nature genetics

Identification of 15 genetic loci associated with risk of major depression in individuals of European descent

Craig L Hyde[1], Michael W Nagle[2], Chao Tian[3], Xing Chen[1], Sara A Paciga[2], Jens R Wendland[2], Joyce Y Tung[3], David A Hinds[3], Roy H Perlis[4] & Ashley R Winslow[2,5]

Despite strong evidence supporting the heritability of major depressive disorder (MDD), previous genome-wide studies were unable to identify risk loci among individuals of European descent. We used self-report data from 75,607 individuals reporting clinical diagnosis of depression and 231,747 individuals reporting no history of depression through 23andMe and carried out meta-analysis of these results with published MDD genome-wide association study results. We identified five independent variants from four regions associated with self-report of clinical diagnosis or treatment for depression. Loci with a *P* value $<1.0 \times 10^{-5}$ in the meta-analysis were further analyzed in a replication data set (45,773 cases and 106,354 controls) from 23andMe. A total of 17 independent SNPs from 15 regions reached genome-wide significance after joint analysis over all three data sets. Some of these loci were also implicated in genome-wide association studies of related psychiatric traits. These studies provide evidence for large-scale consumer genomic data as a powerful and efficient complement to data collected from traditional means of ascertainment for neuropsychiatric disease genomics.

MDD remains one of the most important contributors to morbidity and mortality[1–3]. Efforts to develop novel interventions have been hindered by a limited understanding of the underlying neurobiology. Despite strong evidence of heritability[4,5], efforts to clarify this biology through common or rare variant association studies have been unsuccessful, with this lack of success attributed to heterogeneity of disease and absence of a biological gold-standard diagnosis. One recent study of a Han Chinese population identified two risk loci, in the *LHPP* gene and near the *SIRT1* gene, but neither was supported in European populations where the risk alleles are extremely rare[6].

If a reasonable strategy to overcome disease heterogeneity is to develop more precise or refined phenotypes, another is to efficiently identify much larger cohorts for study despite less intensive phenotyping. The latter strategy has been validated in multiple non-psychiatric diseases but not in psychiatric illness, where phenotyping is presumed to require more detailed interview. Here we identified 75,607 individuals (62% female) who endorsed a previous clinical diagnosis of or treatment for major depression and 231,747 individuals (44% female) reporting no clinical diagnosis of or treatment for depression. All subjects participated in the consumer genomics company 23andMe's optional research initiative (for population sociodemographic features, see **Table 1**). These individuals were genotyped on one of four custom arrays containing genome-wide content, and genotypes were imputed using the September 2013 release of the 1000 Genomes Project Phase 1 reference haplotypes. Research participants with >97% European ancestry, excluding close relatives, were included in the genome-wide association study (GWAS) analysis. The Manhattan plot and quantile–quantile plot for the analysis are shown in **Supplementary Figure 1a**,**b**; *P* values were adjusted for inflation using LD score regression (**Supplementary Table 1**).

RESULTS

New major depression loci in a self-report population

From the discovery 23andMe data set, we identified two distinct regions containing SNPs with *P* values $<1 \times 10^{-8}$ and five additional loci with *P* values $<5 \times 10^{-8}$ (**Supplementary Table 1**) associated with self-report of depression. We have chosen to consider only SNPs with *P* values $<1 \times 10^{-8}$ to be genome-wide significant in this GWAS because of correction for 15 million SNPs in the 23andMe data. The most significant locus yielded an association at rs2806933 (adjusted *P* value $= 8.53 \times 10^{-13}$; odds ratio (OR) = 0.955, 95% confidence interval (CI) = 0.943–0.968; effect allele frequency in controls = 0.61) in a region spanning the 3′ UTR of the *OLFM4* gene (encoding olfactomedin-4). This gene has not previously been implicated in neuropsychiatric disease but is known to be expressed in brain, including in the amygdala and medial temporal lobe[7]. The second most significant locus, with peak association at rs768705 ($P = 2.91 \times 10^{-12}$; OR = 1.051, 95% CI = 1.036–1.067; effect allele frequency in controls = 0.25), spans a locus containing *MEF2C* (myocyte enhancer factor 2C) and *TMEM161B* (transmembrane protein 161B). Variants in *MEF2C*

[1]Statistics, Pfizer Global Research and Development, Pfizer, Inc., Cambridge, Massachusetts, USA. [2]Human Genetics and Computational Biomedicine, Pfizer Global Research and Development, Pfizer, Inc., Cambridge, Massachusetts, USA. [3]23andMe, Inc., Mountain View, California, USA. [4]Center for Experimental Drugs and Diagnostics, Center for Human Genetic Research and Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA. [5]Present address: Orphan Disease Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. Correspondence should be addressed to A.R.W. (awinslow@mail.med.upenn.edu), R.H.P. (rperlis@mgh.harvard.edu) or D.A.H. (dhinds@23andme.com).

Received 11 March; accepted 20 June; published online 1 August 2016; doi:10.1038/ng.3623

Table 1 Cohort demographics for the primary and replication 23andMe data sets

	Discovery		Replication	
	MDD	Controls	MDD	Controls
Total (*n*)	75,607	231,747	45,773	106,354
Age, percentage (years)				
Under 30	12.1	11.6	13.8	13.4
30–45	29.9	27.5	29.8	25.4
45–60	28.8	27.2	29.6	27.7
Over 60	29.3	33.7	26.7	33.3
Sex, percentage				
Male	38.0	56.2	33.8	52.6
Female	62.0	43.8	66.2	47.4

have previously been associated with multiple central nervous system (CNS) phenotypes including epilepsy and intellectual disability[8,9] and have been implicated in regulation of synaptic function[10]. *TMEM161B* is also expressed in the brain, and *Tmem161b* exhibits decreased levels of repressive dimethylation of histone H3 at lysine 9 and lysine 27 in response to social isolation in a mouse model of depression[11]. Whereas schizophrenia and Alzheimer disease GWAS both identify the *MEF2C* region as a disease susceptibility locus, the peak schizophrenia- and Alzheimer disease–associated SNPs are not in strong linkage disequilibrium (LD) with the MDD-associated SNP (schizophrenia: rs181900 (ref. 12), $r^2 = 0.001$; Alzheimer disease: rs190982, $r^2 = 0.016$). Using a population prevalence of 15% for MDD estimated by the Psychiatric Genomics Consortium (PGC)[13], we calculated heritability using LD score regression of $h^2_{\text{liability}} = 0.0528$ for this data set. When using the 23andMe observed population prevalence for MDD of 25%, the heritability is $h^2_{\text{liability}} = 0.0612$.

Results from the meta-analysis of the 23andMe data set with the previously reported PGC meta-analysis of MDD, which encompassed 9,240 cases and 9,519 controls of European descent, are presented in **Figure 1a**,**b** (**Supplementary Table 2**). In the PGC cohort, only 1.22 million SNPs overlapped with the 23andMe MDD data (no results were reported for the X or Y chromosome)[14], and only these SNPs were used for downstream analysis. As a result, several lead SNPs from the discovery 23andMe GWAS were absent, including rs77741769 (*SPPL3–HNF1A*), rs144294997 (*N6AMT1*), rs1432639 (*NEGR1*), and rs67744457 (*EP300–L3MBTL2*). Each cohort was individually adjusted for test statistic inflation using LD score regression (Online Methods), and the combined cohorts were subsequently subjected to meta-analysis using a standard fixed-effects, inverse-variance-weighted approach[15]. The final results from the meta-analysis were further adjusted for the meta-analysis LD score regression intercept of 1.0025.

