

1 **Genomic predictors of response to antidepressants in Major depressive disorder** 2 **(MDD): A GWAS-Based Study on Indian cohort**

3
4 Isha Goel ^{1,2}, Gagan Hans ¹, Nand Kumar ^{1,*}, Subhradip Karmakar ^{2,*}

5
6 1. Department of Psychiatry, All India Institute of Medical Sciences, New Delhi 110029, India

7 2. Department of Biochemistry, All India Institute of Medical Sciences, New Delhi 110029, India

8 * Corresponding Authors

9 10 11 **Abstract**

12
13 Background: Major depressive disorder (MDD) affects a significant portion of the population,
14 with approximately 70% of patients not achieving adequate remission using anti-depressant
15 monotherapy, a condition known as inadequate treatment response (ITR). In the Indian population,
16 few studies have explored the genetic associations of this response, and those that exist are
17 underpowered. This study aims to identify single nucleotide variations associated with ITR in the
18 Indian population. The goal is to develop a population-specific gene panel that can identify
19 subjects at risk for ITR, allowing alternative treatment modalities.

20
21 Methods: Subjects satisfying inclusion-exclusion criteria were enrolled for the study following
22 institutional ethical approval. We recruited 120 subjects with MDD and assessed their response to
23 monotherapy (Escitalopram) using the Hamilton Depression Rating Scale (HAMD) before and 4-
24 6 weeks after medication. Responders (n=45) and non-responders (n=75) were genotyped using
25 the Infinium Global Screening SNP-Array-24 v3.0. A genome-wide association study (GWAS)
26 was performed to identify SNPs associated with ITR.

27
28 Results: Twelve lead-significant SNPs with a threshold of $< 1E-05$ were identified, suggesting an
29 association with ITR. Among these, four SNPs were located in the intronic/regulatory regions of
30 the genes *LRSAM1*, *EFCAB2*, *TRIM56* and *ZNF17*. The remaining eight SNPs were near genes
31 *ALDH1A2*, *LIPC*, *MYOCD*, *SPRY2*, *ANKRD18B*, *CCDC54*, *TNS3*, *ANKRD46*, and *FCRL2*. These
32 genes are involved in critical functions related to cell signalling, immune response,
33 neurodevelopment, regulating intracellular levels, and transcription factor binding.

34
35 Conclusions: Our study identified several novel SNPs that may be associated with ITR in MDD
36 patients, reported for the first time in an Indian cohort. Further investigations are underway to
37 determine their clinical significance and potential to be used for screening individuals responding
38 to drug interventions as user-friendly gene panels.

39 40 **Keywords**

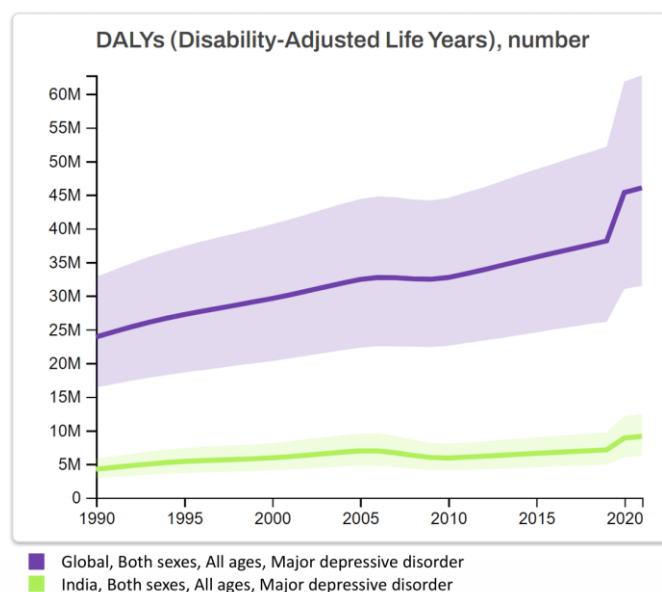
41 Depression, Drug response, Personalized treatment, Genome-wide association study, Selective
42 serotonin reuptake inhibitors

43 1. Introduction

44

45 Unipolar depression, also known as Major Depressive Disorder (MDD), is among the leading
46 causes of disability worldwide, ranking fourth in terms of disease burden and accounting for 46
47 million disability-adjusted life-years (DALYs) in 2021. According to the Global Burden of
48 Disease Study 2021, depression is the fifth leading cause of disability in India, as measured by
49 DALYs. The prevalence of depressive disorders increased from 3,745 cases per million in 2019
50 to 4,418 cases per million in 2021, with Major Depressive Disorder (MDD) alone contributing to
51 over 70% of these cases, affecting nearly 45.4 million individuals in 2021 (Global Burden of
52 Disease Collaborative Network, 2024). MDD affects individuals across all genders, ages, and
53 social backgrounds. MDD is characterized by persistent sadness, loss of interest in activities, and
54 a range of emotional and physical issues, often leading to suicidal thoughts (WHO, 2023).
55 Symptoms of MDD can vary widely among individuals. First-line antidepressant monotherapy
56 prescribed includes Selective Serotonin Reuptake Inhibitors (SSRIs) such as citalopram and
57 escitalopram, and Serotonin-Norepinephrine Reuptake Inhibitors (SNRIs) like duloxetine and
58 venlafaxine (Johnson CF et al., 2022). Despite their widespread use, the chronic and episodic
59 nature of MDD, combined with poor psychosocial functioning, significantly contributes to the
60 global economic burden. While medications and psychotherapy are effective for many,
61 approximately 30-50% of depression patients do not achieve sufficient symptom relief with
62 monotherapy, a condition known as inadequate treatment response (ITR). The situation is
63 particularly alarming in India, where more than 70% of patients experience inadequate treatment
64 response (Rafeyan R., et al., 2020; Arvind BA., et al., 2019).

