



OPEN

Exome-wide association study of treatment-resistant depression suggests novel treatment targets

Shrey B. Shah^{1,2,8}✉, Teja N. Peddada^{1,3,8}, Christopher Song¹, Maame Mensah¹, Heejong Sung¹, Mani Yavi⁴, Peixiong Yuan⁴, Carlos A. Zarate Jr⁴, Brian J. Mickey^{5,6}, Margit Burmeister^{6,7}, Nirmala Akula¹ & Francis J. McMahon¹✉

Treatment-resistant depression (TRD) is a severe form of major depressive disorder (MDD) with substantial public health impact and poor treatment outcome. Treatment outcome in MDD is significantly heritable, but genome-wide association studies have failed to identify replicable common marker alleles, suggesting a potential role for uncommon variants. Here we investigated the hypothesis that uncommon, putatively functional genetic variants are associated with TRD. Whole-exome sequencing data was obtained from 182 TRD cases and 2021 psychiatrically healthy controls. After quality control, the remaining 149 TRD cases and 1976 controls were analyzed with tests designed to detect excess burdens of uncommon variants. At the gene level, 5 genes, *ZNF248*, *PRKRA*, *PYHIN1*, *SLC7A8*, and *STK19* each carried exome-wide significant excess burdens of variants in TRD cases ($q < 0.05$). Analysis of 41 pre-selected gene sets suggested an excess of uncommon, functional variants among genes involved in lithium response. Among the genes identified in previous TRD studies, *ZDHHC3* was also significant in this sample after multiple test correction. *ZNF248* and *STK19* are involved in transcriptional regulation, *PHYIN1* and *PRKRA* are involved in immune response, *SLC7A8* is associated with thyroid hormone transporter activity, and *ZDHHC3* regulates synaptic clustering of GABA and glutamate receptors. These results implicate uncommon, functional alleles in TRD and suggest promising novel targets for future research.

Major depressive disorder (MDD) is a leading cause of disability worldwide, affecting over a quarter-billion people^{1,2}. With often debilitating symptoms, MDD can pervade multiple aspects of mental, emotional, and physical health. Despite advances in drug and behavioral therapies, many patients suffer from treatment-resistant depression (TRD)^{3–13} that is nonresponsive to various available treatments. The prevalence of TRD ranges from 4 to 21%, depending on how it is defined; the most common definition is non-response to two or more antidepressant trials with adequate dose and duration of 4–6 weeks¹⁴.

Genetic studies of MDD aim to identify genes that point toward potential targets of pharmacological intervention and improve treatment outcomes. While genome-wide association studies (GWAS) have linked multiple common variants with MDD^{15,16}, identifying genetic markers associated with antidepressant treatment outcomes has proven more difficult, possibly due to heterogeneity of the phenotype and limited sample sizes^{17,18}. Although antidepressant response is estimated to be substantially influenced by genetic variation, with heritability up to 40%¹⁹, GWAS have found no replicable, genome-wide significant common variants associated with the phenotype^{20–23}. This heritability gap suggests a role for uncommon genetic variants that are often missed by GWAS.

Here we investigated the hypothesis that uncommon, putatively functional genetic variants are associated with TRD. We performed a genetic association study that used exome sequencing to detect single nucleotide variants

¹Human Genetics Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA. ²Rutgers New Jersey Medical School, Newark, NJ, USA. ³Stanford University School of Medicine, Stanford, CA, USA. ⁴Experimental Therapeutics and Pathophysiology Branch and Section on the Neurobiology and Treatment of Mood Disorders, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA. ⁵Department of Psychiatry, Huntsman Mental Health Institute, University of Utah, Salt Lake City, UT, USA. ⁶Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA. ⁷Michigan Neuroscience Institute and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA. ⁸These authors contributed equally: Shrey B. Shah and Teja N. Peddada. ✉email: sbs201@njms.rutgers.edu; peddada@stanford.edu; mcmahonf@mail.nih.gov

within the protein coding regions of genes, about 2% of the entire genome^{24,25}. To enhance power, we focused on individuals with the most treatment-resistant forms of TRD and filtered for uncommon functional variants. We compared the overall burden of variants in cases to that in psychiatrically healthy controls, testing for differences primarily at the level of individual genes and pre-selected sets of genes representing key neurobiological and disease pathways. The results implicate uncommon genetic variants in TRD and suggest promising novel targets for future research.

Materials and methods

Study participants. A total of 182 unrelated individuals with TRD were selected for exome sequencing from 3 different sources: (1) 150 individuals diagnosed with MDD or bipolar disorder (BPD) whose depression symptoms failed to respond to >2 antidepressants (failed lifetime antidepressant trials: median = 4; average = 5; range = 17) collected by the NIMH Experimental Therapeutics and Pathophysiology Branch (Protocol# NCT00024635 approved and monitored by the NIH institutional review board); (2) 24 individuals referred for electroconvulsive therapy at the University of Michigan (IRB protocol HUM00039846 approved and monitored by the University of Michigan institutional review board; a subset of participants in the Bluebird study²⁶); and (3) 8 individuals selected from Level 4 of the Sequenced Treatment Alternatives to Relieve Depression study (STAR*D; Protocol# NCT00021528 approved and monitored by University of Texas Southwestern Medical Center at Dallas and the institutional review boards at each clinical site and regional center and the Data Coordinating Center and the Data Safety and Monitoring Board of the National Institute of Mental Health) on the basis of non-response to 3 or more consecutive treatments despite good physical health, absence of substance abuse, and reported good treatment adherence²⁷. For controls, we used exome data from 2021 individuals with no psychiatric disorders collected by the Swedish Hospital Discharge Register²⁸ and obtained from dbGaP (phs000473.v2.p2) in accordance with a Data Use Agreement monitored and approved by The Joint Addiction, Aging, and Mental Health Data Access Committee). All participants provided written informed consent and sample acquisition was performed in accordance with the guidelines stipulated by the above listed IRBs.