Of the original 23andMe lead SNPs, only the *N6AMT1* locus was not represented in the meta-analysis results at a *P* value less than 5×10^{-6} because of absence of the lead 23andMe SNP in the meta-data set as well as an absence of significant secondary signals in the region. SNPs in the *OLFM4*, *TMEM161B–MEF2C* (two independent SNPs), *MEIS2–TMCO5A*, and *NEGR1* regions reached genome-wide significance in the meta-analysis ($P < 5 \times 10^{-8}$, correcting for 1.22 million SNPs) (**Supplementary Table 2**). Regional association plots are shown for these regions in **Figure 2**. Heritability for the meta-analysis was estimated at $h^2_{\text{liability}} = 0.059$ and 0.069, with prevalence of 15% and 25%, respectively.

Replication of 15 loci associated with major depression

We assessed the ability of the top signals ($P < 1 \times 10^{-5}$) from the meta-analysis to replicate in a separate cohort of 45,773 cases and 106,354 controls from 23andMe (**Table 1**). All individuals in the replication data set were independent from subjects included in the discovery 23andMe data set and had similar characteristics for sex and age distributions. The replication cohort provided additional support for three of the five genome-wide-significant SNPs, in the *TMEM161B–MEF2C* (two SNPs) and *NEGR1* (one SNP) loci. In a joint analysis of the discovery 23andMe data set, the PGC data set, and the 23andMe replication data set, a total of 15 independent loci (17 SNPs) reached genome-wide significance

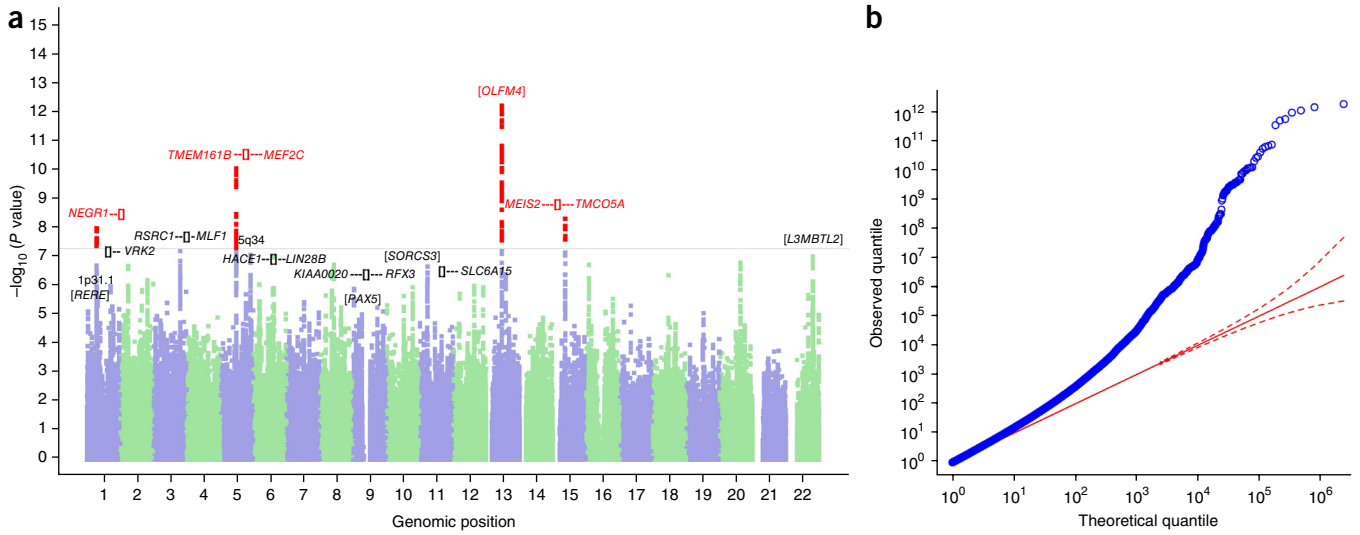

Figure 1 Discovery-phase meta-analysis of 23andMe self-report ascertainment of major depression (75,607 cases and 231,747 controls) and PGC MDD (9,240 cases and 9,519 controls). (**a**) Manhattan plot of the discovery-phase 23andMe GWAS. The intercept calculated by LD score regression was used for inflation correction. The threshold for genome-wide significance ($P < 5 \times 10^{-8}$) is represented by the horizontal purple line. Red dots correspond to SNPs with *P* values smaller than the genome-wide significance threshold. Regions labeled in black denote loci that reached genome-wide significance in the joint analysis. The hg19 release of the UCSC Genome Browser was used for mapping. SNP location is denoted by []. If the SNP occurs between genes, then the distance from those genes is denoted by dashes: no dash, ≤1 kb; -, ≤10 kb; --, ≤100 kb; ---, ≤1,000 kb. (**b**) Quantile–quantile plot for the 23andMe MDD GWAS.

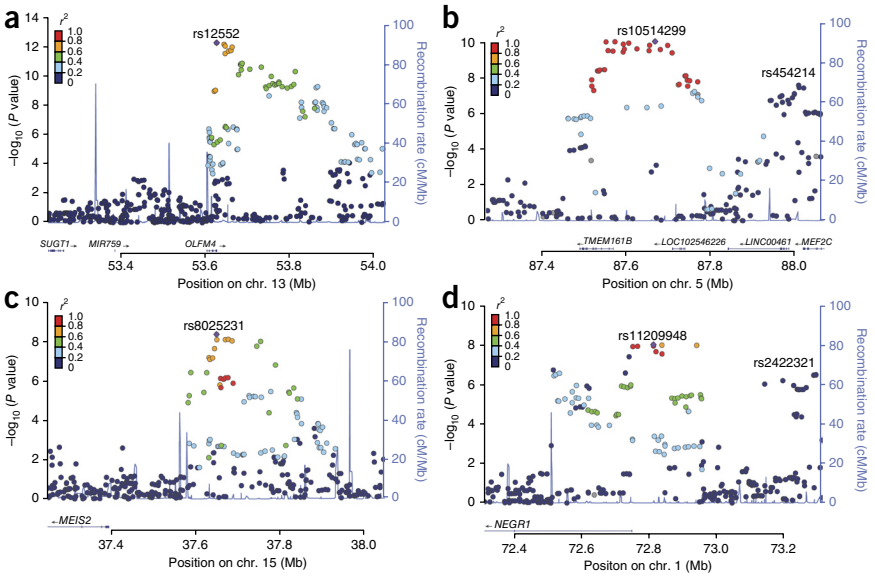

Figure 2 Regional association plots for genome-wide-significant regions and secondary independent signals identified in each region. (**a**–**d**) *OLFM4* locus (rs12552) (**a**), *TMEM161B–MEF2C* locus (rs10514299) (**b**), *MEIS2–TMCO5A* locus (rs8025231) (**c**), and *NEGR1* locus (rs11209948) (**d**). Secondary signals in the *TMEM161B–MEF2C* and *NEGR1* regions (rs454214 and rs2422321, respectively) are shown. Purple diamonds represent the SNP with the smallest *P* value for each locus.