65



66

67

68 Figure 1. Graph showing the numbers of disability-adjusted life-years (DALYs) due to major
69 depressive disorder in both the global and Indian contexts from 1990 to 2021. (Data and figure
70 acquired from Global Burden of Disease Collaborative Network, 2024)

71

72 One of the critical challenges in managing MDD is its multifaceted nature, which arises from
73 diverse causes, coupled with our limited understanding of its pathogenic mechanisms. Apart from
74 environmental influences like stress and trauma, genomic factors play a crucial role in the
75 development of depression. Research indicates that depression can be hereditary, with specific
76 genes involved in neurotransmitter regulation, such as serotonin, linked to a higher risk (Sullivan
77 et al., 2000). Additionally, genes related to neuroplasticity (e.g., *BDNF*), the immune system (e.g.,

78 *SPPI*), and the stress response (e.g., *5-HTTLPR*) contribute significantly to the complexity of
79 depression. (Sforzini, L., et al., 2024; Marlene et al., 2023; Caspi, A. et al., 2003). Variants in these
80 genes can predispose individuals to depression and, importantly, influence how they respond to
81 antidepressants. The National Institute of Mental Health (NIMH) conducted the sequenced
82 Treatment Alternatives to Relieve Depression (STARD) study*, which identified genetic markers
83 associated with the effectiveness of antidepressant treatments helping to tailor personalized
84 therapeutic approaches (Laje G., et al., 2009). Table 1 provides a list of genetic variants that have
85 been associated with the response to antidepressant therapy in STARD* study and few others
86 (Francis J. et al., 2006; Lemonde S. et al., 2004; Hong CJ., et al., 2006; Villafuerte SM., et al.,
87 2009; Kato M., et al., 2009; Paddock S, et al., 2007). The genetic variants identified in the
88 STARD* study were primarily based on a population in the United States (Caucasian, Black, and
89 Hispanic individuals), which may not fully represent the genetic diversity found in the Indian
90 population. Due to significant genetic differences and unique genetic variants in the Indian
91 population, the applicability of these findings may vary.

Gene	Genome build (GRCh38)	Nucleotide Change	Region	rsID
SLC6A4	chr17:30221317-30221615	Stin2.9, Stin2.10 and Stin2.12 (9, 10 or 12 copies of a 16–17 bp repeat)	SLC6A4: intronic VNTR polymorphism	STin2
SLC6A4	chr17:30237299	A>G (l/s allele)	SLC6A4 : 2KB Upstream Variant	rs4795541
SLC6A4	chr17:30237328	T>C / T>G	SLC6A4 : 2KB Upstream Variant	rs25531
HTR1A	chr5:63955024	T>G	-	rs1364043
HTR1A	chr5:63962738	C>A / C>G	HTR1A : 2KB Upstream Variant	rs6295
HTR1B	chr6:77463275	G>A / G>C / G>T	HTR1B : Synonymous Variant	rs6298
HTR2A	chr13:46837850	A>C / A>G / A>T	HTR2A : Intron Variant	rs7997012
HTR2A	chr13:46859220	C>A / C>G / C>T	HTR2A : Intron Variant	rs9316233
HTR2A	chr13:46847701	C>T	HTR2A : Intron Variant	rs1923884
ABCB1	chr7:87531245	A>G	ABCB1 : Intron Variant	rs2032583
DTNBP1	chr6:15650901	G>A	DTNBP1 : Intron Variant	rs760761
GRIK4	chr11:120792654	T>C	GRIK4 : Intron Variant	rs1954787
TPH2	chr12:71938143	G>A / G>T	TPH2 : 2KB Upstream Variant	rs4570625
DAOA	chr13:105489886	C>A / C>T	DAOA : Synonymous Variant	rs778294
DISC1	chr1:231818355	C>T	DISC1 : Missense Variant	rs6675281
DISC1	chr1:231827745	C>A / C>T	DISC1 : Intron Variant	rs1000731
FKBP5	chr6:35702206	A>C / A>G / A>T	FKBP5 : Intron Variant	rs4713916
KCNK2	chr1:215000495	C>T	LOC124904510 : Intron Variant	rs2841616
KCNK2	chr1:215239563	G>A / G>C	KCNK2 : Downstream Variant	rs7549184
KCNK2	chr1:215246328	G>A / G>C / G>T	KCNK2 : Downstream Variant	rs10494996
KCNK2	chr1:215017904	G>C / G>T	KCNK2 : Intron Variant	rs2841608
KCNK2	chr1:215077884	G>A / G>C	KCNK2 : Intron Variant	rs12136349
KCNK2	chr1:215152892	G>A / G>T	KCNK2 : Intron Variant	rs10779646

92

93 Table 1: Table listing the details of the variants reported in literature to be associated with SSRI-
94 treatment response.

95

96 This study aims to address this critical issue within the Indian population, where genetic
97 associations with treatment response have been minimally explored and existing studies lack

98 sufficient power. By leveraging a genome-wide association study (GWAS), this research seeks to
99 identify single nucleotide polymorphisms (SNPs) associated with ITR in Indian patients, paving
100 the way for personalized treatment strategies. Building on this, our study will conduct a
101 retrospective genotypic analysis of depressed patients treated at AIIMS clinics. By comparing
102 patients requiring multi-drug therapy to those responding to monotherapy, we aim to validate
103 reported proprietary markers and discover new ones, enhancing the genetic panel for predicting
104 treatment response. This approach promises to improve clinical outcomes through earlier and
105 more accurate identification of patients at risk for ITR, ultimately guiding more effective and
106 individualized treatment plans.

107

108 **2. Materials and methods**

109 **2.1 Subjects**

110 The study was conducted at the Department of Biochemistry in collaboration with the
111 Department of Psychiatry at the All-India Institute of Medical Sciences (AIIMS), New Delhi,
112 India, from November 2022 to June 2024. Subjects presenting at the psychiatric outpatient
113 department with symptoms of depressive disorder were assessed based on specific inclusion
114 and exclusion criteria. Participants aged 18-65 years, who provided informed consent, and had
115 non-psychotic major depressive disorder (DSM-V), HAMD scores ≥ 17 , and no history of
116 electroconvulsive therapy (ECT) were included. Those recommended for monotherapy
117 (Escitalopram) treatment or already on multiple antidepressants due to non-response to
118 escitalopram were further evaluated.

