Analysis of 23andMe antidepressant efficacy survey data: implication of circadian rhythm and neuroplasticity in bupropion response

QS Li1, C Tian2, GR Seabrook3, WC Drevets1 and VA Narayan1

INTRODUCTION

Despite the Psychiatric Genomic Consortium (PGC)'s collaborative efforts of major depressive disorder (MDD) meta-analysis, variants predictive of disease susceptibility remain elusive, partly due to the heterogeneous nature of the disorder. Treatment for MDD and genetic predictor of treatment response are equally challenging. There are ~30 antidepressants available for MDD patient care and response to treatment varies in time to onset of benefit, overall efficacy, and duration of effect. Genetic variability may contribute to the differences in drug-specific, class-specific or antidepressant-wide treatment resistance. Clinical studies with the genetic data are often limited in sample sizes. Drug response obtained from self-reports may offer an alternative approach to conduct a study with much larger sample size. Using the phenotype data collected from 23andMe ‘Antidepressant Efficacy and Side Effects’ survey and genotype data from 23andMe’s research participants, we conducted genome-wide association study (GWAS) on subjects of European ancestry using four groups of phenotypes (a) non-treatment-resistant depression ($n = 7795$) vs treatment-resistant depression (TRD, $n = 1311$), (b) selective serotonin reuptake inhibitors (SSRI) responders ($n = 6348$) vs non-responders ($n = 3340$), (c) citalopram/escitalopram responders ($n = 2963$) vs non-responders ($n = 2005$), and (d) norepinephrine–dopamine reuptake inhibitor (NDRI, bupropion) responders ($n = 2675$) vs non-responders ($n = 1861$). Each of these subgroups was also compared with controls ($n = 190 000$). The most significant association was from bupropion responders vs non-responders analysis. Variant rs1908557 ($P = 2.6 \times 10^{-8}$, OR = 1.35) passed the conventional genome-wide significance threshold ($P = 5 \times 10^{-8}$) and was located within the intron of human spliced expressed sequence tags in chromosome 4. Gene sets associated with long-term depression, circadian rhythm and vascular endothelial growth factor (VEGF) pathway were enriched in the bupropion analysis. No single-nucleotide polymorphism passed genome-wide significance threshold in other analyses. The heritability estimates for each response group compared with controls were between 0.15 and 0.25, consistent with the known heritability for major depressive disorder.

The largest antidepressant efficacy GWAS meta-analysis performed using the data from three studies among subjects of European ancestry only (STAR*D $n = 980$, GENDEP $n = 706$, and Munich Antidepressant Response Signature including additional samples genotyped $n = 604$ resulting a total sample size of 2256 across three studies) did not identify any variants passing genome-wide significance threshold with primary outcome measurements (percentage improvement on the clinician-rated depression scale and remission in each study after 12 weeks of treatment), although a secondary analysis restricted to 1354 individuals treated with either citalopram or escitalopram revealed a variant rs12054895 ($P = 2.65 \times 10^{-8}$) in the intergenic region of chromosome 5 associated with early improvement after 2 weeks of antidepressant treatment.5 Drug response information obtained from self-reported questionnaires may offer an alternative approach to conduct a study with much larger sample size.

MDD is a heterogeneous disease,6 which partially explained why finding disease risk variants for MDD seems to be more challenging than for schizophrenia and bipolar. Disease subtypes have been defined by clinical features such as MDD with the melancholic, atypical or anxious distress specifier.7 Theoretically, different disease subtypes may be defined by different underlying molecular mechanism, which in turn determines the response to different therapeutic agents. In this study, we explored different disease subtypes as defined by response status and compared the

1Neuroscience Therapeutic Area, Janssen Research & Development, LLC, Titusville, NJ, USA; 23andMe, Inc., Mountain View, CA, USA and 3Janssen Research & Development, Menlo Park, CA, USA. Correspondence: Dr QS Li, Neuroscience Therapeutic Area, Janssen Research & Development, LLC, 1125 Trenton-Harbourton Road, Titusville, NJ 08560, USA. E-mail: ql2@its.jnj.com

Received 26 January 2016; revised 2 July 2016; accepted 18 July 2016
heritability and genetic correlation to reference populations (PGC1 MDD, bipolar and schizophrenia).

