*Psychological Medicine*

**cambridge.org/psm**

# Pathway-based polygene risk for severe depression implicates drug metabolism in CONVERGE

Anna R. Docherty[1,2], Arden Moscati[2,3], Tim B. Bigdeli[4], Alexis C. Edwards[2], Roseann E. Peterson[2], Daniel E. Adkins[1,5], John S. Anderson[1], Jonathan Flint[6,7], Kenneth S. Kendler[2] and Silviu-Alin Bacanu[2]

[1]Department of Psychiatry, University of Utah School of Medicine, Salt Lake City, UT, USA; [2]Department of Psychiatry, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, School of Medicine, Richmond, VA, USA; [3]Department of Psychiatry, Mount Sinai School of Medicine, New York, NY, USA; [4]Department of Psychiatry, SUNY Downstate, Brooklyn, NY, USA; [5]Department of Sociology, University of Utah, Salt Lake City, UT, USA; [6]Center for Neurobehavioral Genetics, UCLA Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA, USA and [7]Department of Psychiatry and Biobehavioral Sciences, UCLA David Geffen School of Medicine, Los Angeles, CA, USA

**Original Article**

**Cite this article:** Docherty AR *et al* (2020). Pathway-based polygene risk for severe depression implicates drug metabolism in CONVERGE. *Psychological Medicine* **50**, 793–798. https://doi.org/10.1017/S0033291719000618

Received: 17 November 2017
Revised: 5 March 2019
Accepted: 6 March 2019
First published online: 2 April 2019

**Key words:**
CONVERGE; depression; drug metabolism; genetic; polygenic

**Author for correspondence:**
Anna R. Docherty,
E-mail: anna.docherty@utah.edu

## Abstract

**Background.** The Psychiatric Genomics Consortium (PGC) has made major advances in the molecular etiology of MDD, confirming that MDD is highly polygenic. Pathway enrichment results from PGC meta-analyses can also be used to help inform molecular drug targets. Prior to any knowledge of molecular biomarkers for MDD, drugs targeting molecular pathways (MPs) proved successful in treating MDD. It is possible that examining polygenicity within specific MPs implicated in MDD can further refine molecular drug targets.

**Methods.** Using a large case–control GWAS based on low-coverage whole genome sequencing ($N = 10\,640$) in Han Chinese women, we derived polygenic risk scores (PRS) for MDD and for MDD specific to each of over 300 MPs previously shown to be relevant to psychiatric diagnoses. We then identified sets of PRSs, accounting for critical covariates, significantly predictive of case status.

**Results.** Over and above global MDD polygenic risk, polygenic risk within the GO: 0017144 drug metabolism pathway significantly predicted recurrent depression after multiple testing correction. Secondary transcriptomic analysis suggests that among genes in this pathway, CYP2C19 (family of Cytochrome P450) and CBR1 (Carbonyl Reductase 1) might be most relevant to MDD. Within the cases, pathway-based risk was additionally associated with age at onset of MDD.

**Conclusions.** Results indicate that pathway-based risk might inform etiology of recurrent major depression. Future research should examine whether polygenicity of the drug metabolism gene pathway has any association with clinical presentation or treatment response. We discuss limitations to the generalizability of these preliminary findings, and urge replication in future research.

## Introduction

Recurrent major depressive disorder (MDD) is a complex phenotype with limited established associations with biological markers and molecular genetic pathways. The China, Oxford and Virginia Commonwealth University Experimental Research on Genetic Epidemiology (CONVERGE) study and the Psychiatric Genomics Consortium (PGC) have made advances in the molecular genetic etiology of MDD, confirming that MDD is highly polygenic, with aggregation of top loci effects accounting for a small proportion of the variance in case–control status (Peterson *et al*., 2017; Major Depressive Disorder Working Group of the PGC *et al*., 2018). Thus, supporting examination of variants across the entire genome, rather than just top loci, will likely be most informative to etiology.