Whole-exome sequencing data generation. DNA was extracted using Gentra Puregene Blood Kit (Qiagen N. V., Hilden, Germany). Libraries were captured using RocheNimbleGen (Roche Sequencing Solutions, Basel, Switzerland), SwiftBio (Integrated DNA Technologies, Inc., Coralville, IA, USA), or SureSelect (Agilent Technologies, Inc., Santa Clara, CA, USA). Libraries were captured from samples collected by the NIMH Experimental Therapeutics and Pathophysiology Branch and STAR*D using RocheNimbleGen (n = 158). Libraries were captured from samples collected by the University of Michigan via the Bluebird Study using SwiftBio (n = 24). Libraries were captured from samples collected by Swedish Hospital Discharge Register using SureSelect (n = 2021). Paired-end sequencing was performed on the Illumina GAI or HiSeq2000 platform (Illumina, Inc., San Diego, CA, USA) for both cases and controls.

Mapping and variant calling. To minimize batch effects, sequence data from cases and controls were processed, mapped, and called jointly. Data quality (fastq files) was determined using Fastqc (v.0.11.9)²⁹. Adapters were trimmed using FASTX-Toolkit (v0.0.13)³⁰ and reads of mixed length (ranging from 63 to 100 bp) were mapped to the Ensembl human genome GRCh38.p13 using BWA-MEM³¹. PCR duplicates were removed using Picard MarkDuplicateSpark in the Genome Analysis Toolkit (GATK) (v4.1.9.0)³². Base quality of reads was recalibrated using BaseRecalibrator and ApplyBQSR in GATK (4.1.9.0). HaplotypeCaller in GATK (4.1.9.0) pipeline was used to call the variants. After joint-calling, variants were filtered via GATK's Variant Quality Score Recalibration. The data analysis pipeline is presented in Supplementary Fig. 1.

Sample quality control (QC). Samples with an atypical distribution of GC content, high (> 165 million) reads, and low (< Q20) Phred-scaled quality scores (n = 43) were excluded from further analyses. Variants with minor allele frequency (MAF) > 1% were used in PLINK (1.9) to verify sex (-check-sex), sample-level heterozygosity (-het), and relatedness (-genome)³³. Samples with sex mismatch, relatedness ($[Z0 < 0.4, Z1 > 0.2]$), or high missingness (> 5%) were excluded (n = 7).

Twelve healthy control samples that were initially misclassified as TRD cases were corrected and included in the healthy control dataset for the downstream analysis, of which only seven remained after QC filtering. One sample from the healthy controls was excluded due to sample-specific artifacts. Samples outside 5 standard deviations from the centroid of the HapMapIII European cluster (n = 28) were excluded from the analysis (Supplementary Fig. 2). Details on samples removed during quality control are presented in Supplementary Table 1. Descriptive details of individuals remaining after quality control for cases and controls are presented in Table 1.

Variant QC and annotation. A total of 1,339,135 variants were initially detected in this analysis. Genotypes with a genotype quality (GQ) < 20 and read depth (DP) < 10 were set to missing. Multi-allelic SNVs, SNVs < 10 bp from an indel, indels, intergenic SNVs, and variants with discordant genotypes between identical samples were excluded from the analysis. Variants not in Hardy-Weinberg Equilibrium (HWE $p < 1E-6$) in cases or controls were removed. Given that the cases and controls came from three different sources and where sequenced separately, locus-level heterozygosity was calculated for each set of individuals separated by source ((1) NIMH Experimental Therapeutics and Pathophysiology Branch and STAR*D; (2) Bluebird Study; (3) Swedish Hospital Discharge Register) and any variants with an observed heterozygosity of 1 were subjected to a permutation-based HWE test (HWperm). We verified the MAF of each SNV in our controls by comparing to the MAF of Non-Finnish European (NFE) in gnomAD genome (v2.1.1), gnomAD Exome (v2.1.1), and ExAC (v0.3) by a binomial test and 9550 variants with binomial test $p < 0.05$ were excluded (Supplementary Fig. 3).

Reported race	Cases	Controls
White	98.0%	100.0%
Black/African American	0.7%	0.0%
Unknown	1.3%	0.0%
Reported sex		
Female	53.0%	49.0%
Male	47.0%	51.0%
Age at enrollment		
Average (Years)	45.5	Not available
SD (Years)	12.3	Not available
Diagnosis		
Major depressive disorder	84.6%	N/A
Bipolar I	9.4%	N/A
Bipolar II	6.0%	N/A

Table 1. Demographic and phenotype data for cases and controls after QC.

A locus missingness filter of 2% per batch and 5% for all cases and controls was applied. Known false-positive exome variants identified by Fajardo et al. were excluded from this study³⁴. Details on variants removed during quality control are presented in Supplementary Table 2. Variant annotation was performed by ANNOVAR (v2020Jun08)³⁵.

Association analyses. We carried out association testing at three levels: single variants, genes, and sets of genes. False Discovery Rate correction was applied to resulting p-values using the Benjamini–Hochberg method.

Variant level analysis and filtering. SKATBinary_Single function in SNP-set (Sequence) Kernel Association Test (SKAT; v2.0.1)³⁶ package in R was used for single variant association analysis. SKAT_Null_Model_Moment Adjust with 100,000 resamplings was used to obtain the model parameters and residuals from the null model for small sample size. The efficient resampling method in the SKATBinary_Single function was used for calibrating p values of variants with low total minor allele counts³⁶. The top 10 principal components of sample ancestry as calculated by PLINK (–pca) were included as covariates in the association model along with sex³³.