($P < 5 \times 10^{-8}$) (**Table 2**). Of the remaining 46 SNPs with a *P* value less than 1×10^{-5} in the meta-analysis of the 23andMe discovery data set and the 23andMe replication, 41 had a consistent direction of effect in the meta-analysis and replication cohorts (*P* values across all analyses, including joint analysis, are shown in **Supplementary Table 2** for SNPs that reached $P < 1 \times 10^{-5}$ in the meta-analysis).

To explore the biological implications of our findings, we used DEPICT to derive tissue enrichment, gene set enrichment, and gene predictions (**Supplementary Table 3**) for SNPs with a *P* value less than 1×10^{-5} in the meta-analysis. Although identification of the functional variant or gene for each locus is not straightforward, many of the top associations in our data set appeared in or near genes encoding transcription factors with known CNS developmental functions (for additional gene predictions from DEPICT and functional annotation for each region, see **Supplementary Table 4**). Gene set enrichment analysis prioritized the *MEIS2* subnetwork ($P = 2.30 \times 10^{-6}$). *MEIS2* encodes a TALE homeodomain transcription factor known to function in development. Although most studies implicate *MEIS2* in peripheral tissue development, recent studies have shown a role for Meis2-regulated pathways in neurogenesis through interactions with Pax6, as well as interactions with Pax3 and Pax7 (refs. 16,17). Notably, our analysis identified significant associations with MDD in the *MEIS2*, *PAX6*, and *PAX5* regions ($P = 2.04 \times 10^{-8}$, 3.94×10^{-7}, and 2.59×10^{-5}, respectively, in the 23andMe discovery data set). Tissue enrichment analysis showed an over-representation of the CNS, and 12 of the 19 nominally associated tissues were from different brain regions (with 'nervous system' as a second-level MeSH term). Although these associations did not pass multiple-testing correction, the top results from our MDD GWAS correspond to genes enriched for CNS expression and transcriptional function important for CNS development or neurogenesis. Further functional annotations of predicted genomic and molecular functions, annotations of

Table 2 Summary statistics for 17 SNPs reaching genome-wide significance ($P < 5 \times 10^{-8}$) in the joint analysis (23andMe discovery, PGC, and 23andMe replication)

rs ID	Gene context	23andMe	PGC	Meta-analysis	Replication	Joint	Sign match
rs10514299	*TMEM161B*--[]---*MEF2C*	4.35×10^{-12}	5.73×10^{-1}	8.50×10^{-11}	1.15×10^{-4}	9.99×10^{-16}	Yes
rs1518395	[]--*VRK2*	1.45×10^{-7}	5.83×10^{-1}	2.01×10^{-7}	1.50×10^{-5}	4.32×10^{-12}	Yes
rs2179744	[*L3MBTL2*]	4.34×10^{-8}	7.74×10^{-1}	9.24×10^{-8}	7.26×10^{-4}	6.03×10^{-11}	Yes
rs11209948	*NEGR1*--[]	4.41×10^{-8}	8.57×10^{-2}	1.01×10^{-8}	9.39×10^{-4}	8.38×10^{-11}	Yes
rs454214	*TMEM161B*---[]--*MEF2C*	6.28×10^{-8}	1.62×10^{-1}	2.42×10^{-8}	6.39×10^{-3}	1.09×10^{-9}	Yes
rs301806	[*RERE*]	3.72×10^{-6}	8.68×10^{-1}	7.28×10^{-6}	2.52×10^{-4}	1.90×10^{-9}	Yes
rs1475120	*HACE1*--[]--*LIN28B*	2.32×10^{-6}	2.29×10^{-1}	1.14×10^{-6}	9.27×10^{-4}	4.17×10^{-9}	Yes
rs10786831	[*SORCS3*]	1.75×10^{-6}	2.98×10^{-1}	1.09×10^{-6}	2.33×10^{-3}	8.11×10^{-9}	Yes
rs12552	[*OLFM4*]	1.23×10^{-12}	1.57×10^{-1}	5.74×10^{-13}	8.70×10^{-1}	8.16×10^{-9}	No
rs6476606	[*PAX5*]	2.59×10^{-5}	1.52×10^{-1}	9.30×10^{-6}	1.94×10^{-4}	1.20×10^{-8}	Yes
rs8025231	*MEIS2*---[]---*TMCO5A*	2.04×10^{-8}	8.49×10^{-2}	4.66×10^{-9}	7.50×10^{-2}	1.23×10^{-8}	Yes
rs12065553	[]	8.53×10^{-7}	8.67×10^{-1}	2.88×10^{-6}	6.79×10^{-3}	1.32×10^{-8}	Yes
rs1656369	*RSRC1*--[]-*MLF1*	8.19×10^{-8}	3.22×10^{-1}	6.05×10^{-8}	3.56×10^{-2}	1.34×10^{-8}	Yes
rs4543289	[]	1.19×10^{-6}	1.15×10^{-2}	8.23×10^{-8}	5.26×10^{-3}	1.36×10^{-8}	Yes
rs2125716	[]---*SLC6A15*	4.33×10^{-7}	8.78×10^{-1}	9.58×10^{-7}	2.24×10^{-2}	3.05×10^{-8}	Yes
rs2422321	*NEGR1*---[]	5.12×10^{-6}	6.91×10^{-3}	3.28×10^{-7}	3.13×10^{-3}	3.18×10^{-8}	Yes
rs7044150	*KIAA0020*---[]---*RFX3*	3.97×10^{-7}	9.64×10^{-1}	1.24×10^{-6}	3.05×10^{-2}	4.31×10^{-8}	Yes

The corresponding *P* values are shown for each phase of analysis: 23andMe MDD (discovery data set), PGC MDD, meta-analysis (23andMe discovery + PGC), 23andMe replication (23andMe replication cohort), and joint analysis. *P* values were adjusted for the 23andMe discovery, PGC, and meta-analysis data (but meta-analysis and replication *P* values were not adjusted for inflation in joint analysis). Adjustment was performed using the intercepts calculated by LD score regression. Sign match indicates whether the directions of effect were matched for the meta-analysis and 23andMe replication data sets. Text representation of SNP location in relation to genes in each region is shown, with SNP location denoted by []. If the SNP occurs between genes, the distance from these genes is denoted by dashes: no dash, ≤1 kb; -, ≤10 kb, --, ≤100 kb; ---, ≤1,000 kb. The hg19 release of the UCSC Genome Browser was used for mapping. The GWAS results were peak pruned by distance (<300 kb) and LD (r^2 >0.1).