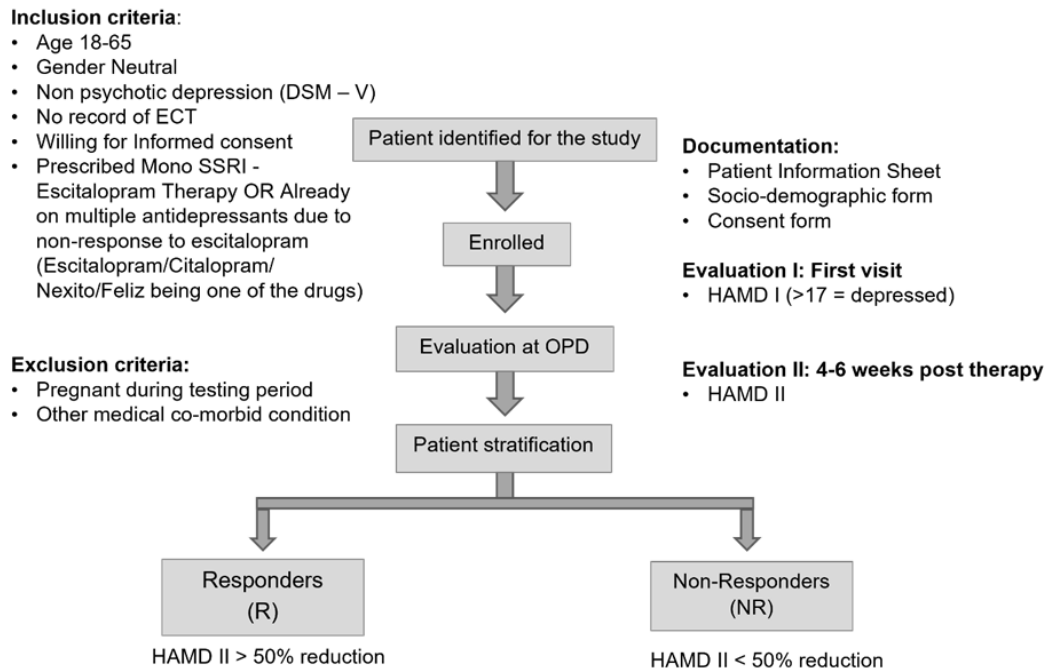
119

120 Exclusion criteria included major comorbidities such as cancer, surgery, heart disease, and
121 known or suspected pregnancy during the study period. Subjects unwilling to provide consent
122 were also excluded. Detailed patient information was recorded in both physical files and
123 electronic documents, encompassing the patient information sheet, informed consent form,
124 referral form with socio-demographic details, medical history, primary clinical diagnosis, and
125 medication information. Depression severity was assessed using the 17-item HAMD scale
126 through structured interviews.

127

128 To determine responders and non-responders to antidepressant treatment, the severity of
129 patients' depression was evaluated twice using the aforementioned scales at two time points.
130 The first assessment was conducted on the day of patient enrolment, with those having a
131 HAMD score ≥ 17 proceeding to further evaluation and blood collection (Evaluation I). Blood
132 samples (4-5 ml) were drawn from consenting patients by an experienced practitioner and
133 stored in sterilized EDTA vials. Patients were re-evaluated for depression severity after 4-6
134 weeks of antidepressant treatment using the HAMD scale (Evaluation II). A reduction in
135 HAMD score of $\geq 50\%$ categorized patients as responders, while those failing to meet this
136 criterion were classified as non-responders. Patients already on multiple antidepressant
137 therapies due to non-response to escitalopram were classified as non-responders to
138 monotherapy.

139



140
141 Figure 2: Work flow of sampling with details of Inclusion/Exclusion criteria and Evaluation
142 metrics for the stratification of patients into responders and non-responder group.
143

144 2.2 Sample preparation

145 Blood samples collected at the OPD were immediately sent to the molecular lab in the
146 Department of Biochemistry for DNA isolation. Using a genomic DNA extraction kit (Qiagen
147 QIAamp Blood Midi-Kit – Catalogue no. 51185), genomic DNA (gDNA) was isolated from
148 the blood following the standard kit protocol. The quality and quantity of the isolated gDNA
149 were assessed using a spectrophotometer (Nano Drop Technologies, Wilmington, DE, USA)
150 and gel electrophoresis. A DNA sample with a concentration of 100 ng/ μ L and an OD value
151 (260/280 ratio) of approximately 1.8 was considered acceptable in terms of quality.

152 Subsequently, 2 μ g of purified gDNA was used for genotyping with the Illumina Infinium
153 Global Screening BeadChip Array v3.0 (Illumina, San Diego, CA, USA). This array enables
154 the genotyping of over 600,000 highly optimized multi-ethnic, curated clinical variants (based
155 based on ClinVar, NHGRI-EBI and PharmGKB databases) spanning both coding and non-
156 coding regions of the human genome. The genotyping data was generated in the form of .idat
157 files, which were then analysed using Genome Studio 2.0.

158 2.3 Quality Control

159 Raw genotyping data files (.dat files) generated from the Illumina GSA bead-chip array were
160 imported into GenomeStudio (v2.0) software (GenomeStudio Software (illumina.com)) using
161 the Illumina hg38 manifest file and sample information sheet. SNP calls were generated and
162 clustered, followed by quality control and filtering based on call rates, strand orientation, and
163 gender estimation. Metadata was enriched with phenotypes of interest, and the cleaned data
164 was exported into PLINK format for further processing. Using PLINK v1.9, the genotyping
165 data underwent quality control procedures, including assessments of SNP and individual
166 missingness, minor allele frequency, gender correction, heterozygosity, relatedness, and
167 Hardy-Weinberg equilibrium (HWE) parameters. Highly heterozygous individuals,
168 duplicates, closely related individuals, and SNPs in build-specific high linkage disequilibrium
169 (LD) regions were removed.