Gene level analysis. Gene level analyses were conducted using three nested sets of variants: (1) Uncommon (control MAF < 5%), functional (annotated as non-synonymous and missense) variants (86,475 variants); (2) Uncommon (MAF < 5%), functional, damaging variants as annotated by ANNOVAR (v20210122)³⁵ (59,549 variants); (3) Uncommon, functional, nonsense variants (annotated as “start loss”, “stop loss”, or “stop gain” in Ensembl) (2655 variants). For each set, the variants were grouped by gene, and processed with the SKATBinary function in SKAT. Utilizing the similar analysis as described above in variant level analysis, SKAT_Null_Model_MomentAdjust with a 100,000 resamplings was used as the null model. The efficient resampling method in SKAT was used to calculate empirical p-values of gene-level associations³⁶.

Gene-set level analysis. Based on disease relevance and literature review, 41 gene sets were pre-defined and manually curated and selected for testing. Two of these gene sets were selected using DAVID, a functional annotation tool, using genes with p value < 0.05 as input³⁷. Contents of tested gene-sets are presented in Supplementary Table 3. Uncommon functional variants from these selected gene-sets were input into Multi-marker Analysis of GenoMic Annotation (MAGMA; v1.10)³⁸, a gene-set analysis tool to test for association, using the modifier multi = snp-wise to determine p values.

Overlaps between putative TRD-associated genes and those within PsychENCODE Cross Disorder Gene Coexpression modules and modules from Hartl et al. were tested utilizing a hypergeometric test^{39,40}.

Replication of top hits. Gene-level results were compared with the published top hits reported by three previous TRD GWAS^{41–43}. Since full summary statistics were not available for those studies, the Sidak method was used for multiple test correction.

Validation—Sanger sequencing. Uncommon variants in ZNF248, the only gene to be significant after Bonferroni correction, were validated in TRD cases using Sanger sequencing. Primers were designed using NCBI Primer-Blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). PCR was performed using Taq PCR Core Kit (Qiagen N. V., Hilden, Germany). For size verification, an amplified PCR product was run on a 2% Agarose gel on E-Gel Power Snap Electrophoresis System (Thermo Fisher Scientific, Waltham, MA, USA). The PCR product was purified using the QIAquick PCR purification kit (Qiagen N. V., Hilden, Germany). Sanger sequencing was performed by Psomagen (Rockville, MD, USA).

Results

After QC, exome sequencing data of 114 billion reads (~ 54 million reads per sample) in a total of 149 TRD cases and 1976 controls were included in this analysis. The Ti/Tv ratio and GC percentage were 2.75 and 49%, respectively. After filtering, a total of 189,497 variants were included.

Variant level. There was no evidence of bias in the analysis, with an exome-wide lambda value of 0.9619 (Supplementary Fig. 4). No exome-wide significant variants were identified. This was expected in this sample of unrelated people given the heterogeneous nature of TRD. The Manhattan plot is shown in Supplementary Fig. 5. Nominally significant variants are presented in Supplementary Table 4.

Gene level. Uncommon, functional variants in 14,083 genes were tested for association with TRD. There was no evidence of bias in the analysis (lambda=0.9333; Fig. 1A). The gene-level Manhattan plot is shown in Fig. 1D. Five genes (*ZNF248*, *PRKRA*, *PYHIN1*, *SLC7A8*, and *STK19*) carried excesses of uncommon, functional variants in TRD cases at an FDR < 5%. These genes represent a variety of biological themes, consistent with the likely heterogeneous architecture of TRD (Table 2). Further filtering for damaging variants included few vari-