Table 3 MDD gene risk score association with secondary phenotypes

Phenotype	*n*	Effect (s.e.)	*P* value	FDR
Early onset	94,891	0.283 (0.095)	2.90×10^{-3}	3.20×10^{-3}
Age of onset	94,891	−1.49 (0.372)	6.10×10^{-5}	8.40×10^{-5}
Anxiety	250,528	0.323 (0.061)	1.00×10^{-7}	2.50×10^{-7}
Panic attacks	247,167	0.319 (0.072)	9.80×10^{-6}	1.50×10^{-5}
Insomnia	248,576	0.272 (0.051)	1.10×10^{-7}	2.50×10^{-7}
Taking an SSRI	52,698	0.448 (0.162)	5.50×10^{-3}	5.50×10^{-3}
Medication for mental health	349,287	0.421 (0.057)	1.40×10^{-13}	1.50×10^{-12}
Prescription sleep aid	350,119	0.184 (0.05)	2.70×10^{-4}	3.20×10^{-4}
Prescription pain medication	346,989	0.236 (0.041)	5.60×10^{-9}	3.10×10^{-8}
Overweight (BMI >27)	401,552	0.212 (0.038)	3.00×10^{-8}	1.10×10^{-7}
Obesity (BMI >30)	401,552	0.216 (0.045)	1.50×10^{-6}	2.70×10^{-6}

The GRS is explained in **Supplementary Table 6**. Analysis of association with age of onset for MDD was conducted in subjects with MDD. Association analysis for all other traits was conducted in cases and controls, and results were adjusted for case/control status from the general 23andMe research community. s.e., standard error.

brain tissue or monocyte expression quantitative trait loci (eQTLs), gene predictions for each region using DEPICT[18], and annotations of disease associations from publicly available GWAS data sets and the Online Mendelian Inheritance in Man (OMIM) database are presented in **Supplementary Table 4** for all 17 SNPs reaching genome-wide significance in the joint analysis (**Table 2**).

After SNPs were pruned for distance (separated by <300 kb) and LD (r^2 >0.1), three regions had multiple SNPs with *P* values less than 1×10^{-6} in the meta-analysis results. We tested SNPs in each of these regions for independence in the replication data set using Wald and likelihood-ratio tests. We conducted this analysis in the replication data set to avoid SNP selection bias from the original findings. By conditioning on each SNP in the models at each locus, we found that two SNPs each in the *TMEM161B–MEF2C* and *NEGR1* regions are likely independent (rs10514299 and rs454214 for *TMEM161B–MEF2C* and rs11209948 and rs2422321 for *NEGR1*), with the variance in the region being explained best by both SNPs, whereas most of the variance in the *MLF1* region was explained by rs1656369 alone (with no additional significance provided by inclusion of rs4645169) (**Supplementary Table 5**).

Validity of self-report phenotype for major depression

As the PGC cohort is substantially smaller than the 23andMe single cohort, power in the PGC MDD GWAS to detect the effect sizes for the two genome-wide-significant loci observed in the preliminary 23andMe GWAS was less than 0.6 at a nominal level of significance ($P < 0.05$, uncorrected), and the analogous power to replicate the remaining 23andMe loci in the PGC data set declined further below this level[19]. However, the probability of the loci in the PGC data set showing the same direction of effect as in the 23andMe data set exceeded 90% for each of the top ten independent 23andMe loci that were also evaluated in the PGC data (corresponding to all overlapping peak-pruned 23andMe loci with unadjusted $P < 1 \times 10^{-7}$ in the 23andMe data). We therefore conducted a sign test examining concordance between the PGC effect direction and the 23andMe effect direction for the top overlapping 23andMe peak loci. Nine of the top ten loci had matching signs (Fisher's exact test, $P = 0.033$). The test results deviated significantly from chance at a range of thresholds, suggesting consistent signal for the PGC results and 23andMe data. For the 82 independent SNPs with nominal *P* values less than 1×10^{-5} in 23andMe, the *P* value for the sign test was 2×10^{-6} with the odds ratio for a sign match of 10.6 (95% CI = 3.5–37.1). Furthermore, the effect sizes for the top independent 23andMe loci were correlated with the effect sizes for these SNPs in PGC (removing loci with minor allele frequency (MAF) <5% to avoid highly variable values). This correlation peaked at the 39th peak 23andMe locus with 68% correlation ($P = 2.5 \times 10^{-9}$). Additionally, we calculated the genetic correlation between the two data sets using LD score regression[20] and found that the two major depression data sets were highly and positively correlated ($r_g = 0.725$, s.e.m. = 0.093; $P = 7.05 \times 10^{-15}$).

Associations of lead SNPs with related phenotypes

To investigate the polygenic nature of MDD, we generated a genetic risk score (GRS) from 17 SNPs (**Supplementary Table 6**) with *P* values $< 5 \times 10^{-8}$ in the joint analysis (discovery 23andMe, PGC, and replication 23andMe) and tested for association of the weighted MDD GRS with reporting of related phenotypes, medication use, and age at onset (**Table 3**) in the combined discovery and replication cohort, adjusting for depression case/control status. The GRS was significantly associated (false discovery rate (FDR) < 0.05) with each of these phenotypes. Notably, the MDD GRS was significantly associated with an earlier age of onset in cases (effect = −1.49 years per unit of log odds, standard error = 0.37; $P = 6.1 \times 10^{-5}$).

The independent effect of each GRS SNP on this set of related phenotypes is presented in **Supplementary Table 7**. Although rs12552 in the *OLFM4* region was not strongly supported in the replication data set, this SNP was associated with increased reporting of panic attacks; use of medication to treat mental health problems, prescription sleep aids, and pain medication; body mass index (BMI) greater than 27; and earlier age at onset of MDD and was commensurately associated with lower continuous age of onset. Individually, rs12552 and rs4543289 had the largest effects on age at onset, with a total of five SNPs having nominal significance ($P < 0.05$).

Table 4 Cross-trait genetic correlation with 23andMe MDD (LD score regression)

Phenotype	r_g (s.e.)	Nominal *P* value	Cohort observed h^2	Significance after Bonferroni correction
PGC MDD	0.725 (0.093)	7.05×10^{-15}	0.128	*
PGC SCZ1	0.23 (0.042)	4.028×10^{-8}	0.543	*
PGC SCZ1 + SWE	0.261 (0.036)	8.132×10^{-13}	0.411	*
PGC SCZ2	0.282 (0.03)	2.182×10^{-21}	0.371	*
PGC bipolar disorder	0.264 (0.049)	7.446×10^{-8}	0.350	*
IGAP AD	−0.069 (0.071)	0.3331	0.039	NS
IPDGC PD (2012)	0.185 (0.091)	0.04123	0.200	NS
GLGC LDL	0.056 (0.031)	0.072	0.191	NS

The observed heritability score (h^2) for the 23andMe discovery cohort is 0.038. Genetic correlation (r_g) of the 23andMe discovery MDD data set with related psychiatric disorders (PGC MDD, PGC SCZ1, PGC SCZ1 + SWE, PGC SCZ2, PGC bipolar disorder), non-psychiatric neurological disorders (IGAP Alzheimer disease (AD), IPDGC Parkinson disease (PD)), and a non-psychiatric and non-neurological trait (GLGC LDL) are shown. NS, not significant. An asterisk indicates significance after Bonferroni correction ($\alpha = 0.00625$). PGC, Psychiatric Genomics Consortium data sets; MDD, major depressive disorder; SCZ1, Schizophrenia (2011 data set); SCZ1 + SWE, SCZ1 and Schizophrenia Swedish National Study meta-analysis; SCZ2, Schizophrenia (2014 data set); IGAP, International Genomics of Alzheimer's Project; IPDGC, International Parkinson Disease Genomics Consortium; GLGC LDL, Global Lipids Genetics Consortium LDL data set.

Sex effects

Because of known disparities between the sexes in the presentation of depression and incidence rate and because of the suggestion of differences in underlying biology, we tested for sex-specific effects for our top SNPs as well as genotype–sex interaction for each SNP in the 23andMe discovery cohort (**Supplementary Table 8**). In the discovery cohort, four SNPs had nominal association at $P < 0.05$, but none survived multiple-testing correction. No results reached nominal significance at $P < 0.05$ in the replication cohort. Our GWAS results thus provide no support for differences between the sexes in genetic predisposition to depression.