MATERIALS AND METHODS

Cohort

Approximately 48,000 research participants, drawn from 23andMe (Mountain View, CA, USA), provided saliva samples for genetic testing, informed consent and answered surveys (‘Antidepressant Efficacy and Side Effects’ and ‘Your Profile and Health History’) online between June 2013 and June 2015 according to a human subjects protocol reviewed and approved by Ethical & Independent Review Services, an AARHPP-accredited Institutional Review Board (http://www.analdireview.com). Studies on other selected phenotypes from 23andMe were previously reported.8–14

Sample genotyping and SNP data imputation

DNA extraction and genotyping were performed as described before,9 on saliva samples by National Genetics Institute (NGI, Los Angeles, CA, USA). Samples have been genotyped on one of four genotyping platforms. The V1 and V2 platforms were variants of the Illumina Human660W-Quad BeadChip (Illumina, San Diego, CA, USA), including about 25,000 custom single-nucleotide polymorphisms (SNPs) selected by 23andMe, with a total of about 560,000 SNPs. The V3 platform was based on the Illumina OmniExpress+ BeadChip, with custom content to improve the overlap with the V2 array, with a total of about 950,000 SNPs. The V4 platform in current use is a fully custom array, including a lower redundancy subset of V2 and V3 SNPs, with additional coverage of lower-frequency-variant variation, and about 570,000 SNPs. Samples that failed to reach 98.5% call rate were reanalyzed. Participant genotype data were imputed against the September 2013 release of 1000 Genomes15 Phase I reference haplotypes, phased with Shapeit2. Prior to imputation, we excluded SNPs with Hardy–Weinberg equilibrium P < 10−20, call rate < 95%, or with large allele frequency discrepancies compared to European 1000 Genomes reference populations.9

Sample sizes for each phenotypic group together with demographic variables such as gender and age as well as genotyping platform used and GWAS genomic control inflation factor lambda are listed in Supplementary Table S1. Overall, the prevalence rate

Genetic heritability and genetic correlation estimates

PGC phase 1 disease susceptibility summary association statistics for MDD, bipolar, and schizophrenia16–19 were downloaded from PGC website (http://www.med.unc.edu/pgc/downloads) and included together with the summary statistics from this study as reference data sets for genetic heritability and genetic correlation estimates. Phenotypic variance explained by variants (both genotyped and imputed, mostly SNPs) (HR) for each of the phenotype groups and the genetic correlation between traits (rG) were estimated using association statistics as implemented in LD Score regression.20

Gene set enrichment analysis

INRICH is a pathway analysis tool for genome-wide association studies, designed for detecting enriched association signals of linkage disequilibrium (LD)-independent genomic regions within biologically relevant gene sets.23 Reference gene sets used in the INRICH analysis include KEGG, Gene Ontology and Molecular Signature Database (v5.0). Top variants from responder vs non-responder analyses with nominal association P-value < 0.0005, 0.0001, 0.00005, 0.00001 were separately fed into PLINK to clump the variants into LD-independent genomic intervals (r2 threshold using 0.2, 0.3 and 0.5, respectively), then LD-independent genomic regions were used for INRICH (version 1.0) analyses. No multiple testing corrections were applied for running INRICH against multiple reference gene sets or for using multiple parameters (P-value cutoff and LD threshold).

RESULTS

Sample sizes for each phenotypic group together with demographic variables such as gender and age as well as genotyping platform used and GWAS genomic control inflation factor lambda are listed in Supplementary Table S1.
for depressed patients was greater in females than in males, consistent with the extant epidemiological literature. Lambda scores ($\lambda_{1000}$) between 1.002 and 1.013 revealed no departures from uniform distributions of $P$-values across ~12 million genotyped and imputed markers.

Out of a total of 12 genome-wide association analyses performed, only the bupropion responders vs non-responders analysis yielded a locus reaching genome-wide significance threshold ($P = 5 \times 10^{-8}$; Figure 1a). The most significant association variant rs1908557 ($P = 2.6 \times 10^{-8}$; OR =1.35) was located...
within the intron of human spliced expressed sequence tags known between tags known among genes GPRIN family member 3 (GPRIN3) and synuclein alpha (SNCA) in chromosome 4 (Figure 1b) overlapping with enhancer/promoter-associated histone mark H3K4Me1. Each copy of rs1908557-C allele was associated with higher odds of being bupropion non-responder. The frequency of C allele was relatively common (minor allele frequency = 25%) in the study population. No SNP passed genome-wide significance threshold in any other GWAS analyses. Manhattan and quantile–quantile plots for all analyses (Supplementary Figure S1 and S2) as well as variants with a nominal association P-value in any of the 12 analyses < 5 x 10^-5 (Supplementary Dataset 1) are available online. The indexed SNPs with P < 1 x 10^-5 for the responders vs non-responders analyses and the responder subgroup vs healthy control group analyses are listed in Tables 1 and 2, respectively.