170

171 2.4 Population Stratification and Imputation

172 The LD-pruned data was then checked for outliers using multidimensional scaling (MDS)
173 analysis based on population stratification methodology, referencing population data from the
174 1000 Genomes Project phase 3. Outliers were identified using component-wise outlier
175 detection and manual inspection of the MDS plot. Subjects that failed to follow population
176 homogeneity were discarded. Phasing and imputation were performed using the Michigan
177 Imputation Server 2 (Das S., et al., 2016) with the 1000 Genomes Project phase 3 dataset. This
178 platform uses Eagle v2.4 to estimate haplotype phase and Minimac4 for imputation using a
179 specific reference panel (HRC or 1000 Genomes). SNPs with $R^2 < 0.3$ and minor allele
180 frequency (MAF) < 0.05 were excluded.

181

182 2.5 Statistical Analysis and Annotation

183 The filtered SNPs were analysed for binary traits using the PLINK --assoc command, which
184 employs a chi-squared test to identify SNPs significantly associated with the phenotype of
185 interest. Quality and association results were visualized through QQ plots and Manhattan plots.
186 Following this, SNP results were functionally annotated using Annovar to identify known and
187 novel variants along with their associated genes.

188

189 3. Results

190 3.1 Patient enrolment

191 Total, 257 patients were screened at the psychiatric outpatient department of AIIMS. Out of
192 these, 213 patients were enrolled in the study, with the remaining excluded due to age criteria,
193 comorbidities, or refusal of consent. Within the enrolled group, 11 patients were excluded due
194 to unsuccessful sample preparation. 82 patients dropped out during follow-up. Ultimately, 120
195 patients successfully completed all the assessments and medication. Table 2 shows
196 demographic details of the patients enrolled and their response group based on the HAMD
197 scale scores after successful evaluations.

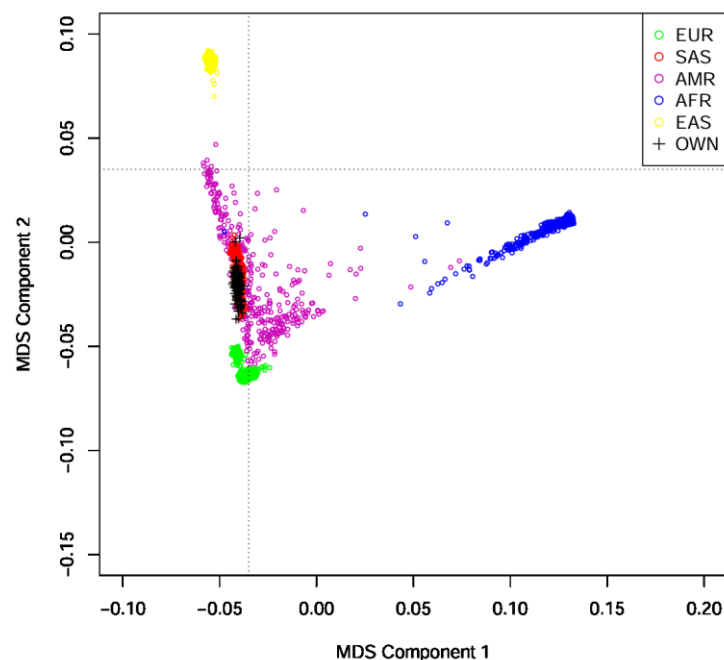
Characteristics	Patients N=120	
	Responders N=45	Non-responders N=75
Gender		
Male (57)	29	28
Female (63)	16	47
Age		
18-40 years (81)	31	50
>40 years (39)	14	25

198 Table 2: Demographic characteristics of the patients who have completed assessment.

199 Among the 120 patients studied, 37.5% were categorized as responders, whereas 62.5% were
200 non-responders. Gender distribution showed that males constituted 47.5% of the total sample,
201 with 24.2% of them being responders and 23.3% non-responders. Conversely, females made
202 up 52.5% of the sample, with 13.3% responders and 39.2% non-responders, reflecting a higher
203 prevalence of depression in females. Age-wise, 67.5% of the patients were between 18-40
204 years, with 25.8% responders and 41.7% non-responders, indicating a higher impact of
205 depression in the younger age group. In the older age group (above 40 years), 32.5% of the
206 patients were represented, with 11.7% being responders and 20.8% non-responders. These
207 statistics suggest that depression disproportionately affects younger individuals more
208 significantly.

209 3.2 QC and Population stratification

210 Genotyping calls from the raw data of all 120 samples (45 True-Responders, 65 Non-
211 Responders, and 10 Resistant) were generated using Genome Studio 2.0. Out of the total 120
212 individuals, two were removed due to high heterozygosity, and one was removed due to
213 duplicate sampling, resulting in 117 samples being retained for association analysis. SNP call
214 rates for all individuals exceeded 98%. Of the initial 654,027 SNPs, 400,460 remained after
215 filtering based on the aforementioned quality control metrics. Utilizing 1000g phase 3
216 population data, the remaining SNPs were subjected to MDS analysis for the assessment of
217 population structure and identification of outliers. As shown in figure 3, all the 117 samples
218 (labelled as OWN) from our study overlapped with individuals from SAS (South Asian)
219 population that is composed of five sub-populations including Sri Lankan Tamils from the UK,
220 Indian Telugu from the UK, Gujrati Indian from Houston, Texas, Punjabi from Lahore and
221 Bengali from Bangladesh; the closest representative of Indian population in 1000 Genome
222 phase 3 dataset. As inspected manually as well as computationally there were no significant
223 outliers among the subjects under study, and all were processed further for imputation and
224 association analysis.



225

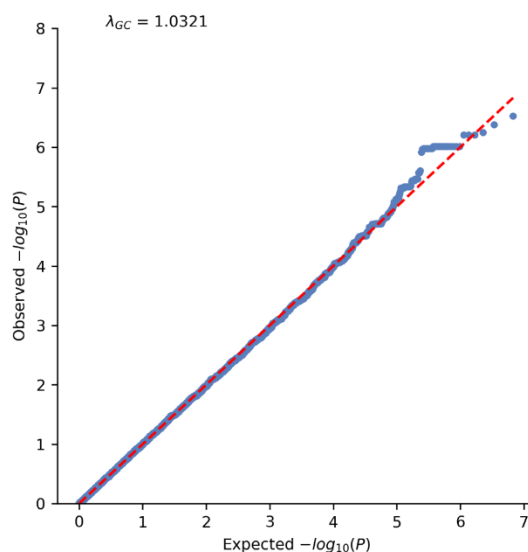
226 Figure 3: MDS plot showing the population structure of the 1000 Genome phase 3 data with
227 individuals from different global populations including European (EUR), American (AMR),
228 African (AFR), South Asian (SAS), and East Asian (EAS) along with the data of 117
229 individuals from this study labelled as OWN.