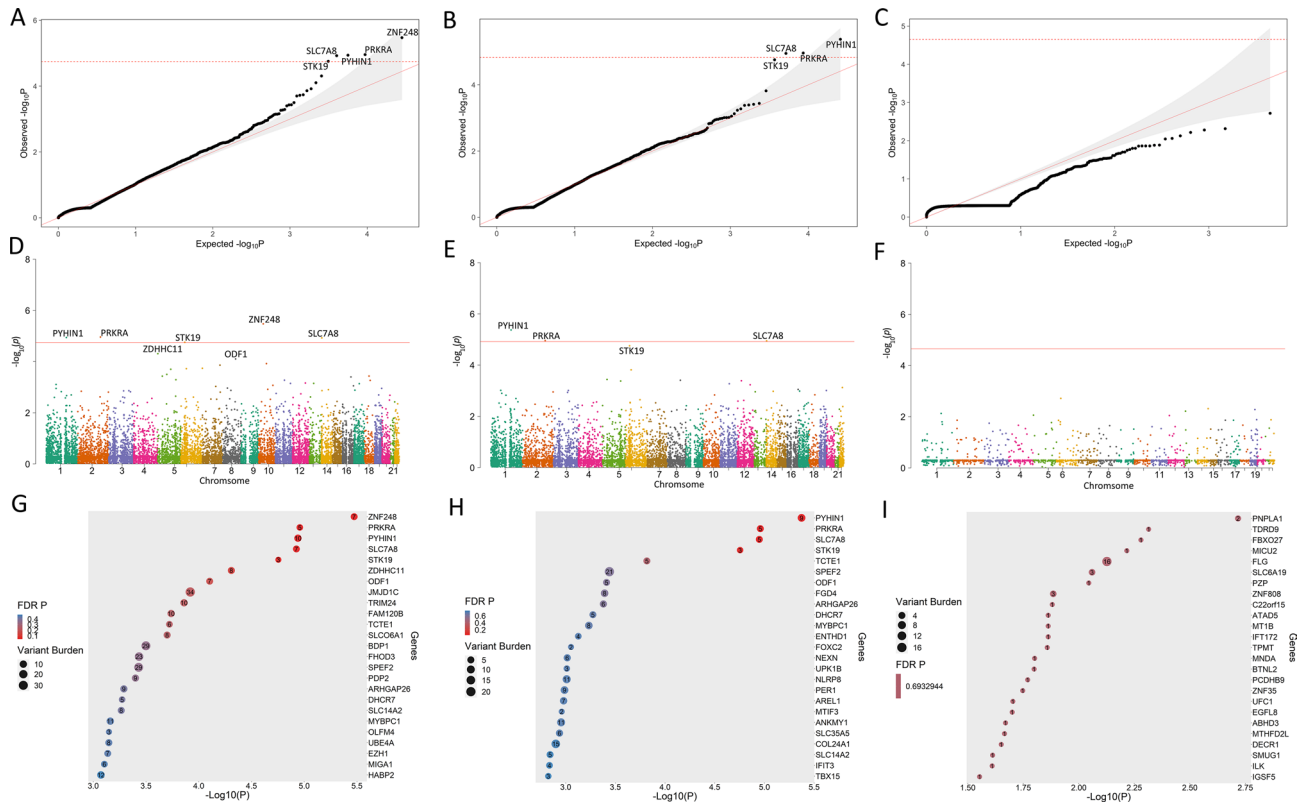


Figure 1. Gene level results. The red lines denote FDR 5% significance threshold. Top: Quantile–Quantile Plots of Uncommon (A) Functional, (B) Damaging, and (C) Nonsense Variants. Middle: Manhattan Plots of uncommon (D) Functional, (E) Damaging, and (F) Nonsense Variants. Bottom: Bubble Plots of uncommon (G) Functional, (H) Damaging, and (I) Nonsense Variants. The number inside of each circle denotes the number of variants scored in each gene.

Symbol	Gene name	Variants	Total carriers		P value	FDR Q	Depth of coverage	
			Cases (%)	Controls (%)			Cases	Controls
<i>ZNF248</i>	Zinc Finger Protein 248	7	18 (12.1)	52 (2.6)	3.4E-06	0.0418	52	37
<i>PRKRA</i>	Protein Activator of Interferon-Induced Protein Kinase	5	27 (18.1)	137 (6.9)	1.1E-05	0.0418	121	176
<i>PYHIN1</i>	Pyrin and HIN Domain Family Member 1	10	15 (10.1)	67 (3.4)	1.1E-05	0.0418	43	40
<i>SLC7A8</i>	Solute Carrier Family 7 Member 8	7	9 (6)	40 (2.0)	1.2E-05	0.0418	43	46
<i>STK19</i>	Serine/Threonine-Protein Kinase 19	3	9 (6)	12 (0.6)	1.8E-05	0.0495	45	35

Table 2. Gene-level association results.

ants in each gene and did not yield any additional significant results (Fig. 1B–C, E–F, H–I). The top 50 genes associated with TRD are shown in a bubble plot in Fig. 1G and all nominally significant genes are shown in Supplementary Table 5.

Gene set level. Of the 41 gene sets tested (Fig. 2A), only one carried a significant excess burden of uncommon functional variants in cases after FDR correction. This gene set, GO:0010226: Response to Lithium Ions, comprised 75 variants in 16 genes and was associated with TRD at $q=5.038E-03$. One additional gene set, GO:0032731: Positive Regulation Of Interleukin 1 Beta Production, was nominally significant before FDR correction. Additional filtering for damaging variants included too few variants in each gene-set and did not yield any additional significant results (Fig. 2B). Details on the gene sets tested are shown in Supplementary Table 6.

We also investigated 30 Cross Disorder Gene Coexpression Modules previously reported by PsychENCODE and 11 Coexpression network modules from Hartl et al. for association with TRD using the hypergeometric test^{39,40}. None were significantly associated with TRD after FDR correction (Supplementary Table 7).

Replication. Among the 54 unique genes associated with TRD reported in previous GWAS studies^{41–43}, Zinc Finger DHHC-Type Palmitoyltransferase 3 (*ZDHHC3*) was supported in the present study at a Sidak-corrected p value < 0.05 (Supplementary Table 8).

None of the 5 significant genes were replicated in the only other large published exome-wide association study of antidepressant response⁴³ (minimum $p=0.1339$ for damaging variants in TRD vs response).

Validation. The 12 cases with a contributing variant (rs11011379) in the gene *ZNF248* were selected for Sanger sequencing. All 12 putative heterozygotes were validated for this variant.

Discussion

We conducted an exome-wide association study at the single variant, gene, and gene-set levels among individuals diagnosed with TRD and unaffected controls. The goal was to identify potentially causal variants and increase our understanding of the neurobiology of TRD.

The results suggest a role for uncommon, functional alleles in TRD. Five genes were found to carry significant excesses of uncommon, functional variants in TRD cases. Of the 22 genes in the lithium-ion response gene set, 16 genes also carried an excess of uncommon functional variants in TRD. Among the 54 genes most strongly associated with TRD in previous studies, *ZDHHC3* was also significantly associated with TRD in the present study.

The five genes that carried excesses of uncommon, functional variants in TRD cases represent a variety of biological themes, consistent with the likely heterogeneous architecture of TRD.