The heritability estimates for responders vs non-responders or non-TRD vs TRD analyses were generally unreliable with large standard errors and smaller h^2 estimates (Table 3). The heritability estimates for each responder, non-responder, non-TRD, TRD group in comparison to healthy controls were between 0.14 and 0.22, consistent with the known heritability estimates for MDD (h^2 = 0.19) estimated from PGC1 MDD samples. The genetic correlation between responder vs non-responder and non-TRD vs TRD in general are less similar than genetic correlations among responder/non-responder/non-TRD/TRD vs healthy controls, as expected.

Gene set enrichment analysis may yield signals of enriched gene sets in GWAS analysis despite the individual variants not reaching genome-wide significance. Applying INRICH enrichment analysis to the NDRI (bupropion) responders vs non-responders GWAS results implicated gene sets such as circadian rhythm, long-term depression (LTD), and vascular endothelial growth factor (VEGF) pathway being enriched among the bupropion suggestive association hits (P < 0.005); Table 4, Supplementary Table S2 for gene lists belonging to each enriched gene sets in Table 4, Supplementary Dataset 2 for SNP lists corresponding to gene lists from Supplementary Table S2). Additional enriched gene sets for other GWAS analyses are listed in Supplementary Table S3.

**DISCUSSION**

We identified a candidate genetic marker rs1908557 in the intergenic region between GPRIN3 and SNCA for bupropion response with an association P-value passing genome-wide significance using a phenotype derived from survey data. This finding ultimately will require replication in clinically ascertained samples to further dissect the genetic basis of treatment response to bupropion among depression patients. The biological significance of rs1908557 is unknown except that rs1908557 is also marginally associated with two brain regions known to exhibit volumetric difference between MDD subjects and healthy controls. Specifically, patients carrying the common T allele are associated with lower hippocampal (P = 0.016), amygdala (P = 0.067) and nucleus accumbens volumes (P = 0.067). In the recent ENIGMA analysis of 1728 MDD patients and 7199 controls, MDD (especially recurrent MDD and patients with earlier age of onset (AOO ≤ 21 years)) had significantly lower hippocampal volumes, although patients with earlier AOO also showed a trend toward smaller amygdala volumes. In our analysis the T allele was associated specifically with better bupropion response. It remains unclear whether the intergenic variant has any biological impact on the neighboring gene SNCA. So far the limited sample size from GTEx did not
support any eQTL relationship between rs1908557 and the neighboring genes; however, the overlap of enhancer/promoter-associated histone mark H3K4Me1 cannot rule out that there may be still more subtle relationship between neighbor genes and rs1908557. However, eQTL data from BRAINEAC\textsuperscript{25} (http://braineac.org/) suggested a subtle cis-eQTL relationship between rs1908557 and genes further away from rs1908557 such as family with sequence similarity 13 member A (FAM12A), multimerin 1 (MMRN1), HECT and RLD domain containing E3 ubiquitin protein ligase 5 (HERC5), and coiled-coil serine rich protein 1 (COCS1).
### Table 3. Heritability and genetic correlation estimates