Cohort characteristics

We further validated the new self-reported phenotype by assessing expected characteristics of medication use, comorbid symptoms, and risk factors commonly seen in MDD in the 23andMe self-report cohort (**Supplementary Table 9**). Reporting of anxiety, panic attacks, and insomnia was significantly increased ($P < 5.0 \times 10^{-243}$ for all traits tested) among subjects reporting depression as well as a BMI greater than 27 (overweight) or a BMI greater than 30 (obese). Reporting of current selective serotonin reuptake inhibitor (SSRI) use, medication for mental health problems, prescription sleep aids, and pain medication was also increased, with the highest odds ratios for any trait tested observed for SSRI and psychotropic use (13.35 and 44.83, respectively), further supporting the validity of the phenotype ascertainment. Cohort characteristics were also tested separately in males and females, with no evidence of sex-specific differences (**Supplementary Table 10**).

Studies have shown a degree of shared genetic liability for different psychiatric disorders, which is likely a result of multiple factors, including genetic pleiotropy, diagnostic overlap, comorbid disease, or disease progression. To initially assess shared genetic risk across psychiatric disorders, we obtained *P* values across five psychiatric traits (schizophrenia, bipolar disorder, neuroticism, depressive symptoms, and subjective well-being) for SNPs with *P* values less than 1×10^{-5} in the MDD meta-analysis (**Supplementary Table 11**; data for neuroticism, depressive symptoms, and subjective well-being are from ref. 21). The MDD-associated SNPs showed the highest degree of overlap (smallest *P* values) in the schizophrenia data set, followed by neuroticism, with less replication in the bipolar disorder, depressive symptoms, and subjective well-being phenotypes. Schizophrenia and bipolar disorder GWAS data were from the publicly available PGC data sets[12,22], and corresponding *P* values for neuroticism, depressive symptoms, and subjective well-being were provided by the Social Science Genetics Association Consortium (SSGAC). The lack of correlation with SSGAC depressive symptoms self-report data may arise from the differences between acute and chronic depression or clinical severity, given that subjects in the SSGAC study were asked to report on depressive symptoms arising in the 2 weeks before the survey, as compared to lifetime major depression in the primary cohorts. Conversely, the trait measure of neuroticism has previously been show to overlap with major depression, consistent with our results.

To rigorously assess the genetic correlation of the MDD GWAS findings with those for other neuropsychiatric disease, we used available GWAS from PGC, including a bipolar disorder and three schizophrenia GWAS (different versions of the PGC schizophrenia data sets), as well as neurodegenerative disease GWAS, to test pairwise genetic correlation with the 23andMe MDD GWAS data set using LD score regression. Because shared controls are used for the PGC data sets, we did not use the results from the meta-analysis of the 23andMe and PGC data. The highest correlation with the primary 23andMe GWAS results was observed for the PGC SCZ2 schizophrenia GWAS ($r = 0.282$, s.e.m. = 0.03; $P = 2.18 \times 10^{-21}$), followed by the bipolar disorder and additional schizophrenia GWAS (**Table 4**); however, we observed little to no correlation for the Parkinson disease and Alzheimer disease data sets. Additionally, we checked for correlation between MDD in the 23andMe data set and a trait with no known epidemiological correlation with depression (LDL cholesterol) and observed no genetic correlation between the two traits.

DISCUSSION

In this study, we present a complementary approach to collecting large-scale genotypic data on major depression. By using the self-report data on major depression from 23andMe, we were able to identify SNPs at a genome-wide level of significance associated with risk for depression in a cohort of European descent. Through a meta-analysis of the 23andMe data with PGC MDD GWAS and a joint analysis with an independent 23andMe replication cohort, we identify 17 independent SNPs significantly associated with diagnosis of major depression. Through tissue and gene set enrichment analyses using DEPICT, we find that these SNPs are predicted to be enriched in genes that are expressed in the CNS and function in transcriptional regulation related to neurodevelopment. We find no robust evidence for sex specificity of effects among our top results, but this study combined both sexes and only adjusted for sex as a covariate and was therefore not structured to identify sex-specific loci. This would ideally be done through a sex-stratified GWAS.

Although the variance explained by these SNPs is small, we find that our cohorts identified by self-report of major depression are highly genetically correlated with cohorts identified by clinical interview, a result further corroborated by significant sign tests and effect size matching between the top 23andMe SNPs (nominal $P < 1 \times 10^{-5}$) and their counterparts in PGC in self-report and clinical interview data sets. To better understand the phenotypic characteristics of the 23andMe self-report subjects, we assessed reporting of medication use and comorbidities and found that all tested characteristics were significantly increased in the subjects reporting depression, similar to what is seen in clinically ascertained subjects. Many of the most significant SNPs show evidence of pleiotropy when examined in other clinically ascertained psychiatric disorders, with the smallest *P* values among individual SNPs seen for MDD-associated SNPs in the PGC schizophrenia and neuroticism data sets. This finding is unsurprising given the pleiotropy reported by other GWAS and cross-psychiatric-disorder analyses[13] and lends further support to the relevance of a self-reported phenotype to clinical disease.

We were unable to replicate the genome-wide-significant loci identified in the recent CONVERGE study[6], although we identified modest associations in each region (*LHPP*: rs145655839, minimum $P = 0.0024$ out of 6,204 SNPs in the region; *SIRT1*: rs187810158, minimum $P = 0.0102$ out of 5,111 SNPs in the region). This result is unsurprising given that our study looked for genetic determinants of susceptibility in both males and females of European descent and likely represented a very different population structure than that of the CONVERGE study of Han Chinese women.

Taken together, our results indicate the usefulness of a strategy complementary to intensive phenotyping for identifying common variant associations with phenotypically heterogeneous neuropsychiatric diseases. The inter-rater reliability of lifetime MDD diagnosis even with structured interview is modest, with a κ value of 0.32–0.57 (refs. 23,24); conversely, the reliance on treatment-seeking patients in the present analysis rather than volunteers responding to advertisements lends additional face validity to the phenotype[25]. The finding

in other large-scale analyses that cohorts ascertained on the basis of treatment versus structured interview yield similar associations[12] and that such phenotypes are consistent with those defined by structured interview[26] adds confidence to the validity of this approach[12]. In light of the massive impact of such disorders on health worldwide, any approach that can help elucidate pathophysiology merits consideration. The finding that a locus previously linked to other neuropsychiatric disease increases MDD risk also adds to a growing literature indicating the pleiotropy of such risk-associated genes.

URLs. Psychiatric Genomics Consortium (PGC MDD, bipolar disorder, SCZ1, SCZ1 + SWE, and SCZ2 data sets), https://www.med.unc.edu/pgc; International Genomics of Alzheimer's Project (IGAP AD data set), http://www.pasteur-lille.fr/en/recherche/u744/igap/igap_download.php; International Parkinson Disease Genomics Consortium (IPDGC PD data set), http://www.pdgene.org/; Global Lipids Genetics Consortium (GLGC LDL data set), http://csg.sph.umich.edu//abecasis/public/lipids2013/.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We would like to thank the research participants and employees of 23andMe for making this work possible. The authors thank the investigators and patient participants of the Psychiatric Genomic Consortium Major Depressive Disorder study for making the PGC MDD phase 1 results available for download. This study was supported in part by Pfizer, Inc. R.H.P. is supported in part by the National Institute of Mental Health and the National Human Genome Research Institute (P50 MH106933). We also thank the Social Science Genetics Association Consortium (SSGAC) for sharing results for subjective well-being, depressive symptoms, and neuroticism.