230 **3.3 Phasing and Imputation:** Additional pre-processing steps were performed for
231 imputation as recommended on the Michigan Server of Imputation, following which 3,82,562
232 SNPs and 117 individuals passed all the filters and QC. These SNPs were then subjected to
233 imputation using all the populations in the 1000 Genome phase 3 data. Variants with R2 score
234 < 0.3 and MAF < 0.05 were filtered out, retaining 67,78,587 imputed SNPs with an average
235 genotype calling rate of 1.0.

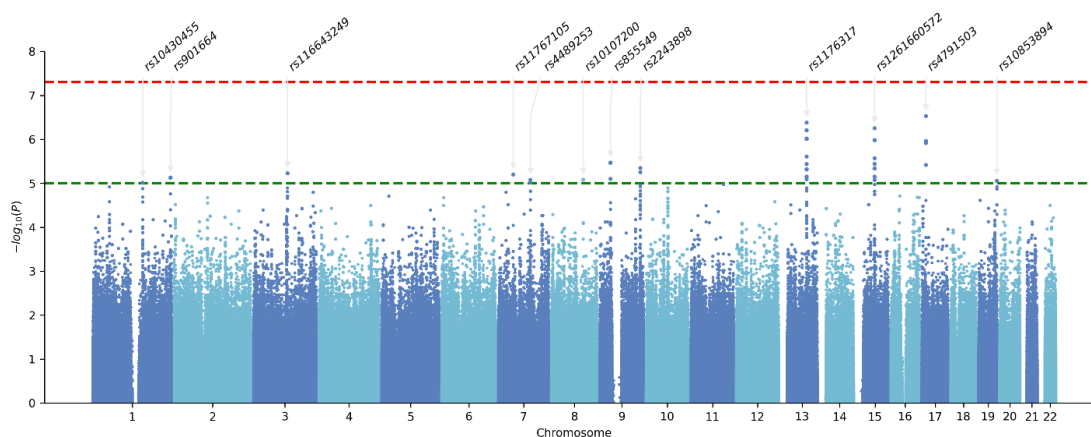
236

237 **3.4 SNPs associated with ITR**

238 Of the total 117 individuals, 73 belonged to the non-responder (non-responders to monotherapy
239 + those on multi-antidepressant treatment due to non-response to monotherapy) groups and 44
240 belonged to the responder (responders to monotherapy) group. Association based on binary
241 (Case/Control) trait was performed using --assoc option in PLINK v1.7. A Quantile-Quantile
242 Plot was utilized to assess the quality of the data used for association based on Genomic
243 inflation estimate: lambda (based on median chisq) = 1.0321, which was well within the
244 acceptable range and represents good quality data, as shown in Figure 4. SNP results were
245 represented in the form of a Manhattan plot (figure 5) and with a p-value cut-off of 1E-05 (1 x
246 10⁻⁵), a total of 78 significant variants including 12 lead SNPs were identified, suggestive of a
247 significant association with ITR. Table 3-4 lists the details of the 12 lead SNPs identified
248 suggestive of an association with ITR. Complete list of 78 significant variants is provided in
249 the supplementary file S1.



250 Figure 4: QQ (Quantile-Quantile) plot of Non-Responder vs Responder comparison showing
251 the quality of the data with Genomic Inflation rate Estimate: Lambda = 1.0321, representing
252 high-quality data.



253 Figure 5: Manhattan plot demonstrating the relationship between SNPs and non-responders,
254 highlighting some of the lead SNPs (rsIDs) with pval < 1 x 10⁻⁵. The green horizontal line at
255 (P = 1E-05) represents the genome-wide significance criterion. Red line represents p-value
256 cut-off of 5E-08.

Sno	CHR	POS	SNPID	P-value	Minor Allele	Major Allele	Frequency of Minor Allele in Cases	Frequency of Minor Allele in Controls	Odds Ratio	Asian Population Frequency
1	17	12569992	17:12569992:T:C	2.96E-07	C	T	0.1575	0.4659	0.2144	T=0.7148; C=0.2852
2	13	80651301	13:80651301:G:A	4.18E-07	A	G	0.2397	0.5682	0.2396	G=0.66; A=0.34
3	15	58359744	15:58359744:C:CACAA	5.62E-07	CACAA	C	0.03425	0.25	0.1064	ACA=1.00; ACAAACA=0.00
4	9	33653459	9:33653459:G:A	3.42E-06	A	G	0.08904	0.3295	0.1989	G=0.85; A=0.15
5	9	127457711	9:127457711:G:C	4.50E-06	G	C	0.3288	0.6364	0.2799	G=0.94; C=0.06
6	3	107162120	3:107162120:G:A	5.93E-06	A	G	0.2055	0.4886	0.2706	G=0.64; A=0.36
7	7	46940824	7:46940824:T:C	6.33E-06	C	T	0.05479	0.2614	0.1638	T=0.8554; C=0.1446
8	1	245020277	1:245020277:C:G	7.49E-06	C	G	0.06849	0.2841	0.1853	C=0.15; G=0.85
9	8	100440707	8:100440707:C:T	8.30E-06	T	C	0.3699	0.6705	0.2885	C=0.4960; T=0.5040
10	7	101098969	7:101098969:G:A	8.39E-06	A	G	0.5959	0.2955	3.516	G=0.48; A=0.52
11	19	57422202	19:57422202:A:G	8.80E-06	G	A	0.3836	0.1136	4.853	A=0.631; G=0.369
12	1	157733448	1:157733448:T:A	9.79E-06	T	A	0.3288	0.625	0.2939	T=0.38; A=0.62