| Analysis group | $h^2$ liability (s.e.) | Genetic correlation (s.e) |
|----------------|------------------------|--------------------------|
|                | C                      | D                        | E                        | F                      | G                        | H                        | I                        | J                        | K                        | L                        | M                        | N                        | O                        |
| NDRI responders vs non-responders (A) | -0.0541 (0.088) | NE | NE | NE | NE | NE | NE | NE | NE | NE | NE | NE | NE | NE |
| SSRI responders vs non-responders (B) | 0.0456 (0.0498) | 0.4414 (0.4605) | -0.3919 (0.4126) | NE | NE | NE | NE | NE | NE | NE | NE | NE | NE | NE |
| citalopram or escitalopram responders vs non-responders (C) | 0.1201 (0.0925) | 0.3214 (0.3902) | -0.1525 (0.2966) | 0.0032 (0.3484) | NE | NE | NE | NE | NE | NE | NE | NE | NE | NE |
| non-responders (C) | 0.11 (0.0995) | -0.4917 (0.3368) | -0.334 (0.2569) | -0.018 (0.3735) | -0.6184 (0.1377)** | -0.3437 (0.3974)** | 0.5017 (0.2409)** | 0.2372 (0.2739)** | NE | NE | NE | NE | NE | NE |
| NDRI non-responders vs healthy controls (E) | 0.1492 (0.0617) | 1.0662 (0.2294)*** | 1.021 (0.2761)*** | 1.0813 (0.2233)*** | NE | NE | 1.0858 (0.2129)*** | 0.8947 (0.2421)*** | NE | NE | NE | NE | NE | NE |
| SSRI non-responders vs healthy controls (F) | 0.1781 (0.0323) | 0.9621 (0.0906)*** | 1.0064 (0.1315)*** | 0.9651 (0.1755)*** | 0.9379 (0.1372)*** | 0.9 (0.1518)*** | 0.9622 (0.1236)*** | 1.0192 (0.1987)*** | NE | NE | NE | NE | NE | NE |
| citalopram or escitalopram responders vs healthy controls (G) | 0.1946 (0.0639) | 0.7342 (0.1935)*** | 1.0803 (0.2411)*** | 0.9699 (0.2264)*** | 0.8184 (0.2201)*** | 0.9351 (0.2183)*** | 1.021 (0.3517)*** | 0.1689 (0.2012)*** | NE | NE | NE | NE | NE | NE |
| TRD vs healthy controls (H) | 0.1673 (0.0537) | 1.0835 (0.1958)*** | 0.8264 (0.2083)*** | 0.6599 (0.2267)*** | 0.8811 (0.2046)*** | 1.1634 (0.2055)*** | 0.8111 (0.1904)*** | 0.091 (0.2576)*** | NE | NE | NE | NE | NE | NE |
| NDRI responders vs healthy controls (I) | 0.2211 (0.0607) | 0.9303 (0.1136)*** | 0.9131 (0.1791)*** | 1.0013 (0.1117)*** | 0.778 (0.2128)*** | 0.1861 (0.1383)*** | 0.091 (0.1111)*** | 0.0678 (0.1119)*** | NE | NE | NE | NE | NE | NE |
| SSRI responders vs healthy controls (J) | 0.1441 (0.0201) | 0.9667 (0.0594)*** | 0.9699 (0.0232)*** | 0.9899 (0.1393)*** | 0.697 (0.0883)*** | 0.0926 (0.0791)*** | 0.1295 (0.1024)*** | 0.0547 (0.0605)*** | NE | NE | NE | NE | NE | NE |
| citalopram or escitalopram responders vs healthy controls (K) | 0.1621 (0.0354) | 0.9871 (0.0718)*** | 0.95962 (0.1455)*** | 0.091 (0.0791)*** | 0.0101 (0.0900)*** | 0.0897 (0.0100)*** | 0.1063 (0.0104)*** | 0.0529 (0.0605)*** | NE | NE | NE | NE | NE | NE |
| non-TRD vs healthy controls (L) | 0.141 (0.0151) | 0.9871 (0.0718)*** | 0.95962 (0.1455)*** | 0.091 (0.0791)*** | 0.0101 (0.0900)*** | 0.0897 (0.0100)*** | 0.1063 (0.0104)*** | 0.0529 (0.0605)*** | NE | NE | NE | NE | NE | NE |
| PGC MDD (M) | 0.1894 (0.0354) | 0.5551 (0.1241)*** | 0.472 (0.0833)*** | 0.0656 (0.0605)*** | 0.472 (0.0833)*** | 0.0656 (0.0605)*** | 0.472 (0.0833)*** | 0.0656 (0.0605)*** | NE | NE | NE | NE | NE | NE |
| PGC BP (N) | 0.2848 (0.0286) | 0.6504 (0.1241)*** | 0.0529 (0.0833)*** | 0.1063 (0.0605)*** | 0.0529 (0.0833)*** | 0.1063 (0.0605)*** | 0.0529 (0.0833)*** | 0.1063 (0.0605)*** | NE | NE | NE | NE | NE | NE |
| PGC SCZ (O) | 0.3227 (0.0248) | 0.6504 (0.1241)*** | 0.0529 (0.0833)*** | 0.1063 (0.0605)*** | 0.0529 (0.0833)*** | 0.1063 (0.0605)*** | 0.0529 (0.0833)*** | 0.1063 (0.0605)*** | NE | NE | NE | NE | NE | NE |

Abbreviations: MDD, major depressive disorder; NDRI, norepinephrine–dopamine reuptake inhibitor; NE, not estimable; SSRI, selective serotonin reuptake inhibitor; TRD, treatment-resistant depression. *P < 0.05; **P < 0.01; ***P < 0.001. $h^2$ estimate assumes population prevalence rate of 8% for responder/non-responders/TRD/non-TRD groups, 15% for MDD, and 1% for bipolar and schizophrenia, respectively.
rs8076666 in solute carrier family 26 (anion exchanger), member 11 (SLC26A11, \( P = 1.59 \times 10^{-5} \)) and rs9373491 in glutamate receptor, metabotropic 1 (GRM1, \( P = 3.01 \times 10^{-5} \)) also showed suggestive volumetric relationships with brain volumes\(^{23}\) (Supplementary Dataset 1).