AUTHOR CONTRIBUTIONS

A.R.W., C.L.H., and J.R.W. conceived the meta-analysis and statistical analysis. A.R.W., C.L.H., and R.H.P. oversaw data set analysis and primary data interpretation. C.L.H. designed and performed meta-analysis and further statistical analysis of the three data sets. X.C. provided statistical support and data visualization for the meta-analysis. M.W.N. provided DEPICT functional annotation and LD score regression analyses. R.H.P., A.R.W., and C.L.H. wrote the manuscript. A.R.W., R.H.P., C.L.H., D.A.H., S.A.P., and M.W.N. provided data interpretation and revised the manuscript. J.Y.T. and D.A.H. conceived and designed the 23andMe MDD GWAS. D.A.H. and C.T. performed GWAS for 23andMe data sets and statistical support.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Angst, F., Stassen, H.H., Clayton, P.J. & Angst, J. Mortality of patients with mood disorders: follow-up over 34-38 years. *J. Affect. Disord.* **68**, 167–181 (2002).
2. Lopez, A.D., Mathers, C.D., Ezzati, M., Jamison, D.T. & Murray, C.J. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* **367**, 1747–1757 (2006).
3. Wittchen, H.U. *et al.* The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur. Neuropsychopharmacol.* **21**, 655–679 (2011).
4. Lichtenstein, P. *et al.* Recurrence risks for schizophrenia in a Swedish national cohort. *Psychol. Med.* **36**, 1417–1425 (2006).
5. Sullivan, P.F., Kendler, K.S. & Neale, M.C. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch. Gen. Psychiatry* **60**, 1187–1192 (2003).
6. CONVERGE Consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* **523**, 588–591 (2015).
7. Hawrylycz, M.J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
8. Le Meur, N. *et al. MEF2C* haploinsufficiency caused by either microdeletion of the 5q14.3 region or mutation is responsible for severe mental retardation with stereotypic movements, epilepsy and/or cerebral malformations. *J. Med. Genet.* **47**, 22–29 (2010).
9. Paciorkowski, A.R. *et al. MEF2C* haploinsufficiency features consistent hyperkinesis, variable epilepsy, and has a role in dorsal and ventral neuronal developmental pathways. *Neurogenetics* **14**, 99–111 (2013).
10. Barbosa, A.C. *et al.* MEF2C, a transcription factor that facilitates learning and memory by negative regulation of synapse numbers and function. *Proc. Natl. Acad. Sci. USA* **105**, 9391–9396 (2008).
11. Wilkinson, M.B. *et al.* Imipramine treatment and resiliency exhibit similar chromatin regulation in the mouse nucleus accumbens in depression models. *J. Neurosci.* **29**, 7820–7832 (2009).
12. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
13. Cross-Disorder Group of the Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381**, 1371–1379 (2013).
14. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol. Psychiatry* **18**, 497–511 (2013).
15. Whitehead, A. *Meta-Analysis of Controlled Clinical Trials* (John Wiley & Sons, 2002).
16. Agoston, Z. *et al.* Meis2 is a Pax6 co-factor in neurogenesis and dopaminergic periglomerular fate specification in the adult olfactory bulb. *Development* **141**, 28–38 (2014).
17. Agoston, Z., Li, N., Haslinger, A., Wizenmann, A. & Schulte, D. Genetic and physical interaction of Meis2, Pax3 and Pax7 during dorsal midbrain development. *BMC Dev. Biol.* **12**, 10 (2012).
18. Pers, T.H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat. Commun.* **6**, 5890 (2015).
19. Purcell, S., Cherny, S.S. & Sham, P.C. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19**, 149–150 (2003).
20. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
21. Okbay, A. *et al.* Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* **48**, 624–633 (2016).
22. Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near *ODZ4*. *Nat. Genet.* **43**, 977–983 (2011).
23. Keller, M.B. *et al.* Results of the DSM-IV mood disorders field trial. *Am. J. Psychiatry* **152**, 843–849 (1995).
24. Regier, D.A. *et al.* DSM-5 field trials in the United States and Canada, Part II: test–retest reliability of selected categorical diagnoses. *Am. J. Psychiatry* **170**, 59–70 (2013).
25. Wisniewski, S.R. *et al.* Can phase III trial results of antidepressant medications be generalized to clinical practice? A STAR*D report. *Am. J. Psychiatry* **166**, 599–607 (2009).
26. Castro, V.M. *et al.* Validation of electronic health record phenotyping of bipolar disorder cases and controls. *Am. J. Psychiatry* **172**, 363–372 (2015).

ONLINE METHODS

Data access. The full GWAS summary statistics for the 23andMe discovery data set will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please contact David Hinds (dhinds@23andme.com) for more information and to apply to access the data. Information for the 10,000 most significant SNPs from the discovery 23andMe GWAS is included in **Supplementary Table 12**.

Population and study design. Participants were part of the customer base of 23andMe, a consumer genetics company. This cohort has been described in detail elsewhere[27,28]. Participants provided informed consent and participated in the research online. The protocol was approved by an external AAHRPP-accredited institutional review board, Ethical and Independent Review Services. The discovery cohort was selected from participant data available in January 2015, and the replication cohort was selected in January 2016 from additional data available at that time.

Genotyping, quality control, and imputation. DNA extraction and genotyping were performed on saliva samples by the National Genetics Institute (NGI), a CLIA-licensed clinical laboratory and a subsidiary of the Laboratory Corporation of America. Samples were genotyped on one of four genotyping platforms. The V1 and V2 platforms are variants of the Illumina HumanHap550+ BeadChip, including about 25,000 custom SNPs selected by 23andMe. The V3 platform is based on the Illumina OmniExpress+ BeadChip, with custom content to improve overlap with theV2 array. The V4 platform used most recently is a fully custom array, including a lower-redundancy subset of V2 and V3 SNPs with additional coverage of lower-frequency coding variation. Platforms V1–V4 contain 586,916, 584,942, 1,008,948, and 570,000 SNPs, respectively. Samples that failed to reach a call rate of 98.5% were reanalyzed. Individuals whose analyses failed repeatedly were recontacted by 23andMe customer service to provide additional samples, as is done for all 23andMe customers.

Participant genotype data were imputed against the September 2013 release of the 1000 Genomes Project Phase 1 reference haplotypes, phased with SHAPEIT2 (ref. 29). We phased using an internally developed phasing tool, Finch, which implements the Beagle haplotype graph–based phasing algorithm[30], modified to separate the haplotype graph construction and phasing steps. Finch extends the Beagle model to accommodate genotyping error and recombination, to handle cases where there are no consistent paths through the haplotype graph for the individual being phased. We constructed haplotype graphs for European samples on each 23andMe genotyping platform from a representative sample of genotyped individuals and then performed out-of-sample phasing of all genotyped individuals against the appropriate graph.