257 Table 3: List of SNPs significantly associated with ITR in depression patients

Sno	CHR	rsID	Type	Nearest Coding Gene	Nearest Coding Gene Name
1	17	rs4791503	LINC00670 : Intron Variant	MYOCD	myocardin
2	13	rs1176317	Intergenic	SPRY2	sprouty RTK signaling antagonist 2
3	15	rs1261660572	Intergenic	ALDH1A2; LIPC	aldehyde dehydrogenase 1 family member A2; Lipase C
4	9	rs855549	Intergenic	ANKRD18B	ankyrin repeat domain 18B
5	9	rs2243898	LRSAM1 : Intron Variant	LRSAM1	leucine rich repeat and sterile alpha motif containing 1
6	3	rs116643249	LINC00882(ncRNA) : Intron Variant	CCDC54	coiled-coil domain containing 54
7	7	rs11767105	Intergenic	TNS3	tensin 3
8	1	rs901664	EFCAB2 : Intron Variant	EFCAB2	EF-hand calcium binding domain 2
9	8	rs10107200	LOC105375670 : Intron Variant	ANKRD46	ankyrin repeat domain 46
10	7	rs4489253	TRIM56:1000B Downstream Variant	TRIM56	tripartite motif containing 56
11	19	rs10853894	ZNF17 : 500B Downstream Variant	ZNF17	zinc finger protein 17
12	1	rs10430455	Intergenic	FCRL2	Fc receptor like 2

258 Table 4: List of nearest genes of the SNPs significantly associated with ITR in depression
259 patients

260

261 The association analysis of single nucleotide polymorphisms (SNPs) identified several genetic
262 variants significantly associated with antidepressant treatment response. Among the most
263 notable findings, rs4791503 on chromosome 17 exhibited a highly significant p-value of
264 2.96×10^{-7} . The minor allele (C) was observed at a frequency of 0.1575 in cases compared to
265 0.4659 in controls, resulting in an odds ratio of 0.2144. This SNP, an intron variant within the
266 *LINC00670* gene, is situated near the *MYOCD* gene, which encodes myocardin, a critical
267 regulator of smooth muscle gene expression. Another significant SNP, rs1176317 on
268 chromosome 13, displayed a p-value of 4.18×10^{-7} . Here, the minor allele (A) was present in
269 23.97% of cases and 56.82% of controls, yielding an odds ratio of 0.2396. This variant is
270 adjacent to the *SPRY2* gene, known for its role in modulating receptor tyrosine kinase signaling
271 pathways. The SNP rs1261660572 on chromosome 15, with a p-value of 5.62×10^{-7} , showed
272 the minor allele (CACAA) frequency to be 0.03425 in cases versus 0.25 in controls, leading to
273 an odds ratio of 0.1064. This SNP is located upstream of the *ALDH1A2* gene, which is involved
274 in retinoic acid biosynthesis.

275 Further significant associations were found with rs855549 on chromosome 9 (p-value
276 3.42×10^{-6}), rs2243898 on chromosome 9 (p-value 4.50×10^{-6}), and rs116643249 on
277 chromosome 3 (p-value 5.93×10^{-6}). These SNPs are linked to the *ANKRD18B*, *LRSAMI*, and
278 *CCDC54* genes, respectively, all of which play essential roles in protein interactions and
279 signalling pathways. Additionally, the SNPs rs4489253 on chromosome 7 (p-value 8.39×10^{-6})
280 and rs10853894 on chromosome 19 (p-value 8.80×10^{-6}) demonstrated the highest odds ratios
281 of 3.516 and 4.853, respectively, indicating strong associations with the treatment response.
282 These variants are situated near the *TRIM56* and *ZNF17* genes, involved in ubiquitination and
283 transcription regulation, respectively. These findings underscore the potential genetic markers
284 that can predict antidepressant treatment response. The identified genes are integral to various
285 biological processes, including gene regulation, signal transduction, and cellular metabolism,
286 thereby contributing to the observed variability in treatment response among individuals.

287 4. Discussion and limitations

288 Our study represents one of the pioneering GWAS investigations into inadequate drug response
289 in depression within the Indian population, utilizing the Illumina Infinium SNP-array platform.
290 The objective was to identify key genetic variations contributing to variability in treatment
291 response. Subjects were carefully selected based on thorough clinical assessments.

292 Our study has certain limitations, primarily due to the moderate sample size, which is smaller
293 than typical genome-wide association studies (GWAS). The vast genetic diversity within
294 India's large population adds complexity, as the diverse genetic backgrounds and ethnicities
295 across different regions introduce significant variability not fully captured by the SAS sample
296 in the 1000G population dataset. This variability makes it difficult to generalize our findings
297 to the entire population and can introduce confounding factors, reducing the study's power to
298 detect significant associations. Additionally, cultural stigma around psychological disorders in
299 India often leads to underreporting and a lack of formal diagnoses, affecting the quality and
300 accuracy of the data. We also witnessed significant dropouts in follow-ups, which is typical in
301 longitudinal studies. Despite these shortcomings, our results are intriguing and of clinical
302 significance given the fact that a considerable number of people who die by suicide have a
303 mental health issue. Identifying individuals at risk early on could avoid such disasters through
304 proper counselling and other preventive measures.

305 Our study successfully identified several genetic variants associated with treatment response,
306 underscoring the importance of including diverse populations in genetic research. To our
307 knowledge, this study is one of the first to explore SNP-based heritability of inadequate drug

308 response in depression among the Indian population. Most previous GWAS studies have
309 focused on American ancestries. Our findings highlight the need for larger, more inclusive
310 studies to understand the genetic underpinnings of depression and its treatment response in
311 diverse populations. This research marks a significant step toward identifying genetic
312 variations that influence drug response in depression, laying the groundwork for future research
313 and the potential development of personalized treatment strategies that consider the genetic
314 diversity of the Indian population.