It is noteworthy that the gene set enrichment analysis of bupropion GWAS yielded the most gene sets of interest to MDD. In particular, circadian rhythm, LTD and VEGF pathway genes were enriched among the bupropion suggestive association hits (\( P < 0.0005 \)). Alteration of circadian rhythms and disturbances of sleep are common features of the major depressive syndrome. Variants from period circadian clock 3 (PER3, representative intronic variants rs7528457, \( P = 3.81 \times 10^{-5} \); rs12137927 \( P = 6.30 \times 10^{-5} \), RAR-related orphan receptor A (ROA, rs185937898, \( P = 5.51 \times 10^{-5} \)), and nuclear receptor subfamily 1, group D, member 1 (NR1D1) were accountable for circadian rhythm gene set enrichment in the bupropion response GWAS. Acute and chronic stress, which putatively can be precipitating factors of MDD, also can affect rhythms of the circadian pacemaker. PER3 is a member of the Period family of genes and is expressed in a circadian pattern in the suprachiasmatic nucleus, the primary circadian pacemaker in the mammalian brain. Genes in this family encode components of the circadian rhythms of locomotor activity, metabolism and behavior. PER3 is upregulated by CLOCK/ARNTL heterodimers but then represses this up-regulation in a feedback loop using PER/CRY heterodimers to interact with CLOCK/ARNTL. RORA is a member of the NR1 subfamily of nuclear hormone receptors and aids in the transcriptional regulation of some genes involved in circadian rhythm. Rs228697 in PER3 has been linked to morningness–eveningness preference and circadian rhythm sleep disorders, although the length polymorphism/VNTR in PER3 has been linked to stress response and bipolar disorder.\(^{26-28}\) In addition, rs12137927 (the same SNP implicated in our study) and rs228644 from PER3 and rs11632098 from RORA were reportedly linked to endorsing the presence of both a modest number (\( > 2 \) to \( < 6 \)) and a high number of depressive symptoms (\( \geq 6 \)) on the Geriatric Depression Scale as compared with endorsing none-few depressive symptoms (0–2).\(^{29}\) The variant with association signal of \( P < 0.0001 \) in the NR1D1 interval is variant rs10305315, located downstream of IGFBP4 (\( P = 5.99 \times 10^{-5} \)). It remains unclear whether rs10305315 affects the function of NR1D1 (no eQTL evidence from GTEx portal based on GTEx Analysis Release V6). NR1D1 encodes a ligand-sensitive transcription factor that negatively regulates the expression of core clock proteins. The expression of this gene represses the circadian clock transcription factor aryl hydrocarbon receptor nuclear translocator-like protein 1 (ARNTL).

Variants from glutamate receptor, ionotropic, delta 2 (GRID2, rs76800659 \( P = 4.21 \times 10^{-5} \)), glutamate receptor, metabotropic 1 (GRM1, rs2328741 \( P = 1.12 \times 10^{-5} \); rs2268666 \( P = 4.75 \times 10^{-5} \), glutamate receptor, metabotropic 5 (GRMS, rs308873 \( P = 2.24 \times 10^{-5} \)), corticotropin releasing hormone receptor 1 (CHHR1) and protein kinase C, alpha (PRKCA, rs34337960 \( P = 1.49 \times 10^{-5} \)) were accountable for LTD gene set enrichment. \( \eta \)-glutamate is the major excitatory neurotransmitter in the central nervous system and activates both ionotropic and metabotropic glutamate receptors. Glutamatergic neurotransmission is involved in most aspects of normal brain function and putatively is perturbed in many neuropathological conditions. In particular, glutamate has been implicated in the pathophysiology and treatment of mood disorders, particularly with respect to altered transmission in limbic–thalamocortical circuits.\(^{30}\) Rs2268666 in GRM1 was shown to be associated with unipolar depression (UPD) phenotype in a discovery cohort of 350 patients and 370 matched controls (\( P = 7.0 \times 10^{-5} \) in allelic test and \( P = 0.0002 \) in genotypic test with T being the risk allele, both passing multiple testing correction threshold), and was partially replicated in an
independent cohort of 904 patients and 1012 controls (genotypic test \( P = 0.02 \), allelic test \( P = 0.59 \)). Furthermore, patients homozygous for the non-risk genotypes (C/C) showed reduced hippocampal glutamate levels as measured by \(^{1}H\)-MR-spectroscopy, a more pronounced normalization of HPA-axis hyperactivity using a combined dexamethasone suppression/CRH-challenge (DEX/CRH) test,\(^{32}\) and a better antidepressant treatment outcome.\(^{33}\) Likewise, patients with each copy of C allele of rs2268666 also showed an increased likelihood of being bupropion responders in this study (\( P = 4.75 \times 10^{-9} \), OR = 1.22). Knockout mice mGluR5(-/-) displayed more depression-like behaviors (learned helplessness, social withdrawal and anhedonia) than control mice following exposure to stressful stimuli, whereas lentiviral rescue of mGluR5 in the nucleus accumbens decreased these depression-like behaviors in mGluR5(-/-) mice.\(^{34}\) mGluR5 may be involved in the regulation of neural network activity and synaptic plasticity. \(^{d}\) Serine is an endogenous co-agonist for N-methyl-d-aspartate receptors and regulates neurotransmission and synaptic plasticity including long-term potentiation and LTD.\(^{43,35}\) \(^d\) Serine serves as an endogenous ligand for GluD2, the predominant excitatory neurotransmitter receptors in the mammalian brain, to regulate LTD at synapses.