Statistical genomics can be used to examine both global polygenic risk and risk within established gene pathways. With genomic data, it is increasingly feasible to begin to isolate and examine molecular drug targets for MDD. One way to work toward this aim is by inspecting molecular pathway (MP) enrichment in relation to severe MDD diagnosis, or in relation to symptom presentation in severe cases. It is also important to present caveats: MP analyses are fraught with potential for spurious findings and require (1) large sample sizes, (2) strict correction for (a) multiple comparison and (b) average background polygenic signal, (3) a grain of salt, given the diffuse and quite polygenic signal from even very large GWAS meta-analyses, and thus (4) replication. It is also important to note that available database information about MPs is far from complete.

© Cambridge University Press 2019

CAMBRIDGE UNIVERSITY PRESS

Keeping these important points in mind, we also know that many MPs are enriched for MDD signal, and that medication targeting-specific neurotransmitter pathways are efficacious. Thus, it is possible that MPs have some influence on the presentation of and risk for severe MDD, and potentially on drug response. Given the aim of genomic research to identify promising drug targets for MDD, the identification of potential molecular classifiers of drug response is highly desirable.

To this aim, we examined global polygenic risk for MDD, and also polygenic risk for MDD specific to MPs, in the presentation of severe MDD in a large case–control sample of Han Chinese descent ($N = 10\,060$) (CONVERGE Consortium, 2015). CONVERGE aimed at identifying genetic risk factors for recurrent MDD among a rigorously ascertained cohort, and presents several advantages for genetic study. The sample is entirely female, thus reducing heterogeneity from sex differences. Further, unlike most studies of its size, CONVERGE subjects all have four Han Chinese grandparents, thus reducing noise stemming from population stratification. Finally, to increase severity and potential genetic signal, all subjects were ascertained for recurrent depression, rather than a single episode. This ascertainment method, and the size of the study, allowed CONVERGE to identify and replicate risk loci for MDD (CONVERGE Consortium, 2015).

This analysis used summary statistics from a leave-one-batch-out GWAS of MDD within CONVERGE. We constructed Bayesian subject-level genome-wide MDD polygenic risk scores (PRS) and MDD PRS specific to pruned MPs. Scores and covariates were regressed onto case status, and pathway scores with significant signal above and beyond global MDD polygenic risk were then examined by within-pathway, gene-based PRS.

## Methods and materials

### Sampling

Recurrent MDD cases and healthy controls were recruited from 51 mental health centers and psychiatric departments of general medical hospitals across 21 provinces of China. Details have been previously published (CONVERGE Consortium, 2015). We controlled for potential clinical heterogeneity by recruiting only female participants, and to control for ethnic stratification, only participants whose grandparents (all four) were of Han Chinese descent were recruited to participate. Cases and controls [age $M$ (S.D.) = 44.4 (8.9) and 47.7 (5.6), respectively] were excluded for diagnosis of bipolar disorder, any psychotic symptoms occurring outside of the context of a major depressive episode, and any diagnosis of intellectual disability. Description of the clinical features of the cases can be found in Edwards *et al.* (2018). Cases had at least two major depressive episodes with the first episode occurring prior to age 50, and could not have abused drugs or alcohol prior to their first episode of depression. Controls were clinically screened to rule out prior depressive episodes and had to be at least 40 years of age, past the age window of typical MDD onset. The study protocol was approved centrally by the Ethical Review Board of Oxford University, and by the ethics committees of all of the participating hospitals in China. All participants provided written informed consent.

### DNA sequencing

DNA extraction, sequencing, genotyping, and imputation details have been reported (Cai *et al.*, 2017). Briefly, DNA was extracted from saliva using Oragene and sequenced reads were obtained from Illumina Hiseq machines aligned to Genome Reference Consortium Human Build 37 patch release 5 (GRCh37.p5) with Stampy (v1.0.17) with default parameters (Lunter and Goodson, 2011). Reads consisting of base quality ⩽5 or containing adaptor sequencing were filtered out. Alignments were indexed in BAM format (Lunter and Goodson, 2011) using Samtools (v0.1.18) (Li *et al.*, 2009), and PCR duplicates were marked for downstream filtering using Picardtools (v1.62). The Genome Analysis Toolkit's (GATK, version 2.6) BaseRecalibrator (McKenna *et al.*, 2010) created recalibration tables to screen known SNPs and INDELs in the BAM files from dbSNP (version 137, excluding all sites added after version 129). GATKlite (v2.2.15) was used for subsequent base quality recalibration and removal of read pairs with improperly aligned segments as determined by Stampy.