315 In our study, we identified over 25 significant intergenic variants on chromosome 15, located
316 near the regulatory regions of genes *ALDH1A2* and *LIPC*. *ALDH1A2* (Aldehyde
317 Dehydrogenase 1 Family Member A2) encodes an enzyme involved in retinoate signalling,
318 ESR-mediated signalling, and cofactor metabolism. *LIPC* (Lipase C, hepatic type) plays a
319 crucial role in plasma lipoprotein assembly, remodelling, and cholesterol metabolism, with
320 previous studies linking lipoprotein levels to depression and cognitive behaviour (Jia QF et al.,
321 2020). Additionally, one significant variant on chromosome 9 was located in the intronic region
322 of the gene *LRSAMI*, which encodes a ring finger protein involved in intracellular trafficking,
323 antigen processing, signal transduction, cell adhesion, and the adaptive immune response.
324 Mutations in *LRSAMI* have been linked to neurodegenerative diseases such as Parkinson's
325 disease (Mishra R et al., 2020).

326 We also discovered four significant intronic variants on chromosome 1 associated with the
327 *EFCAB2* gene (EF-Hand Calcium Binding Domain 2), which is involved in calcium ion
328 binding. Although its function is not yet fully determined, the GENDEP clinical trial suggests
329 its role in neurodevelopment and cognitive function, particularly in axon cargo transport and
330 neurotransmitter release (Uher, R., 2009; Ren H et al., 2018). Furthermore, a downstream
331 variant on chromosome 19 (rs10853894) was significantly associated with the *ZNF17* gene, a
332 DNA-binding transcription factor implicated in numerous essential cellular processes,
333 including transcriptional regulation, ubiquitin-mediated protein degradation, signal
334 transduction, actin targeting, DNA repair, and cell migration. ZNF proteins have been
335 associated with various neuro-related diseases, including Alzheimer's, schizophrenia,
336 depression, anxiety, trauma, epilepsy, and autism (Bu S et al., 2021).

337 Other notable genes, such as *TRIM56*, *MYOCD*, *SPRY2*, *ANKRD18B*, *CCDC54*, *TNS3*,
338 *ANKRD46*, and *FCRL2*, were also identified, highlighting their involvement in innate immune
339 response, nervous system development, PDGFR-beta signalling pathway, FGFR2/3 signalling,
340 and receptor tyrosine kinase signalling. These findings provide insights into the genetic
341 underpinnings of drug response in depression, suggesting potential targets for personalized
342 treatment strategies.

343 Notably, none of the identified SNPs, genes, or loci have been previously reported in earlier
344 GWAS studies on depression. This discrepancy can be attributed to the fact that most GWAS
345 research has predominantly focused on populations of American and Caucasian ancestry, with
346 limited studies conducted on the Asian population, particularly the Indian demographic.
347 Despite this, the major functional ontologies of the identified genes such as Immune response,
348 cell signalling, metabolic pathways, and transcriptional activities were found to overlap
349 significantly with those reported in previous research.

350 5. Acknowledgments

351 The authors thank the clinical staff (Psychologists) Praggapti Ghosh, Dr. Pratibha Lamba,
352 Tanya Bhatia, Geetanjali Arora, and Hafsa Masroor for their contributions to the psychological
353 assessments. We also acknowledge Saurabh Dubey for his technical support, Shubham Dhiryan
354 and Komal Kashyap for data entry.

355 6. Author Contributions

356 Conceptualization, S.K., N.K. and G.H.; methodology, I.G., S.K., N.K. and G.H.; software,
357 I.G. and S.K.; formal analysis, I.G.; investigation, S.K., G.H.; data curation, I.G., G.H.;
358 writing—original draft preparation I.G.; writing—review and editing, I.G., S.K., N.K. and
359 G.H.; supervision, S.K., N.K. and G.H.; project administration, S.K., N.K. and G.H.; funding
360 acquisition, S.K., N.K.

361

362 7. Institutional Review Board Statement

363 Institute Ethics Committee of All India Institute of Medical Sciences, New Delhi gave ethical
364 approval for this work, file no. IEC-523/05.06.2020, RP-16/2020

365

366 8. Informed Consent Statement

367 Informed consent was obtained from all subjects involved in the study.

368

369 9. Data Availability Statement

370 All data produced in the present study are available upon reasonable request to the authors.

371

372 10. Conflicts of Interest

373 The authors declare no conflict of interest.

374

375 11. Funding Statement

376 This research was funded by ICMR to vide project No. 5/4-4/187/M/2020-NCD-II

377

378 12. References

379 1. Institute of Health Metrics and Evaluation (2024). Global Health Data Exchange
380 (GHDx). <https://vizhub.healthdata.org/gbd-results> (Accessed 20 November 2024).

381

382 2. World Health Organization: WHO. (2023, March 31). *Depressive disorder*
383 (*Depression*). <https://www.who.int/news-room/fact-sheets/detail/depression> (Accessed 20
384 November 2024).

385

386 3. Johnson CF, Maxwell M, Williams B, Dougall N, MacGillivray S. Dose-response effects of
387 selective serotonin reuptake inhibitor monotherapy for the treatment of depression:
388 systematic review of reviews and meta-narrative synthesis. *BMJ Med.* 2022 Dec
389 1;1(1):e000017 <https://doi.org/10.1136/bmjmed-2021-000017>.

390

391 4. Rafeyan R, Papakostas GI, Jackson WC, et al. Inadequate response to treatment in major
392 depressive disorder: augmentation and adjunctive strategies. *J Clin*
393 *Psychiatry.* 2020;81(3):OT19037BR3. <https://doi.org/10.4088/JCP.OT19037BR3>.

394

395 5. Arvind BA, Gururaj G, Loganathan S, Amudhan S, Varghese M, Benegal V, Rao GN,
396 Kokane AM, B S C, P K D, Ram D, Pathak K, R K LS, Singh LK, Sharma P, Saha PK, C
397 R, Mehta RY, T M S; NMHS collaborators group. Prevalence and socioeconomic impact of
398 depressive disorders in India: multisite population-based cross-sectional study. *BMJ Open.*
399 2019 Jun 27;9(6):e027250. <https://doi.org/10.1136/bmjopen-2018-027250>.

400

401 6. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and
402 meta-analysis. *Am J Psychiatry.* 2000 Oct;157(10):1552-62.
403 <https://doi.org/10.1176/appi.ajp.157.10.1552>.