In preparation for imputation, we split phased chromosomes into segments of no more than 10,000 genotyped SNPs, with overlaps of 200 SNPs. We excluded SNPs with Hardy–Weinberg equilibrium P value $<1 \times 10^{-20}$, call rate <95%, or large allele frequency discrepancies in comparison to European 1000 Genomes Project reference data. Frequency discrepancies were identified by computing a 2×2 table of allele counts for European 1000 Genomes Project samples and 2,000 randomly sampled 23andMe customers with European ancestry and then identifying SNPs with a χ^2 test P value $<1 \times 10^{-15}$. We imputed each phased segment against all-ancestry 1000 Genomes Project haplotypes (excluding monomorphic and singleton sites) using Minimac2 (ref. 31), with five rounds and 200 states for parameter estimation.

For the X chromosome, we built separate haplotype graphs for the non-pseudoautosomal region and each pseudoautosomal region; these regions were phased separately. We then imputed males and females together using Minimac2, as with the autosomes, treating males as homozygous pseudodiploids for the non-pseudoautosomal region.

Ancestry determination. We restricted the analysis to individuals who had >97% European ancestry, as determined through an analysis of local ancestry[32]. Briefly, our algorithm first partitions phased genomic data into short windows of about 100 SNPs. Within each window, we use a support vector machine (SVM) to classify individual haplotypes into one of 31 reference populations. The SVM classifications are then fed into a hidden Markov model (HMM) that accounts for switch errors and incorrect assignments, and gives probabilities for each reference population in each window. Finally, we used simulated admixed individuals to recalibrate the HMM probabilities so that the reported assignments were consistent with the simulated admixture proportions. The reference population data were derived from public data sets (the Human Genome Diversity Project, HapMap, and 1000 Genomes Project), as well as 23andMe customers who have reported having four grandparents from the same country.

A maximal set of unrelated individuals was chosen for each analysis using a segmental identity-by-descent (IBD) estimation algorithm[33]. Individuals were defined as related if they shared more than 700 cM identical by descent, including regions where the two individuals shared either one or both genomic segments identical by descent. This level of relatedness (sharing of roughly 20% of the genome) corresponds approximately to the minimal expected sharing between first cousins in an outbred population. When constructing the replication cohort, we identified unrelated individuals who were also unrelated to all individuals used in the discovery analysis.

We used principal-component analysis (PCA) to characterize residual population structure in the subset of 23andMe participants with European ancestry. We computed principal components using 82,654 SNPs that were genotyped on all 23andMe array designs, with Hardy–Weinberg P value $>1 \times 10^{-40}$, MAF >0.01, and call rate >99%, excluding regions of extended long-range LD. We used the ARPACK library[34] to compute principal components using data for 519,914 individuals across all array designs; additional individuals were then projected onto this set of eigenvectors.

The proportion of variance explained by each principal component is shown in **Supplementary Figure 2a** and the proportion of each component's variance that is explained by country of ancestry is shown in **Supplementary Figure 2b**, for a set of individuals reporting four grandparents from a single country. The first five principal components were largely explained by geographical ancestry, whereas higher-order principal components were not.

GWAS and meta-analysis. In the GWAS and replication analysis, we computed association test results by logistic regression assuming additive allelic effects. For tests using imputed data, we used the imputed dosages rather than best-guess genotypes. We included covariates for age, sex, and the top five principal components to account for residual population structure. Although we could justify the choice of five principal components on the basis of the preceding ancestry analysis, we actually chose to use five because of computational considerations, and others have noted this to be a reasonable choice[35].

For quality control of genotyped GWAS results, we removed SNPs that were only genotyped on our V1 and/or V2 platform because of small sample size and SNPs on the mitochondrial or Y chromosome because many of these are not genotyped reliably. Using family trio data, we flagged SNPs that failed a test for parent–offspring transmission; specifically, we regressed the child's allele count against the mean parental allele count and flagged SNPs with fitted $\beta<0.6$ and $P<1 \times 10^{-20}$ for a test of $\beta<1$. We removed SNPs with a Hardy–Weinberg P value $<1 \times 10^{-20}$ in Europeans or a call rate of <90%. We also tested genotyped SNPs for genotype date effects and removed SNPs with $P < 1 \times 10^{-50}$ by ANOVA of SNP genotypes against a factor dividing genotyping date into 20 roughly equally sized buckets.

For imputed GWAS results, we removed SNPs with average r^2 <0.5 or minimum r^2 <0.3 in any imputation batch, as well as SNPs that had strong evidence of an imputation batch effect. The batch effect test is an F test from an ANOVA of the SNP dosages against a factor representing imputation batch; we removed results with $P < 1 \times 10^{-50}$. Before GWAS analysis, we identified, for each SNP, the largest subset of the data passing these criteria, on the basis of the original genotyping platform—either V2 + V3 + V4, V3 + V4, V3, or V4 only—and computed association test results for whatever was the largest passing set. After quality control, the 23andMe discovery GWAS included results for 13,474,321 imputed variants and 60,949 genotyped variants that did not have imputed results passing our filters, for a total of 13,535,270 variants. Of these, 15,774 could not be used to compute test results because of problems with logistic regression fitting, leaving 13,519,496 tests. Hardy–Weinberg equilibrium and batch effect P values are presented in **Supplementary Table 13**.

Results from 23andMe were adjusted for variance inflation by multiplying the variance (that is, the square of the standard error) of each genetic effect estimate by the intercept of 1.0598, as calculated by LD score regression[20]. Meta-analysis with PGC data was conducted by inverse-variance-weighted,

fixed-effects meta-analysis on overlapping SNPs after adjusting the standard errors of each individual analysis for its own λ value (the LD score regression intercept in the PGC data was 1.0243). Final results from the meta-analysis were further adjusted for the overall LD score regression intercept of 1.0025.

LD score regression. We calculated LD scores (LD Score (LDSC) version 1.0.0) as previously described using the European 1000 Genomes Project reference panel (phase 3 version 5a) with a MAF cutoff for SNP inclusion greater than 5%. GWAS summary statistics data were collected from the following resources: PGC (MDD, bipolar disorder, SCZ1, SCZ1 + SWE, and SCZ2), the International Genomics of Alzheimer's Project (IGAP AD), the International Parkinson Disease Genomics Consortium (IPDGC PD), and the Global Lipids Genetics Consortium (GLGC LDL). GWAS data were harmonized using the munge_sumstats.py function (using the SNP list derived from LD score calculation), and genomic inflation control intercepts were calculated for the 23andMe MDD data, PGC MDD data, and PGC + 23andMe meta-analysis data using the ldsc.py function (with all default settings and options). Additionally, we calculated liability heritability estimates for the meta-analysis using the same function, with a population prevalence estimation of 15% or 25%, as previously described[13]. Finally, we calculated the cross-trait regression between the 23andMe MDD GWAS and PGC data sets, the IGAP data, the IPDGC data, and the GLGC data.