### Calling and imputation of genotypes

GATK's UnifiedGenotyper VariantRecalibrator (version 2.7-2-g6bda569) was used on post-BSQR files for variant discovery and genotyping at all polymorphic SNPs in 1000 G Phase1 ASN panel (The 1000 Genomes Project Consortium *et al.*, 2012), as well as variant quality score recalibration. A sensitivity threshold of 90% was applied for imputation after optimizing for transition to transversion ratios. Genotype likelihoods were calculated using a binomial mixture model implemented in SNPtools (version 1.0) (Wang *et al.*, 2013), and imputation was performed at sites with no reference panel using BEAGLE (version 3.3.2) (Browning and Browning, 2007). A second round of imputation was performed at biallelic polymorphic SNPs using 1000G Phase 1 ASN haplotypes as a reference panel. To determine the final number of SNPs, we applied a conservative inclusion threshold for SNPs: (1) a $p$ value for violation $HWE$ >10–6, (2) information score >0.9, and (3) minor allele frequency in CONVERGE >0.5%.

### Diagnostic assessments

Participant interview sessions lasted approximately 2 h. Interviewers were largely trained psychiatrists with a small number representing post-graduate medical students or psychiatric nurses, and all were clinically trained by the CONVERGE team for at least one week. Interviews were recorded and included an assessment of psychopathology, demographic characteristics, and psychosocial functioning. Trained editors listened to a portion of the interviews to provide ratings of interview quality. An assessment of lifetime adversity (binary yes/no) reflected a composite of endorsements on self-reported stressful life event and childhood sexual abuse measures (Peterson *et al.*, 2018). In CONVERGE, adversity has been observed to contribute significant variance in genetic signal, and was included as a feature in our follow-up analyses only to determine whether genetic signal is influenced, in any direction, by presence of lifetime adversity. We excluded participants who had incomplete assessment information or were lacking high-quality genetic data, to arrive at a final 4728 controls and 5612 case samples for analysis.

### Polygenic risk for MDD and pathway-based risk as a predictor of case status

MDD PRSs were constructed globally (genome-wide) and within each of the MPs, using summary statistics from leave-one-batch-

https://doi.org/10.1017/S0033291719000618 Published online by Cambridge University Press

out GWAS analyses of the CONVERGE sample. A list of pathways was derived from independent, previously published pathway enrichment statistics from the PGC (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). These pathways were deemed to likely be the most relevant to psychiatric disorders broadly. Gene lists corresponding to each pathway were extracted from the relevant gene ontology database, with the number of genes per pathway ranging from 8 to 598 (mean = 9). Python-based LDpred (Vilhjalmsson *et al*., 2015), a polygenic risk scoring software, allows for the modeling of LD to weight the relative contributions of syntenic variants to the outcome phenotype. LDpred uses postulated proportions of causal variants in the genome as Bayesian prior probabilities for PRS calculations, and a range of eight different priors were used (proportions of 1, 0.3, 0.1, 0.03, 0.01, 0.003, and 0.001) to construct scores. All pathway scores were LD pruned and merged based on excessive correlation between scores ($R^2 < 0.5$). Stepwise regressions were run with all PRS variables and critical covariates as determined by the CONVERGE Consortium (2015). MDD was treated as a dichotomous variable (despite likely heterogeneity) in order to avoid predicting additional clinical information within cases such as symptom count.