- 404
405 7. Sforzini, L., Marizzoni, M., Bottanelli, C. *et al.* Transcriptomic profiles in major depressive
406 disorder: the role of immunometabolic and cell-cycle-related pathways in depression with
407 different levels of inflammation. *Mol Psychiatry* (2024). [https://doi.org/10.1038/s41380-](https://doi.org/10.1038/s41380-024-02736-w)
408 [024-02736-w](https://doi.org/10.1038/s41380-024-02736-w)
409
410 8. Marlene, Santos., Luís, Lima., Serafim, Carvalho., Jorge, Mota-Pereira., Paulo, Pimentel.,
411 Dulce, Maia., Diana, Correia., M., Fátima, Barroso., S., Gomes., Agostinho, Cruz., Rui,
412 Medeiros. (2023). The Impact of BDNF, NTRK2, NGFR, CREB1, GSK3B, AKT, MAPK1,
413 MTOR, PTEN, ARC, and SYN1 Genetic Polymorphisms in Antidepressant Treatment
414 Response Phenotypes. *International Journal of Molecular Sciences*, 24(7):6758-6758.
415 <https://doi.org/10.3390/ijms24076758>
416
417 9. Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., & Poulton, R.
418 (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT
419 gene. *Science*, 301(5631), 386-389. <https://doi.org/10.1126/science.1083968>.
420
421 10. Laje G, Perlis RH, Rush AJ, McMahon FJ. Pharmacogenetics studies in STAR*D: strengths,
422 limitations, and results. *Psychiatr Serv*. 2009 Nov;60(11):1446-57.
423 <https://doi.org/10.1176/appi.ps.60.11.1446>.
424
425 11. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, Vrieze S, Chew EY, Levy S,
426 McGue M, Schlessinger D, Stambolian D, Loh PR, Iacono WG, Swaroop A, Scott LJ, Cucca
427 F, Kronenberg F, Boehnke M, Abecasis GR, Fuchsberger C. Next-generation genotype
428 imputation service and methods. *Nature Genetics* 48, 1284–1287 (2016)
429
430 12. HAMILTON M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960
431 Feb;23(1):56-62. <https://doi.org/10.1136/jnnp.23.1.56>.
432
433 13. Paddock S, Laje G, Charney D, Rush AJ, Wilson AF, Sorant AJ, Lipsky R, Wisniewski SR,
434 Manji H, McMahon FJ. Association of GRIK4 with outcome of antidepressant treatment in
435 the STAR*D cohort. *Am J Psychiatry*. 2007 Aug;164(8):1181-8.
436 <https://doi.org/10.1176/appi.ajp.2007.06111790>.
437
438 14. Kato M, Fukuda T, Wakeno M, Okugawa G, Takekita Y, Watanabe S, Yamashita M, Hosoi
439 Y, Azuma J, Kinoshita T, Serretti A. Effect of 5-HT1A gene polymorphisms on
440 antidepressant response in major depressive disorder. *Am J Med Genet B Neuropsychiatr*
441 *Genet*. 2009 Jan 5;150B(1):115-23. <https://doi.org/10.1002/ajmg.b.30783>.
442
443 15. Villafuerte SM, Vallabhaneni K, Sliwerska E, McMahon FJ, Young EA, Burmeister M.
444 SSRI response in depression may be influenced by SNPs in HTR1B and HTR1A. *Psychiatr*
445 *Genet*. 2009 Dec;19(6):281-91. <https://doi.org/10.1097/YPG.0b013e32832a506e>.
446
447 16. Hong CJ, Chen TJ, Yu YW, Tsai SJ. Response to fluoxetine and serotonin 1A receptor (C-
448 1019G) polymorphism in Taiwan Chinese major depressive disorder. *Pharmacogenomics J*.
449 2006 Jan-Feb;6(1):27-33. <https://doi.org/10.1038/sj.tpj.6500340>.
450
451 17. Lemonde S, Du L, Bakish D, Hrdina P, Albert PR. Association of the C(-1019)G 5-HT1A
452 functional promoter polymorphism with antidepressant response. *Int J*

- 453 Neuropsychopharmacol. 2004 Dec;7(4):501-6.
454 <https://doi.org/10.1017/S1461145704004699>.
455
- 456 18. Francis J. McMahon, et al. Variation in the Gene Encoding the Serotonin 2A Receptor Is
457 Associated with Outcome of Antidepressant Treatment. *The American Journal of Human*
458 *Genetics* 2006, 78(5):804-814.
459
- 460 19. Uher, R., Huezio-Diaz, P., Perroud, N. et al. Genetic predictors of response to antidepressants
461 in the GENDEP project. *Pharmacogenomics J* 9, 225–233 (2009).
462 <https://doi.org/10.1038/tpj.2009.12>.
463
- 464 20. Jia QF, Yang HX, Zhuang NN, Yin XY, Zhu ZH, Yuan Y, Yin XL, Wang Y, Cheung EFC,
465 Chan RCK, Hui L. The role of lipoprotein profile in depression and cognitive performance:
466 a network analysis. *Sci Rep.* 2020 Nov 26;10(1):20704. [https://doi.org/10.1038/s41598-](https://doi.org/10.1038/s41598-020-77782-9)
467 [020-77782-9](https://doi.org/10.1038/s41598-020-77782-9).
468
- 469 21. Mishra, R., Amanullah, A., Upadhyay, A., Dhiman, R., Dubey, A. R., Singh, S., Prasad, A.,
470 & Mishra, A. (2020). Ubiquitin ligase LRSAM1 suppresses neurodegenerative diseases
471 linked aberrant proteins induced cell death. *International Journal of Biochemistry & Cell*
472 *Biology*, 120, 105697. <https://doi.org/10.1016/j.biocel.2020.105697>.
473
- 474 22. Ren, H., Fabbri, C., Uher, R. et al. Genes associated with anhedonia: a new analysis in a
475 large clinical trial (GENDEP). *Transl Psychiatry* 8, 150 (2018).
476 <https://doi.org/10.1038/s41398-018-0198-3>.
477
- 478 23. Bu S, Lv Y, Liu Y, Qiao S, Wang H. Zinc Finger Proteins in Neuro-Related Diseases
479 Progression. *Front Neurosci.* 2021 Nov 18;15:760567.
480 <https://doi.org/10.3389/fnins.2021.760567>.
481
- 482