The VEGF pathway is another gene set enriched in bupropion response analysis. The neurotransmitter hypothesis of MDD hypothesizes that the neuropathology of MDD involves a downregulation of neurotrophin signaling, involving both brain-derived neurotrophic factor and the multi-competent angiogenic and neurogenic molecule VEGF in hippocampal neurogenesis. A recent meta-analysis of fourteen studies (\( n = 1633 \)) showed that VEGF levels were significantly elevated in individuals with MDD when compared to healthy controls,\(^{36}\) supporting that this mediator may be involved in neuroplasticity mechanisms or compensating for the pathophysiology of MDD. Variants from fms-related tyrosine kinase 4 (FLT4, rs189869480, \( P = 6.12 \times 10^{-7} \)), PPKCA, and phospholipase C, gamma 1 (PLCG1, rs56012336, \( P = 2.05 \times 10^{-10} \)) was accountable for VEGF pathway enrichment in the bupropion response GWAS. The protein encoded by PLCG1 catalyzes the formation of inositol 1, 4, 5-trisphosphate and diacylglycerol from phosphatidylinositol 4, 5-bisphosphate, and is a major enzyme of the phosphatidylinositol second messenger system. Polymorphisms in PLCG1A were associated with bipolar disorder\(^{37}\) and response to lithium.\(^{18,39}\)

Many variants from candidate gene studies and GWAS analyses have been associated with antidepressant treatment response with various strength of association in mostly small sample sizes. Findings of selected variants reported from prior candidate gene and GWAS studies are discussed in Supplementary Text S3.

The consistent heritability estimates between epidemiologically ascertained depression samples and PGC1 MDD samples (which in large part were clinically ascertained) and the high genetic correlation estimated between our phenotypes and PGC1 MDD patients suggest that self-reported samples do not significantly bias the recruitment of diagnostic class, although the self-reported samples are certainly more heterogeneous and may include both MDD and minor depressive disorder samples. The apparent overlap in genetic architecture between drug/class-specific and antidepressant-wide responders/non-responders is not surprising given that these research participants have a history of self-reported depression. There are inherent challenges of interpreting single arm retrospective study where self-reported outcome assessment is the only data collection modality. These challenges include lack of diagnostic certainty, recall biases and qualitative nature of outcome assessment, and whether patients are optimally dosed and medication compliance for minimal dose exposure (for example, 6 weeks). The Antidepressant Efficacy and Side Effect Survey contained a question ‘For how long did you take Wellbutrin/bupropion?’ with the following four answers ‘Less than 4 months; 4 months–1 year; 1–5 years; and Over 5 years’. The shortest time duration < 4 months will not allow us to disen-<ref>talgle minimal medication exposure of 6 weeks. However, among the ~ 2500 patients reporting efficacious bupropion response, only ~ 150 (~6%) took the medication for < 4 months. In contrast, among the ~ 1800 patients who reported lack of efficacy to bupropion, ~ 830 (~46%) took the medication for longer than 4 months.