Trait ascertainment. Subjects with depression were identified through self-report in web-based surveys. A total of six survey data sources were used to compose the depression phenotype:

1. (Your Medical History survey: 2009–2013): "Have you ever been diagnosed by a doctor with any of the following psychiatric conditions?" (options for Depression: Yes, No, I don't know)
2. (Research Snippet: 2010–2014): "Have you ever been diagnosed with clinical depression?" (answers: Yes, No, I'm not sure)
3. (Health Intake survey, unbranched: 2014–2015): "Have you ever been diagnosed with or treated for any of the following conditions?" (options for Depression: Yes, No, I'm not sure)
4. (Health Intake survey, branched: 2013)
   - 4a. "Have you ever been diagnosed or treated for any of the following conditions?" (options for "A mental health or psychiatric condition": Yes, No, I'm not sure)
   - 4b. "What mental health problems have you had? Please check all that apply" (checkbox: Depression)
5. (Health Profile survey: 2015–2016)
   - 5a. "Have you ever been diagnosed with or treated for any of the following conditions? Anxiety, Attention deficit disorders, Bipolar disorder/manic depression, Depression, Eating disorder (such as anorexia or bulimia)" (answers: Yes, No, I'm not sure)
   - 5b. "Have you ever been diagnosed with or treated for depression?" (answers: Yes, No, I'm not sure)
6. (Health Followup survey: 2014–2015): "In the last 2 years, have you been newly diagnosed with or started treatment for any of the following conditions?" (options for Depression: Yes, No, I'm not sure)

Sources 1 and 3–5 represent four different iterations of a general medical history survey, administered over successive time periods from 2009 to 2016. Source 2 used a different mechanism for presenting individual questions to participants outside of the context of a formal survey. Source 6 was a survey administered to a subset of participants at least 1 year after they had completed one of the Health Intake surveys.

Sources 1–5 were combined by keeping the first non-missing response among these sources for each participant, evaluated in the specified order (the 'coalesced response'). We then incorporated responses to source 6, by defining cases as the union of cases from the coalesced response and cases from source 6; we defined controls as individuals who met the criteria for controls for either of the above conditions (1–5 or 6) and were not defined as cases for either condition (1–5 or 6).

For the branched data sources, sources 4 and 5, participants were first asked a screening question (4a or 5a) and, if they answered affirmatively, were asked a specific follow-up question (4b or 5b). Cases were defined by having positive responses to the follow-up question, and controls were the union of the individuals with "No" responses to either the screening or follow-up question.

As a result of the staging of the discovery and replication analyses, the discovery cohort did not include any responses from source 5 and the replication cohort consisted almost entirely of responses from source 3 or 5.

In sources 1 and 3–5, we also asked for an age of first diagnosis of depression. These data were provided by a majority of participants, including 74% of cases in the discovery cohort and 85% of cases in the replication cohort.

We used Cohen's κ to assess agreement across responses for sources 1–5, taking advantage of participants who had responded to more than one of the survey data sources (**Supplementary Table 14**).

Agreement was good in most comparisons ($\kappa > 0.7$) but was somewhat worse for comparisons with branched source 4 (κ between 0.5 and 0.7). Source 4 systematically undercalled cases in comparison to the other sources, apparently because of the wording of the screening question. This tendency was partially mitigated in the logic for the combined phenotype, where we preferentially used responses to sources 1–3 if available.

The logic for composing the depression phenotype in this way was based on several considerations. For most participants (>95%), we had just one response or the available responses were all in agreement, so a deeper analysis of the mismatch data was unlikely to substantially affect downstream results. Our strategy of selecting one response per participant without regard for the other responses of this individual also seemed least likely to introduce bias in classification of participants who provided multiple responses.

Secondary phenotypes. A set of common comorbidities of depression were defined on the basis of responses to single questions, as follows:

- Anxiety (Health Intake survey, unbranched, 2014–2015): "Have you ever been diagnosed with or treated for anxiety?" (options: Yes, No, I don't know)
- Panic attacks (Health Intake survey, unbranched, 2014–2015): "Have you ever been diagnosed with or treated for panic attacks?" (options: Yes, No, I don't know)
- Insomnia (Research Snippet, 2013–2016): "Do you routinely have trouble getting to sleep at night?" (options: Yes, No, I don't know)
- Taking an SSRI (Research Snippet, 2013–2016): "Are you currently taking an SSRI (selective serotonin reuptake inhibitor) for any reason?" (options: Yes, No, I don't know)
- Ever taken medication for a mental health condition, prescription sleep aids, or prescription pain medication (Health Intake survey, unbranched, 2014–2015): "Have you ever taken these medications?" "Medications to treat depression or anxiety or another mental health condition," "Prescription sleep aids," "Prescription pain medications" (checkbox for each category)

Overweight and obesity were defined on the basis of BMI (>27 and >30, respectively), computed from self-reported height and weight, which were collected using fill-in forms in multiple survey contexts.

Associations with secondary phenotypes and age of onset. We computed GRSs on the basis of the 17 SNPs with P values $<5 \times 10^{-8}$ in the joint analysis of the 23andMe discovery, PGC, and 23andMe replication results, as a linear combination of independent single-SNP effect sizes estimated from that joint analysis (**Supplementary Table 2**). We tested each secondary phenotype for association with these scores in the combined 23andMe discovery and replication cohort; we tested for effects on age of onset in depression cases only (**Table 3**). For age of onset, we defined 'early onset' as onset before 30 years of age and fit this binary outcome by logistic regression; we also fit a model for continuous age of onset using linear regression. In all these tests, we included covariates for age, sex, five principal components, and depression case/control status. In this way, we were testing for residual association not explained by depression status, and thus these associations are independent of the data that were used to identify these 17 variants. Separately, we tested each of the 17 SNPs individually for association with this same set of phenotypes, including the same covariates (**Supplementary Table 7**).

doi:10.1038/ng.3623

DEPICT functional analysis. We used DEPICT[18] to determine the most likely causal gene at each of the depression-associated loci and to assess reconstituted gene sets enriched for and tissues highly expressing these genes. The reconstituted gene sets used in the analysis were derived from publicly available gene set annotations, which were then integrated with data from 77,840 gene expression arrays[36], to predict which other genes were likely to be part of these gene sets.

For the analysis, we selected SNPs significantly associated with depression at $P < 1 \times 10^{-5}$. After clumping these SNPs using 500-kb flanking regions and an LD cutoff of $r^2 > 0.1$, 63 independent SNP signals were identified from 816 top variants. These 63 top SNPs were further merged into 59 nonoverlapping loci containing 157 genes, which were then assessed using the DEPICT algorithm for gene set and tissue enrichment[18]. The results shown in **Supplementary Table 3** were not corrected for multiple testing.

27. Eriksson, N. *et al.* Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet.* **6**, e1000993 (2010).
28. Tung, J.Y. *et al.* Efficient replication of over 180 genetic associations with self-reported medical data. *PLoS One* **6**, e23473 (2011).
29. 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
30. Browning, S.R. & Browning, B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* **81**, 1084–1097 (2007).
31. Fuchsberger, C., Abecasis, G.R. & Hinds, D.A. minimac2: faster genotype imputation. *Bioinformatics* **31**, 782–784 (2015).
32. Durand, E.Y., Do, C.B., Mountain, J.L. & Macpherson, J.M. Ancestry Composition: a novel, efficient pipeline for ancestry deconvolution. Preprint at *bioRxiv* http://dx.doi.org/10.1101/010512 (2014).
33. Henn, B.M. *et al.* Cryptic distant relatives are common in both isolated and cosmopolitan genetic samples. *PLoS One* **7**, e34267 (2012).
34. Lehoucq, R.B., Sorensen, D.C. & Yang, C. *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods* (SIAM, 1998).
35. Tucker, G., Price, A.L. & Berger, B. Improving the power of GWAS and avoiding confounding from population stratification with PC-Select. *Genetics* **197**, 1045–1049 (2014).
36. Fehrmann, R.S. *et al.* Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat. Genet.* **47**, 115–125 (2015).