We tested two models for the relationship between PRSs and MDD status. Model 1 included all risk scores and the two primary ancestry principal components. Model 2 included all risk scores, primary ancestry principal components, and the presence/absence of lifetime adversity (a significant moderator of the genetic signal in previous CONVERGE GWAS). Follow-up analyses included similar associations of any significant pathway-based PRS associations with age at onset (AAO), family history of MDD, and number of lifetime episodes within the MDD cases.

## Results

For the primary analysis of this paper, i.e. finding pathways associated with MDD status, we employed a stepwise regression that pared down the number of pathways PRSs to only the significant ones (after adjusting for multiple testing). Statistics for the best-fitting model are presented in Table 1, each parameter in the model remaining significant after FDR correction. The best-fitting stepwise regression included three robust predictors: Global MDD, one ancestry PC, and the MDD PRS for the GO:0017144 drug metabolism pathway. Model 2 results are presented in Table 2. Global MDD risk and GO:0017144 pathway-based risk remained significant in Model 2, where lifetime adversity was entered as a covariate (and was found to be the most robust predictor). Of note, the stepwise models captured no increase in pathway-based risk by genome-wide MDD PRS. Standardized effect sizes, presented as squared Pearson partial correlation coefficients in both Tables 1 and 2, were obtained after adjusting for multiple testing (Bigdeli *et al*., 2016). We mention that these analyses appear to be well powered, given that (even after adjusting for multiple testing) we estimated to have a power >80% to detect PRS explaining only 0.003 of variation in MDD liability, which is half of the in-sample estimates for GO:0017144.

While not the main goal of this work, due to their likely reduced power, we use a gene PRS-type secondary/exploratory analysis to suggest genes in the drug metabolism pathway that might be most important to MDD. Of the 33 genes included in the pathway, most of which reflect protein coding for cytochrome P450 and flavin containing monooxygenase (https://icgcportal.genomics.cn/genesets/GO:0017144), 88 803 variants on nine genes passed variant filtering and LD pruning. Exploratory gene-based analysis included MDD PRS calculation specific to the nine genes involved in the GO:0017144 pathway. Exploratory LDpred analyses of the PRSs associated with the 33 genes yielded three genes with suggestive signals (Table 3). *CYP2C19* and *CBR1* both yielded FDR *q*-values of 0.05. Follow-up analyses of the GO:0017144 pathway in the MDD cases resulted in a significant association with AAO in the expected direction, after correction for multiple testing ($t = 3.388$, $p = 0.0007$ without global MDD PRS in the model; $t = 3.382$, $p = 0.0008$ with global MDD PRS in the model). Gene-level analyses of CBR1 and CYP2C19 predicting AAO were suggestive but not significant ($p < 0.08$).

## Discussion

In the CONVERGE sample, MDD polygenic risk specific to the GO:0017144 pathway significantly predicted MDD case status above and beyond genome-wide polygenic risk for MDD. Exploratory analyses suggest that among genes in this pathway, *CYP2C19* and *CBR1* appear to contribute the most to MDD liability, as their effect on trait remained significant after accounting for both MDD polygenic risk and for an established genetic risk moderator, lifetime adversity. These scores were based on MDD GWAS weights entirely within-ancestry, and did not rely on summary statistics from Northern European ancestry GWAS. This is the first study to test MDD PRS specific to MPs in relation to recurrent MDD, and preliminary results suggest that GO:0017144 may be relevant to recurrent MDD and to clinical presentation.

MDD is genetically complex, and it is important to try to identify areas of focus on the genome with which to inform drug target research. These results are potentially relevant to pharmacogenomics. Several common antidepressants (tricyclic antidepressants and SSRIs, including citalopram, escitalopram, sertraline, amitriptyline, and clomipramine) are substrates for *CYP2C19* (Sim *et al*., 2010). More importantly, *CYP2C19* possesses epoxygenase activity and metabolizes various prostaglandins (arachidonic acid, linoleic acid, eicosapentaenoic acid) to active epoxide products, which have vasodilatory, anti-inflammatory and antinociceptive effects (Fabbri *et al*., 2018). Similarly, *CBR1* plays a protective role in oxidative stress, neurodegeneration, and apoptosis. CBR1 normally breaks down prostaglandin E2, which may mediate the aversive/negative affect component of inflammatory pain (Singh *et al*., 2017).

