ventures. The remaining author declare no conflict of interest.

ACKNOWLEDGMENTS
The CONVERGE study was funded by the Wellcome Trust (WT090532/Z/09/Z, WT083573/Z/02/2Z, WT089269/Z/02/9Z) 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 P7 Posthumus) along with a supplement from the Dutch Brain Foundation and the VU University Amsterdam. We thank TH Pers for support for the DEPICT software. Generation Scotland (GSFSH) was supported by a Wellcome Trust Strategic Award ‘Stratifying Resilience and Depression Longitudinally’ (STRAbD; Reference 104036/2/14/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 Government Health Directorates (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 (IGR) MoDoS (Systematic Investigation of the Molecular Causes of Major Mood Disorders and Schizophrenia; grant 01GS08144 to Markus M Nöthen and Sven Cichon, grant 01GS08147 to Marcela Retišch), under the auspices of the National Genome Research Network plus (NGPfNplus), and through the Integrated Network IntegrArtef (Integrated Understanding of Causes and Mechanisms in Mental Disorders, grant BMBF01ZX1314G), under the auspices of the eMed Programme. The GenRED GWAS project was supported by NIMH RO1 Grants MH061688 (DF Levison), MH059542 (WH Coyne), MH075131 (WB Lawson), MH059552 (JB Potash), MH059541 (WA Scheffner) and MH060912 (WM Weissman). We acknowledge the contributions of Dr George S Zrubenko and Dr Wendy N Zrubenko, 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 the following: 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 (G000647). The GENRED GWAS was funded by a European Commission Framework 6 grant, EC Contract Ref.: LSHP-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 (10150811). 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 (NEDA) and the Netherlands Twin Register (NTR) contributed to GAIN-MDD and to MDD2000. This study was funded by 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, Spinoza premie 56-464-14192; Geestkracht program Grant 10-000-1002); the Center for Medical Systems Biology (NWO Genomics), Biobanking and Biometric Resources Research Infrastructure, VU University Medical Center and Neuroscience Campus Amsterdam, NBio/BasicAssist/ 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 (214994, 339462, 389892, 389891, 389892, 389893, 442915, 442981, 469679, 469739, 552485, 552498, 613602, 613608, 613674 and 196677), the Australian Research Council (FT0991360 and FT0991022), the FP-5 GenomeWide project (QLGZ-C2- 2002-01254) and the US National Institutes of Health (AA07535, AA12248, AA13320, AA13321, AA13326, AA14041, HH66206, DA12854 and 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 a 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 Collaborative Research Network 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 (GAIN_MEDI’) funded by the Federal Ministry of Education and Research (grant 031L0101B, 031L0103B and 031L0105B). The German Center for Neurodegenerative Diseases (GCNDD) and the German Center for Schizophrenia Research (GCN_S) have been supported by the Federal Ministry of Education and Research (grant no. 033KR012) 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 IBD project was supported by Award # U54LM08748 from the National Library of Medicine (to RHP) and R01MH085542 from the National Institute of Mental Health (to RHP and JWS, respectively).

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.
2 Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H et al. The European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand 2004; 109: 589–597.
3 Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H et al. Diagnosing depressive symptoms in the general population: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl 2004; 38–46.
4 Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry 2000; 157: 1552–1562.
5 Flint J, Kendler KS. The genetics of major depression. Neuroen 2014; 81: 484–503.
6 Wray NR, Perfagida ML, Blackwood DH, Penninx BW, Gordon SD, Nyholt DR et al. Detection of wide association in major depressive disorder: meta-analysis, and lessons learned. Mol Psychiatry 2012; 17: 36–48.
7 CONVERGE consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. Nature 2015; 523: 588–591.
Supplementary Information accompanies the paper on the Translational Psychiatry website (http://www.nature.com/tp)