ZNF248 is involved in DNA recognition, transcriptional activation, regulation of protein folding and RNA packaging, and lipid binding. A previous TRD GWAS reported association with a common SNP, rs2505705, that is an eQTL for *ZNF248* and other genes, suggesting that common variants in this gene may also contribute to TRD⁴¹. Although the function of zinc finger proteins, including *ZNF248*, in MDD remains to be fully determined⁴⁴, *ZNF248* is highly expressed in the brain. Another zinc-finger gene implicated by this study, *ZDHHC3*, regulates synaptic clustering of GABA and glutamate receptors. *STK19* resides in the MHC locus on chr 6p, a significant GWAS locus for schizophrenia, depression, and bipolar disorder⁴⁵. *STK19* is one of the 4 genes in the *RCCX* cassette, previously implicated in schizophrenia⁴⁶ and congenital adrenal hyperplasia, which itself is associated with increased risk for variety of psychiatric disorders⁴⁷. *PYHIN1* (pyrin and HIN domain family member 1) is related to interferon regulation, involved in transcriptional regulation of genes important for cell cycle control, and has been linked to depressive behaviors induced by lipopolysaccharide in mouse models⁴⁸.

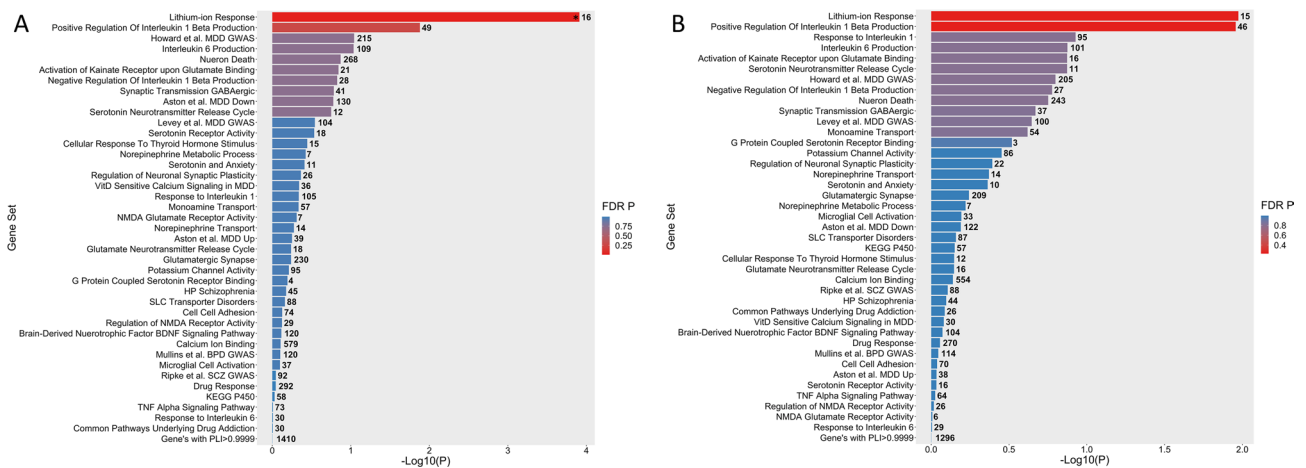


Figure 2. Gene set level results. The number adjacent to each bar denotes the number of genes contained within that gene set. (A) Bar Chart of Gene Set Analysis based on (A). Uncommon Functional and (B) Uncommon Damaging Variants. * denotes FDR $q < 0.05$.

The *PRKRA* gene plays an important role in the inhibition of translation and induction of apoptosis pathways, and is implicated in neurodegenerative diseases related to inflammation^{49–52}. *SLC7A8* encodes the Large neutral Amino acids Transporter small subunit 2 (LAT2) which enables thyroid hormone transporter activity. Deficient thyroid function can cause depressed mood and is generally treatable with adequate thyroid replacement. The use of triiodothyronine (T3) as a potential augmentation strategy in cases of TRD even without thyroid dysfunction, has been highlighted in several studies^{53,54}.

Our gene-set analyses found that the lithium-ion response gene set (GO:0010226) was significantly associated with TRD. Lithium is a first-line treatment option in bipolar disorder, in either manic or depressed states, and often used as an augmentation treatment for antidepressants. Thus, a significant percentage of the cases in our study have used lithium. We unfortunately do not have data that indicates whether these cases are responders or non-responders. Furthermore, we only have these data for cases from the NIMH Experimental Therapeutics and Pathophysiology Branch and not for all our cases. With these limitations, our results suggest that lithium-response may explain why some individuals with MDD or BPD become treatment resistant, or that lithium neurobiology may be involved in treatment response in major mood disorders.

These results also highlight a role for immune response in TRD. Among the significant associations we detected, two genes (*PYHIN1* and *PRKRA*) were involved in immune response. Interestingly, one gene set (“Up Regulation of Interleukin 1 Beta”) was associated with TRD in the present study at nominal significance levels; however, due to the number of gene sets tested this association is very limited. These results are consistent with the growing body of evidence that neuroinflammation plays a role in psychiatric disorders^{55–59}.

There have only been two previous exome wide association studies of TRD^{43,60}. In the first study, Tammiste et al. in 2013 conducted a small (n = 10) exome wide association study of treatment non-response in MDD and found variants in bone morphogenetic protein 5 (*BMP5*) to be associated with non-response to a single treatment trial of escitalopram⁶⁰. In the second study, Fabbri et al. in 2020 conducted a larger (n = 1209) exome wide association study comparing typical antidepressant response with TRD and found no exome-wide significant variants, genes, or gene sets⁴³. The significant results in the present study, with a smaller case sample than Fabbri et al., may be attributable to power gains through a larger healthy control group compared to their smaller control group that also had MDD (responsive to antidepressants). Furthermore, the selection of severe TRD cases may enhance enrichment of uncommon, high-risk variants⁶¹. We selected severe samples from the fourth level of STAR*D, samples treated with electroconvulsive therapy (ECT), and samples with multiple failed treatment trials, in contrast to the more typical TRD cases studied by Fabbri et al.