This study grouped escitalopram and citalopram into one analysis group to increase the sample size. Citalopram is a racemic mixture of a pharmacologically active S-enantiomer (that is, escitalopram) and the R-enantiomer, which is putatively inactive. The study explored both drug-specific and class-specific effects and focused on the groups with the largest sample sizes as well as on the treatment-resistant subgroup. Despite the interesting finding from the bupropion GWAS analysis and the consistency (though not significant at the genome-wide significance threshold) between this study and previously reported studies for key MDD, pharmacokineticians, and pharmacodynamics genes relevant to antidepressant efficacy response, the genetic variants alone are unlikely to deliver clinically actionable predictive diagnostic tests. A more comprehensive approach using a composite signature of predictors ultimately may be required to predict treatment outcome to a particular drug class with sufficient sensitivity and specificity to warrant its use in the clinic.

**CONFLICT OF INTEREST**

QSL, GRS, WCD and VAN are employees of Janssen Research & Development, LLC. All authors may be shareholders in Johnson & Johnson, which is the parent company of the Janssen companies. QSL, GRS, WCD and VAN declare that except for income received from their primary employer, no financial support or compensation has been received from any individual or corporate entity over the past 3 years for research or professional service, and there are no personal financial holdings that could be perceived as constituting a potential conflict of interest. CT is an employee of 23andMe.

**ACKNOWLEDGMENTS**

We thank the customers, research participants and employees of 23andMe for making this work possible. The genotype-phenotype association analysis was funded by Janssen Research and Development, LLC.