Because MDD is likely quite heterogeneous, it is important to try to identify genetic classes of cases that differ phenotypically (Milaneschi *et al*., 2017). This view is supported by research attempting to genetically stratify severe psychopathology (Howard *et al*., 2017). Recent research by Howard *et al.* has successfully replicated a genetic classification of MDD using PRS of ~20 medical and psychiatric phenotypes. Future research would benefit from examination of whether high-risk PRS classes are phenotypically distinct (whether clinical features differ) and also whether they are distinct with respect to risk within drug-relevant MPs.

Were there enough non-European GWAS to complete a similar stratified analysis to CONVERGE, we would have attempted this in conjunction with our analysis of MP-based genetic MDD risk. PRS for several medical and psychiatric phenotypes can currently be calculated for European samples based on European GWAS, but unfortunately the field lacks sufficient

https://doi.org/10.1017/S0033291719000618 Published online by Cambridge University Press

**Table 1.** Best-fitting stepwise regression model predicting major depressive disorder case status from global major depressive disorder polygenic risk score, molecular pathway-based polygenic risk score, and ancestry

| | Estimate | Standard error | *t*-value | *p* value | Effect size | $R^2$ |
|---|---|---|---|---|---|---|
| Ancestry principal component 2 | 3.176 | 0.565 | 5.624 | $1.86*10^{-8}$ | 0.04 | 0.007 |
| MDD global L1O PRS | 0.038 | 0.003 | 13.177 | $1.19*10^{-39}$ | 0.13 | 0.041 |
| GO:0017144 Drug metabolism L1O PRS | 0.056 | 0.012 | 4.828 | $1.38*10^{-6}$ | 0.03 | 0.006 |

MDD, major depressive disorder; L1O, leave-one-out cross-validation scoring procedure; PRS, polygenic risk score; GO, gene ontology database.

**Table 2.** Best-fitting stepwise regression model accounting for adversity

| | Estimate | Standard error | *t*-value | *p* value | Effect size | $R^2$ |
|---|---|---|---|---|---|---|
| Lifetime adversity | 0.176 | 0.011 | 15.843 | $1.57*10^{-56}$ | 0.15 | 0.059 |
| MDD global L1O PRS | 0.031 | 0.003 | 11.855 | $2.04*10^{-32}$ | 0.11 | 0.033 |
| Ancestry principal component 2 | 3.989 | 0.503 | 7.933 | $2.13*10^{-15}$ | 0.07 | 0.015 |
| GO:0017144 Drug metabolism L1O PRS | 0.050 | 0.010 | 4.845 | $1.27*10^{-6}$ | 0.03 | 0.006 |

MDD, major depressive disorder; L1O, leave-one-out cross-validation scoring procedure; PRS, polygenic risk score; GO, gene ontology database.
Lifetime adversity is binary (1 = lifetime adversity and 0 = no lifetime adversity).

**Table 3.** Suggestive genes in drug metabolism pathway after false discovery rate correction

| Suggestive genes in GO:17144 | Chr:Mbp start-end | FDR | Targets/relevance |
|---|---|---|---|
| CYP2C19: Cytochrome P450 Family 2 Subfamily C Member 19 | 10:94.763–10:94.853 | 0.05 | • Tricyclic antidepressants and SSRIs<br>• Vasodilatory, anti-inflammatory and antinociceptive effects |
| CBR1: Carbonyl reductase 1 | 21:36.070–21:36.073 | 0.05 | • Protective role in oxidative stress, neurodegeneration, apoptosis |
| FMO2: Flavin containing monooxygenase 2 | 1:171.185–1:171.213 | 0.09 | |

GO, gene ontology database; Chr:Mbp, chromosome and position; FDR, false discovery rate corrected *p* value.

East Asian GWAS for analysis in CONVERGE. We hope that future GWAS of non-European samples will facilitate additional meaningful insights.