This study has several important limitations. We focused on a relatively rare form of MDD, resulting in a relatively small sample of cases. Larger samples are needed, but the discovery of genes and gene sets in this study suggests that the “extreme phenotype” strategy we employed can be an effective means to identify uncommon genetic variants—even in modestly sized samples^{62–64}. The varying definition of TRD across the case samples used in this study leaves some uncertainty as to which aspects of antidepressant treatment resistance were most influenced by uncommon variants. This heterogeneity is a necessary consequence of aggregating samples to increase sample size and power. The controls we used were psychiatrically healthy and did not include cases of typical antidepressant treatment responders. Thus, we cannot rule out that some of our findings reflect genetic risk for MDD itself, rather than TRD, but this is unlikely to account for most of our findings. We found no evidence that genes previously associated with MDD were over-represented among the TRD-associated genes we identified. Furthermore, the forms of TRD represented by the cases we studied are likely to be very uncommon in the general population and in the non-psychiatric controls we employed. Although the samples we studied used different capture libraries for cases and controls, we found no evidence of bias in the form of inflated p-values at the variant, gene, or gene-set level, with exome-wide lambda values close to 1. The cases and controls we used were of European ancestry; studies in more diverse populations are needed and may implicate different genes and mechanisms⁶⁵.

In summary, by focusing on individuals with severe TRD and filtering for uncommon variants, we identified genes and gene-sets likely to predispose to this important psychiatric illness. The genes and gene-sets identified in this study suggest that transcriptional regulation, lithium response, thyroid hormone transporter activity, and immune response are all involved with TRD. These results highlight potential biological markers and mechanisms that could lead to new therapeutic targets in TRD. We hope that our findings will provide novel hypotheses for future investigations into the etiology of TRD. With the continual decrease in sequencing costs and a growing effort to build large biobanks and cohorts of patients with TRD⁶⁶, exome sequencing analysis will be an increasingly important tool to better understand the complex trait of antidepressant treatment outcome.

Data availability

The datasets generated and/or analyzed for the current study’s controls are available in the dbGaP repository, [phs000473.v2.p2]. The datasets generated and/or analyzed for this study’s cases are available in dbGaP, [phs003329.v1.p1].

Received: 25 March 2023; Accepted: 18 July 2023

Published online: 01 August 2023

References

1. Shadrina, M., Bondarenko, E. A. & Slominsky, P. A. Genetics factors in major depression disease. *Front. Psychiatry* **9**, 334 (2018).
2. Bains, N. & Abdijadid, S. Major depressive disorder. In *StatPearls* (StatPearls Publishing, 2022).
3. Voineskos, D., Daskalakis, Z. J. & Blumberger, D. M. Management of treatment-resistant depression: Challenges and strategies. *Neuropsychiatr. Dis. Treat.* **16**, 221–234 (2020).
4. Otte, C. et al. Major depressive disorder. *Nat. Rev. Dis. Primer* **2**, 1–20 (2016).