**REFERENCES**

1. Ising M, Lucase S, Binder EB, Bettecken T, Uhr M, Ripke S et al. A genomewide association study points to multiple loci that predict antidepressant drug treatment outcome in depression. Arch Gen Psychiatry 2009; 66: 966–975.
2. Uher R, Perroud N, Ng MY, Hauser J, Henigsberg N, Maier W et al. Genome-wide pharmacogenetics of antidepressant response in the GENDEP project. Am J Psychiatry 2010; 167: 555–564.
3. Garriock HA, Kraft JB, Shyn SJ, Peters EJ, Yokoyama JS, Jenkins GD et al. A genomewide association study of citalopram response in major depressive disorder. Biol Psychiatry 2010; 67: 133–138.
4. Ji Y, Biernacka JM, Hebbring S, Chai Y, Jenkins GD, Batzler A et al. Pharmacogenomics of selective serotonin reuptake inhibitor treatment for major depressive disorder: genome-wide associations and functional genomics. Pharmacogenomics J 2013; 13: 456–463.
5. GENDEP Investigators, MARS Investigators, STAR*D Investigators. Common genetic variation and antidepressant efficacy in major depressive disorder: a meta-analysis of three genome-wide pharmacogenetic studies. Am J Psychiatry 2013; 170: 207–217.
6. Krishnan V, Nestler EJ. The molecular neurobiology of depression. Nature 2008; 455: 894–902.
7. American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders. 5th edn Washington DC, USA, 2013.
8. Do CB, Tung JY, Dorfman E, Kiefer AK, Drabant EM, Francke U et al. Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson’s disease. PLoS Genet 2011; 7: e1002141.
9. Eriksson N, Macpherson JM, Tung JY, Hon LS, Naughton B, Saxonov S et al. Web-based, participant-driven studies yield novel genetic associations for common traits. PLoS Genet 2010; 6: e1000993.
10 Eriksson N, Tung JY, Kiefer AK, Hinds DA, Francke U, Mountain JL et al. Novel associations for hypothyroidism include known autoimmune risk loci. PLoS ONE 2012; 7: e34477.
11 Kiefer AK, Tung JY, Do CB, Hinds DA, Mountain JL, Francke U et al. Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. PLoS Genet 2013; 9: e1003299.
12 Tung JY, Do CB, Hinds DA, Kiefer AK, Macpherson JM, Chourdhy AB et al. Efficient replication of over 180 genetic associations with self-reported medical data. PLoS ONE 2011; 6: e23473.
13 Hromatka BS, Tung JY, Kiefer AK, Do CB, Hinds DA, Eriksson N. Genetic variants associated with motion sickness point to roles for inner ear development, neurological processes and glucose homeostasis. Hum Mol Genet 2014; 23: 2700–2708.
14 Hu Y, Shmygelska A, Tran D, Eriksson N, Tung JY, Hinds DA. GWAS of 89,283 individuals identifies genetic variants associated with self-reporting of being a morning person. Nat Commun 2016; 7: 10448.
15 Genomes Project C, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM et al. A map of human genome variation from population-scale sequencing. Nature 2010; 467: 1061–1073.
16 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 2013; 381: 1371–1379.
17 Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM et al. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry 2013; 18: 497–511.
18 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 2011; 43: 969–976.
19 Schizophrenia Psychiatric Genome-Wide Association Study Consortium. Genome-wide association study identifies five new schizophrenia loci. Nat Genet 2011; 43: 969–976.
20 Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J et al. Schizophrenia Working Group of the Psychiatric Genomics Consortium. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015; 47: 291–295.
21 Lee PH, O'Dushlaine C, Thomas B, Purcell SM. INRICH: interval-based enrichment analysis for genome-wide association studies. Bioinformatics 2012; 28: 1797–1799.
22 Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 2010; 26: 2336–2337.
23 Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivieres S, Jahanshad N et al. Common genetic variants influence human subcortical brain structures. Nature 2015; 520: 224–229.
24 Schmaal L, Velman DJ, van Erp TG, Samann PG, Frodl T, Jahanshad N et al. Subcortical brain alterations in major depressive disorder: findings from the ENIGMA Major Depressive Disorder working group. Mol Psychiatry 2015; 21: 806–812.
25 Ramasamy A, Trabzuni D, Gueff S, Varghese V, Smith C, Walker R et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. Nat Neurosci 2014; 17: 1418–1428.
26 Hida A, Kitamura S, Katayose Y, Kato M, Ono H, Kadotani H et al. Sequencing of clock gene polymorphisms demonstrates association of a PER3 polymorphism with morningness–eveningness preference and circadian rhythm sleep disorder. Sci Rep 2014; 4: 6309.
27 Wang X, Mozhu K, Li Z, Mulligan MK, Ingels JF, Zhou X et al. A promoter polymorphism in the Per3 gene is associated with alcohol and stress response. Transl Psychiatry 2012; 2: e73.
28 Karchikeyan R, Marimuthu G, Ramasubramanian C, Arunachal G, BaHammam AS, Spence DW et al. Association of Per3 length polymorphism with bipolar I disorder and schizophrenia. Neuropsychiatr Dis Treat 2014; 10: 2325–2330.
29 Maglione JE, Niervergelt CM, Parimi N, Evans DS, Ancoli-Israel S, Stone KL et al. Associations of PER3 and RORA circadian gene polymorphisms and depressive symptoms in older adults. Am J Geriatr Psychiatry 2015; 23: 1075–1087.
30 Abdallah CG, Jiang L, De Feyter FM, Fasula M, Krystal JH, Rothman DL et al. Glutamate metabolism in major depressive disorder. Am J Psychiatry 2014; 171: 1320–1327.
31 Menke A, Samann P, Kloorb S, Czamara D, Lueca S, Hennings J et al. Polymorphisms within the metabotropic glutamate receptor 1 gene are associated with depression phenotypes. Psychoneuroendocrinology 2012; 37: 565–575.
32 Heuser I, Yassouridis A, Hofboer F. The combined dexamethasone/CRH test: a refined laboratory test for psychiatric disorders. J Psychiatr Res 1994; 28: 341–356.
33 Shin S, Kwon O, Kang JI, Kwon S, Oh S, Choi J et al. mGluR5 in the nucleus accumbens is critical for promoting resilience to chronic stress. Nat Neurosci 2015; 18: 1017–1024.
34 Oliet SH, Muthet JP. Regulation of N-methyl-D-aspartate receptors by astrocytic D-serine. Neuroscience 2009; 158: 275–283.
35 Hennenberger C, Papouin T, Oliet SH, Rusakov DA. Long-term potentiation depends on release of D-serine from astrocytes. Nature 2010; 463: 232–236.
36 Carvalho AF, Kohler CA, McIntyre RS, Knochel C, Brunoni AR, Thase ME et al. Peripheral vascular endothelial growth factor as a novel depression biomarker: A meta-analysis. Psychoneuroendocrinology 2015; 62: 18–26.
37 Turecki G, Grof P, Cavazzoni P, Dufy A, Grof E, Ahrens B et al. Evidence for a role of phospholipase C-gamma1 in the pathogenesis of bipolar disorder. Mol Psychiatry 1998; 3: 534–538.
38 Rohayem J, Bayle JF, Richa S. [Predictors of prophylactic response to lithium]. Encephale 2008; 34: 394–399.
39 Ikeeda A, Kato T. Biological predictors of lithium response in bipolar disorder. Psychiatry Clin Neurosci 2003; 57: 243–250.

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

© The Author(s) 2016

Supplementary Information accompanies the paper on the Translational Psychiatry website (http://www.nature.com/tp)