### Strengths

A strength of this study is that ascertainment successfully minimized genetic heterogeneity due to sex differences and to ancestry. Moreover, ascertainment was limited to recurrent depression rather than a single episode, potentially increasing the strength of the genetic signal. In addition, this large cohort provided sufficient power to detect even small effects, and deep phenotyping allowed for the additional inspection of lifetime adversity in relation to genetic signal. We included a limited number of pathways – only those implicated in severe psychopathology – in order to limit Type I error.

### Limitations

Strict ascertainment of females with four Han Chinese grandparents may limit the generalizability of GO:0017144 pathway findings. Because the study rigorously ascertained recurrent MDD, and not single episode, it is also possible that the effects observed here are limited to severe and/or recurrent MDD. One source of potential bias stems from the inclusion of the pathways implicated in severe psychopathology in Europeans rather than in individuals of Han Chinese ancestry. In addition, we did not detect signal for some candidate pathways (e.g. serotonergic or noradrenergic) that we might have expected to be relevant to MDD. Null results could stem from insufficient coverage across enough genes in the pathway after quality control, or pathway signal being obscured by the global polygenic signal for MDD that we accounted for in our models. This would be consistent with the small genetic effect sizes we find in such a large MDD cohort.

Finally, there are problems inherent to any study of gene pathways. These include the field's very limited molecular understanding of gene pathways and related sequelae, incompleteness of pathway databases, conflicting opinions about the independence/interdependence of gene pathways, and the elevated polygenicity of all psychiatric traits. To the first three points, we assert that it is studies like this that further develop the pathway literature and inform pathway database efforts. To the fourth point, it is critical for studies to account for genome-wide polygenic risk when testing for any signal relevant to a small portion of the

https://doi.org/10.1017/S0033291719000618 Published online by Cambridge University Press

genome. These points underscore the importance of independent replication.

**Author ORCIDs.** Anna R. Docherty, 0000-0001-7139-7007.

**Acknowledgements.** The authors gratefully acknowledge the support of all partners in hospitals across China, without whom this research would not be possible.

**Financial support.** This work was supported by the Wellcome Trust (J.F., grant numbers WT090532/Z/09/Z, WT083573/Z/07/Z, WT089269/Z/09/Z); the National Institutes of Health (A.D., grant number K01MH093731; A.E., grant number AA021399; D.A., grant number KMH109765; K.K., grant number MH100549; R.P., grant number K01MH113848; S.B., grant numbers MH100560, AA022717); the Simons Foundation (A.D., grant number 513631); the Brain & Behavior Research Foundation (A.D.); and the American Foundation for Suicide Prevention (A.D.).

**Conflict of interest.** None.

## References

**Bigdeli TB, Lee D, Webb BT, Riley BP, Vladimirov VI, Fanous AH, Kendler KS and Bacanu SA** (2016) A simple yet accurate correction for winner's curse can predict signals discovered in much larger genome scans. *Bioinformatics (Oxford, England)* **32**, 2598–2603.

**Browning SR and Browning BL** (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics* **81**, 1084–1097.

**Cai N, Bigdeli TB, Kretzschmar WW, Li Y, Liang J, Hu J, Peterson RE, Bacanu S, Webb BT, Riley B, Li Q, Marchini J, Mott R, Kendler KS and Flint J** (2017) 11,670 whole-genome sequences representative of the Han Chinese population from the CONVERGE project. *Scientific Data* **4**, 170011.

**CONVERGE Consortium** (2015) Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588–591.

**Edwards AC, Docherty AR, Moscati A, Bigdeli TB, Peterson RE, Webb BT, Bacanu SA, Hettema JM, Flint J and Kendler KS** (2018) Polygenic risk for severe psychopathology among Europeans is associated with major depressive disorder in Han Chinese women. *Psychological Medicine* **48**, 777–789.