5. Lähteenvuo, M., Taipale, H., Tanskanen, A., Rannanpää, S. & Tiihonen, J. Courses of treatment and risk factors for treatment-resistant depression in Finnish primary and special healthcare: A nationwide cohort study. *J. Affect. Disord.* **308**, 236–242 (2022).
6. Gronemann, F. H. *et al.* Treatment patterns in patients with treatment-resistant depression in Danish patients with major depressive disorder. *J. Affect. Disord.* **287**, 204–213 (2021).
7. Gronemann, F. H., Jorgensen, M. B., Nordentoft, M., Andersen, P. K. & Osler, M. Incidence of, risk factors for, and changes over time in treatment-resistant depression in Denmark: A register-based cohort study. *J. Clin. Psychiatry* **79**, 21247 (2018).
8. Reutfors, J. *et al.* Mortality in treatment-resistant unipolar depression: A register-based cohort study in Sweden. *J. Affect. Disord.* **238**, 674–679 (2018).
9. Hägg, D. *et al.* A register-based approach to identifying treatment-resistant depression—Comparison with clinical definitions. *PLoS ONE* **15**, e0236434 (2020).
10. Fife, D. *et al.* Epidemiology of pharmaceutically treated depression and treatment resistant depression in Taiwan. *Psychiatry Res.* **252**, 277–283 (2017).
11. Mahlich, J., Tsukazawa, S. & Wiegand, F. Estimating prevalence and healthcare utilization for treatment-resistant depression in Japan: A retrospective claims database study. *Drugs Real World Outcomes* **5**, 35–43 (2018).
12. Kubitz, N., Mehra, M., Potluri, R. C., Garg, N. & Cossrow, N. Characterization of treatment resistant depression episodes in a cohort of patients from a US commercial claims database. *PLoS ONE* **8**, e76882 (2013).
13. Shin, D. *et al.* Cost analysis of depression using the national insurance system in South Korea: A comparison of depression and treatment-resistant depression. *BMC Health Serv. Res.* **20**, 286 (2020).
14. Gaynes, B. N. *et al.* Defining treatment-resistant depression. *Depress. Anxiety* **37**, 134–145 (2020).
15. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).
16. Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.* **22**, 343–352 (2019).
17. Ormel, J., Hartman, C. A. & Snieder, H. The genetics of depression: Successful genome-wide association studies introduce new challenges. *Transl. Psychiatry* **9**, 1–10 (2019).
18. Levinson, D. F. *et al.* Genetic studies of major depressive disorder: Why are there no GWAS findings, and what can we do about it?. *Biol. Psychiatry* **76**, 510–512 (2014).
19. Tansey, K. E. *et al.* Contribution of common genetic variants to antidepressant response. *Biol. Psychiatry* **73**, 679–682 (2013).
20. Pain, O. *et al.* Identifying the common genetic basis of antidepressant response. *Biol. Psychiatry Glob. Open Sci.* **2**, 115–126 (2022).
21. Wigmore, E. M. *et al.* Genome-wide association study of antidepressant treatment resistance in a population-based cohort using health service prescription data and meta-analysis with GENDEP. *Pharmacogenomics J.* **20**, 329–341 (2020).
22. Fabbri, C. *et al.* The genetics of treatment-resistant depression: A critical review and future perspectives. *Int. J. Neuropsychopharmacol.* **22**, 93–104 (2019).
23. GENDEP Investigators *et al.* Common genetic variation and antidepressant efficacy in major depressive disorder: A meta-analysis of three genome-wide pharmacogenetic studies. *Am. J. Psychiatry* **170**, 207–217 (2013).
24. Seaby, E. G., Pengelly, R. J. & Ennis, S. Exome sequencing explained: A practical guide to its clinical application. *Brief. Funct. Genomics* **15**, 374–384 (2016).
25. Cheng, S. *et al.* Exome-wide screening identifies novel rare risk variants for major depression disorder. *Mol. Psychiatry* **27**, 3069–3074 (2022).
26. Mickey, B. J., Ginsburg, Y., Jensen, E. & Maixner, D. F. Distinct predictors of short- versus long-term depression outcomes following electroconvulsive therapy. *J. Psychiatr. Res.* **145**, 159–166 (2022).
27. Trivedi, M. H. *et al.* Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. *Am. J. Psychiatry* **163**, 28–40 (2006).
28. Purcell, S. M. *et al.* A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **506**, 185–190 (2014).
29. FastQC: A quality control tool for high throughput sequence data – ScienceOpen. <https://www.scienceopen.com/document?vid=de674375-ab83-4595-afa9-4c8aa9e4e736>.
30. FASTX-Toolkit. http://hannonlab.cshl.edu/fastx_toolkit/.
31. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. <https://doi.org/10.48550/arXiv.1303.3997> (2013).
32. McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
33. Purcell, S. *et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
34. Fuentes Fajardo, K. V. *et al.* Detecting false-positive signals in exome sequencing. *Hum. Mutat.* **33**, 609–613 (2012).
35. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
36. Lee, S., Fuchsberger, C., Kim, S. & Scott, L. An efficient resampling method for calibrating single and gene-based rare variant association analysis in case-control studies. *Biostatistics* **17**, 1–15 (2016).
37. Sherman, B. T. *et al.* DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkac194> (2022).
38. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized gene-set analysis of GWAS data. *PLOS Comput. Biol.* **11**, e1004219 (2015).
39. Gandal, M. J. *et al.* Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* **362**, eaat8127 (2018).
40. Hartl, C. L. *et al.* Coexpression network architecture reveals the brain-wide and multiregional basis of disease susceptibility. *Nat. Neurosci.* **24**, 1313–1323 (2021).
41. Li, Q. S., Tian, C. & Hinds, D. Genome-wide association studies of antidepressant class response and treatment-resistant depression. *Transl. Psychiatry* **10**, 1–12 (2020).
42. Fabbri, C. *et al.* Genome-wide association study of treatment-resistance in depression and meta-analysis of three independent samples. *Br. J. Psychiatry J. Ment. Sci.* **214**, 36–41 (2019).
43. Fabbri, C. *et al.* A polygenic predictor of treatment-resistant depression using whole exome sequencing and genome-wide genotyping. *Transl. Psychiatry* **10**, 1–12 (2020).
44. Squassina, A., Meloni, A., Chillotti, C. & Pisanu, C. Zinc finger proteins in psychiatric disorders and response to psychotropic medications. *Psychiatr. Genet.* **29**, 132–141 (2019).
45. Amare, A. T. *et al.* Bivariate genome-wide association analyses of the broad depression phenotype combined with major depressive disorder, bipolar disorder or schizophrenia reveal eight novel genetic loci for depression. *Mol. Psychiatry* **25**, 1420–1429 (2020).
46. Sekar, A. *et al.* Schizophrenia risk from complex variation of complement component 4. *Nature* **530**, 177–183 (2016).
47. Engberg, H. *et al.* Congenital adrenal hyperplasia and risk for psychiatric disorders in girls and women born between 1915 and 2010: A total population study. *Psychoneuroendocrinology* **60**, 195–205 (2015).
48. Yamawaki, Y. *et al.* Sodium butyrate abolishes lipopolysaccharide-induced depression-like behaviors and hippocampal microglial activation in mice. *Brain Res.* **1680**, 13–38 (2018).

49. Paquet, C. *et al.* The PKR activator PACT is induced by A β : Involvement in Alzheimer's disease. *Brain Pathol.* **22**, 219–229 (2012).
50. Zhu, P. J. *et al.* Suppression of PKR promotes network excitability and enhanced cognition by interferon- γ -mediated disinhibition. *Cell* **147**, 1384–1396 (2011).
51. Burnett, S. B., Vaughn, L. S., Sharma, N., Kulkarni, R. & Patel, R. C. Dystonia 16 (DYT16) mutations in PACT cause dysregulated PKR activation and eIF2 α signaling leading to a compromised stress response. *Neurobiol. Dis.* **146**, 105135 (2020).
52. Zhu, P. J. *et al.* Activation of the ISR mediates the behavioral and neurophysiological abnormalities in Down syndrome. *Science* **366**, 843–849 (2019).
53. Zhou, X. *et al.* Comparative efficacy, acceptability, and tolerability of augmentation agents in treatment-resistant depression: Systematic review and network meta-analysis. *J. Clin. Psychiatry* **76**, e487–498 (2015).
54. Aronson, R., Offman, H. J., Joffe, R. T. & Naylor, C. D. Triiodothyronine augmentation in the treatment of refractory depression. A meta-analysis. *Arch. Gen. Psychiatry* **53**, 842–848 (1996).
55. Najjar, S., Pearlman, D. M., Alper, K., Najjar, A. & Devinsky, O. Neuroinflammation and psychiatric illness. *J. Neuroinflammation* **10**, 816 (2013).
56. Jeon, S. W. & Kim, Y.-K. The role of neuroinflammation and neurovascular dysfunction in major depressive disorder. *J. Inflamm. Res.* **11**, 179–192 (2018).
57. Lee, C.-H. & Giuliani, F. The role of inflammation in depression and fatigue. *Front. Immunol.* **10**, 1696 (2019).
58. Yang, C., Wardenaar, K. J., Bosker, E. J., Li, J. & Schoevers, R. A. Inflammatory markers and treatment outcome in treatment resistant depression: A systematic review. *J. Affect. Disord.* **257**, 640–649 (2019).
59. Lurie, D. I. An integrative approach to neuroinflammation in psychiatric disorders and neuropathic pain. *J. Exp. Neurosci.* **12**, 1179069518793639 (2018).
60. Tammiste, A. *et al.* Whole-exome sequencing identifies a polymorphism in the BMP5 gene associated with SSRI treatment response in major depression. *J. Psychopharmacol. Oxf. Engl.* **27**, 915–920 (2013).
61. Amanat, S., Requena, T. & Lopez-Escamez, J. A. A systematic review of extreme phenotype strategies to search for rare variants in genetic studies of complex disorders. *Genes* **11**, 987 (2020).
62. Povysil, G. *et al.* Rare-variant collapsing analyses for complex traits: Guidelines and applications. *Nat. Rev. Genet.* **20**, 747–759 (2019).
63. Lee, S., Abecasis, G. R., Boehnke, M. & Lin, X. Rare-variant association analysis: Study designs and statistical tests. *Am. J. Hum. Genet.* **95**, 5–23 (2014).
64. Epi25 Collaborative & Epi25 Collaborative. Ultra-rare genetic variation in the epilepsies: A whole-exome sequencing study of 17,606 individuals. *Am. J. Hum. Genet.* **105**, 267–282 (2019).
65. Murphy, E. *et al.* Race, genetic ancestry and response to antidepressant treatment for major depression. *Neuropsychopharmacology* **38**, 2598–2606 (2013).
66. Soda, T. *et al.* International consortium on the genetics of electroconvulsive therapy and severe depressive disorders (Gen-ECT-ic). *Eur. Arch. Psychiatry Clin. Neurosci.* **270**, 921–932 (2020).

Acknowledgements

We would like to thank the all the study participants without whom this research would not have been possible. Data were analyzed on the Biowulf high performance computing platform at NIH. This research was supported in part by the National Institute of Mental Health (NIMH) Intramural Research Program (ZIA-MH002844), K23 MH 092648, The Michigan Institute for Clinical & Health Research (National Center for Advancing Translational Sciences, 2UL1 TR 000433), The Taubman Medical Research Institute, and The University of Michigan Depression Center.

Author contributions

Study was designed by F.J.M. Sequencing was done by C.S. and M.M. Analyses were carried out by N.A., S.S., T.P., M.Y., and H.S. Samples were contributed by C.Z., B.M., and M.B. All co-authors reviewed the manuscript.

Funding

The STAR*D trial (NCT00021528) was funded by the NIMH under contract N01 MH-90003 to the University of Texas Southwestern Medical Center at Dallas, and in part by the Hersh Foundation. Exome sequence data from the control sample were obtained from dbGaP at <http://www.ncbi.nlm.nih.gov/gap> through dbGaP accession number phs000473. Samples used for data analysis were provided by the Swedish Cohort Collection supported by the NIMH grant R01MH077139, the Sylvan C. Herman Foundation, the Stanley Medical Research Institute and The Swedish Research Council (grants 2009-4959 and 2011-4659). Support for the exome sequencing was provided by the NIMH Grand Opportunity grant RCMH089905, the Sylvan C. Herman Foundation, a grant from the Stanley Medical Research Institute and multiple gifts to the Stanley Center for Psychiatric Research at the Broad Institute of MIT and Harvard.

Competing interests

Dr. Carlos A. Zarate, Jr. is a full-time U.S government employee. He is listed as a coinventor on a patent for the use of ketamine in major depression and suicidal ideation. Dr. Zarate is listed as a coinventor on a patent for the use of (2R,6R)-hydroxynorketamine, (S)-dehydronorketamine and other stereoisomeric dehydro and hydroxylated metabolites of (R,S)-ketamine metabolites in the treatment of depression and neuropathic pain. Dr. Zarate is listed as co-inventor on a patent application for the use of (2R,6R)-hydroxynorketamine and (2S,6S)-hydroxynorketamine in the treatment of depression, anxiety, anhedonia, suicidal ideation and post-traumatic stress disorders. Dr. Zarate has assigned his patent rights to the U.S. government but will share a percentage of any royalties that may be received by the government. All other authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-38984-z>.

Correspondence and requests for materials should be addressed to S.B.S. or F.J.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023