**Fabbri C, Tansey KE, Perlis RH, Hauser J, Henigsberg N, Maier W, Mors O, Placentino A, Rietschel M, Souery D, Breen G, Curtis C, Lee SH, Newhouse S, Patel H, O'Donovan M, Lewis G, Jenkins G, Weinshilboum RM, Farmer A, Aitchison KJ, Craig I, McGuffin P, Schruers K, Biernacka JM, Uher R and Lewis CM** (2018) Effect of cytochrome CYP2C19 metabolizing activity on antidepressant response and side effects: meta-analysis of data from genome-wide association studies. *European Neuropsychopharmacology* **28**, 945–954.

**Howard DM, Clarke TK, Adams MJ, Hafferty JD, Wigmore EM, Zeng Y, Hall LS, Gibson J, Boutin TS, Hayward C, Thomson PA, Porteous DJ, Smith BH, Murray AD, MDD Working Group of the PGC**, Haley CS, Deary IJ, Whalley HC and McIntosh AM (2017) The stratification of major depressive disorder into genetic subgroups. bioRxiv 134601.

**Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G and Durbin R** (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* **25**, 2078–2079.

**Lunter G and Goodson M** (2011) Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Research* **21**, 936–939.

**Major Depressive Disorder Working Group of the PGC, Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, Adams MJ, Agerbo E, Air TM, Andlauer TMF, Bacanu SA, Baekvad-Hansen M, Beekman AFT, Bigdeli TB, Binder EB, Blackwood DRH, Bryois J, Buttenschon HN, Bybjerg-Grauholm J, Cai N, Castelao E, Christensen JH, Clarke TK, Coleman JIR, Colodro-Conde L, Couvy-Duchesne B, Craddock N, Crawford GE, Crowley CA, Dashti HS, Davies G, Deary IJ, Degenhardt F, Derks EM, Direk N, Dolan CV, Dunn EC, Eley TC, Eriksson N, Escott-Price V, Kiadeh FHF, Finucane HK, Forstner AJ, Frank J, Gaspar HA, Gill M, Giusti-Rodriguez P, Goes FS, Gordon SD, Grove J, Hall LS, Hannon E, Hansen CS, Hansen TF, Herms S, Hickie IB, Hoffmann P, Homuth G, Horn C, Hottenga JJ, Hougaard DM, Hu M, Hyde CL, Ising M, Jansen R, Jin F, Jorgenson E, Knowles JA, Kohane IS, Kraft J, Kretzschmar WW, Krogh J, Kutalik Z, Lane JM, Li Y, Li Y, Lind PA, Liu X, Lu L, MacIntyre DJ, MacKinnon DF, Maier RM, Maier W, Marchini J, Mbarek H, McGrath P, McGuffin P, Medland SE, Mehta D, Middeldorp CM, Mihailov E, Milaneschi Y, Milani L, Mill J, Mondimore FM, Montgomery GW, Mostafavi S, Mullins N, Nauck M, Ng B, Nivard MG, Nyholt DR, O'Reilly PF, Oskarsson H, Owen MJ, Painter JN, Pedersen CB, Pedersen MG, Peterson RE, Pettersson E, Peyrot WJ, Pistis G, Posthuma D, Purcell SM, Quiroz JA, Qvist P, Rice JP, Riley BP, Rivera M, Saeed Mirza S, Saxena R, Schoevers R, Schulte EC, Shen L, Shi J, Shyn SI, Sigurdsson E, Sinnamon GBC, Smit JH, Smith DJ, Stefansson H, Steinberg S, Stockmeier CA, Streit F, Strohmaier J, Tansey KE, Teismann H, Teumer A, Thompson W, Thomson PA, Thorgeirsson TE, Tian C, Traylor M, Treutlein J, Trubetskoy V, Uitterlinden AG, Umbricht D, Van der Auwera S, van Hemert AM, Viktorin A, Visscher PM, Wang Y, Webb BT, Weinsheimer SM, Wellmann J, Willemsen G, Witt SH, Wu Y, Xi HS, Yang J, Zhang F, Arolt V, Baune BT, Berger K, Boomsma DI, Cichon S, Dannlowski U, de Geus ECJ, DePaulo JR, Domenici E, Domschke K, Esko T, Grabe HJ, Hamilton SP, Hayward C, Heath AC, Hinds DA, Kendler KS, Kloiber S, Lewis G, Li QS, Lucae S, Madden PFA, Magnusson PK, Martin NG, McIntosh AM, Metspalu A, Mors O, Mortensen PB, Muller-Myhsok B, Nordentoft M, Nothen MM, O'Donovan MC, Paciga SA, Pedersen NL, Penninx B, Perlis RH, Porteous DJ, Potash JB, Preisig M, Rietschel M, Schaefer C, Schulze TG, Smoller JW, Stefansson K, Tiemeier H, Uher R, Volzke H, Weissman MM, Werge T, Winslow AR, Lewis CM, Levinson DF, Breen G, Borglum AD and Sullivan PF** (2018) Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics* **50**, 668–681.

**McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M and DePristo MA** (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**, 1297–1303.

**Milaneschi Y, Lamers F, Peyrot WJ, Baune BT, Breen G, Dehghan A, Forstner AJ, Grabe HJ, Homuth G, Kan C, Lewis C, Mullins N, Nauck M, Pistis G, Preisig M, Rivera M, Rietschel M, Streit F, Strohmaier J, Teumer A, Van der Auwera S, Wray NR, Boomsma DI and Penninx B** (2017) Genetic association of major depression with atypical features and obesity-related immunometabolic dysregulations. *JAMA Psychiatry* **74**, 1214–1225.

**Peterson RE, Cai N, Bigdeli TB, Li Y, Reimers M, Nikulova A, Webb BT, Bacanu SA, Riley BP, Flint J and Kendler KS** (2017) The genetic architecture of major depressive disorder in Han Chinese women. *JAMA Psychiatry* **74**, 162–168.

**Peterson RE, Cai N, Dahl AW, Bigdeli TB, Edwards AC, Webb BT, Bacanu SA, Zaitlen N, Flint J and Kendler KS** (2018) Molecular genetic analysis subdivided by adversity exposure suggests etiologic heterogeneity in major depression. *The American Journal of Psychiatry* **175**, 545–554.

**Schizophrenia Working Group of the Psychiatric Genomics Consortium** (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427.

**Sim SC, Nordin L, Andersson TM, Virding S, Olsson M, Pedersen NL and Ingelman-Sundberg M** (2010) Association between CYP2C19 polymorphism and depressive symptoms. *The American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **153b**, 1160–1166.

**Singh AK, Zajdel J, Mirrasekhian E, Almoosawi N, Frisch I, Klawonn AM, Jaarola M, Fritz M and Engblom D** (2017) Prostaglandin-mediated inhibition of serotonin signaling controls the affective component of inflammatory pain. *The Journal of Clinical Investigation* **127**, 1370–1374.

**The 1000 Genomes Project Consortium**, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT and

https://doi.org/10.1017/S0033291719000618 Published online by Cambridge University Press

McVean GA (2012) An integrated map of genetic variation from 1092 human genomes. *Nature* 491, 56–65.

**Vilhjalmsson BJ, Yang J, Finucane HK, Gusev A, Lindstrom S, Ripke S, Genovese G, Loh PR, Bhatia G, Do R, Hayeck T, Won HH, Kathiresan S, Pato M, Pato C, Tamimi R, Stahl E, Zaitlen N, Pasaniuc B, Belbin G, Kenny EE, Schierup MH, De Jager P, Patsopoulos NA, McCarroll S, Daly M, Purcell S, Chasman D, Neale B, Goddard M, Visscher PM, Kraft P, Patterson N and Price AL** (2015) Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *The American Journal of Human Genetics* **97**, 576–592.

**Wang Y, Lu J, Yu J, Gibbs RA and Yu F** (2013) An integrative variant analysis pipeline for accurate genotype/haplotype inference in population NGS data. *Genome Research* **23**, 833–842.

https://doi.org/10.1017/S0033291719000618 Published online by Cambridge University Press