LEWIS UNIVERSITY

BEYOND TRIAL AND ERROR: HARNESSING GENETICS FOR TARGETED ANTIDEPRESSANT TREATMENTS

RESEARCH PROJECT SUBMITTED

IN PARTIAL FULFILMMENT OF THE REQUIREMETS

FOR THE DEGREE OF

MASTER OF SCIENCE, DATA SCIENCE

BY SYDNEY WALTER

DIRECTOR: DR. DAWN GRAUNKE

ROMEOVILLE, IL

JUNE 2024

ABSTRACT

Advancements in genetic sequencing have revolutionized the treatment approach for Major Depressive Disorder (MDD), emphasizing personalized medicine guided by genetic profiles. This study utilized patient DNA (deoxyribonucleic acid) sequences to identify significant single nucleotide polymorphisms (SNPs) influencing antidepressant responses and side effects, integrating them with reference sequences to optimize treatment strategies. Genetic profiling of simulated patients revealed associations between specific variants and treatment outcomes, aligning with existing literature. For instance, variants such as rs762551 in Cytochrome P450 1A (*CYP1A2*) were linked to increased fatigue risk and altered medication requirements. From these patient profiles, a predictive algorithm was developed to recommend the most suitable antidepressant based on individual genetic profiles. This study underscores the potential of pharmacogenomics to refine antidepressant therapy through personalized genetic insights, while acknowledging current limitations. Further research encompassing broader genetic landscapes promises to advance precision psychiatry, optimizing therapeutic interventions and patient care outcomes.

TABLE OF CONTENTS

LIST OF TABLES vii
LIST OF ABBREVIATIONS viii
CHAPTER 1 INTRODUCTION 1
Metabolism and Transport of Antidepressants 2
CYP2D6 (Cytochrome 450 2D6) 2
rs1065852 2
CYP1A2 (Cytochrome P450 1A2) 3
rs4646427 3
rs4646425 4
rs2069526 4
rs2470890 5
rs2472304 5
rs762551 5
ABCB1 (ATP Binding Cassette Subfamily B Member 1) 6
rs1045642 6
rs2032583 6
Neurotransmitter Receptors and Transporters 7
HTR2A (5-Hydroxytryptamine Receptor 2A) 7
rs6313 7
rs7997012 9
rs6314 9
rs2770296 9
HTR1D (5-Hydroxytryptamine Receptor 1D) 10
rs6296 10
SLC6A4 (Solute Carrier Family 6 Member 4) 11
rs57098334 11
OPRM1 (Opioid Receptor Mu 1) 12
rs1799971 12
Neuroplasticity and Neurotrophic Factors 12
BDNF (Brain-Derived Neurotrophic Factor) 12
rs962369 12
rs7103411 13
rs7124442 14

GDNF (Glial Cell-Derived Neurotrophic Factor) ... 14
rs2216711 ... 14
rs2973049 ... 15
CREB1 (cAMP Responsive Element Binding Protein 1) ... 15
rs889895 ... 15
Stress Response and Hormonal Regulation ... 16
FKBP5 (FK506 Binding Protein 5) ... 16
rs17614642 ... 16
COMT (Catechol-O-Methyltransferase) ... 17
rs4680 ... 17
MTHFR (Methylenetetrahydrofolate Reductase) ... 19
rs1801131 ... 19
rs1801133 ... 20
Synaptic Function and Signal Transduction ... 20
NRXN1 (Neurexin 1) ... 20
rs4971678 ... 20
GSK3B (Glycogen Synthase Kinase 3 Beta) ... 20
rs334558 ... 20
CACNA1A (Calcium Voltage-Gated Channel Subunit Alpha1 A) ... 21
rs2112460 ... 21
Other Roles in Antidepressant Response ... 22
MC1R (Melanocortin 1 Receptor) ... 22
rs2228478 ... 22
rs1801133 ... 22
CHAPTER 2 METHODOLOGY ... 23
Data Sources ... 23
Sharable Files ... 23
Gene DNA sequences ... 23
Variant Annotation Datasets ... 23
Gene Selection Process ... 24
Bioinformatic Tools ... 25
Processing MDD Variant Annotation Dataset ... 25
Simulated Patient Profiles ... 26
Data Query and Analysis ... 27

Predictive Algorithm 28
CHAPTER 3 RESULTS 30
Selected Genes and Antidepressant Relevance 30
Identified Variants 35
Patient 1 Results 42
Identified Variants in Patient 1 42
Treatment Plan for Patient 1 Based on Genetic Profile 45
Primary Recommendation: Bupropion 45
Secondary Option: Fluvoxamine 45
Medications to Use with Caution 45
General Considerations 46
Patient 2 Results 46
Identified Variants in Patient 2 46
Treatment Plan for Patient 2 Based on Genetic Profile 49
Primary Recommendation: Citalopram or Fluoxetine 49
Secondary Option: Bupropion 49
Medications to Use with Caution 49
General Considerations 50
Patient 3 Results 50
Identified Variants in Patient 3 50
Treatment Plan for Patient 3 Based on Genetic Profile 53
Primary Recommendation: Bupropion 53
Secondary Option: Fluvoxamine 53
Medications to Use with Caution 53
General Considerations 54
Predictive Model Results 54
CHAPTER 4 DISCUSSION 57
Results Linked to Literature Evidence 57
Limitations 59
Extensions 61
Practical Benefits of Precision Medicine 62
Conclusion 63
CHAPTER 5 LITERATURE CITED 65
CHAPTER 6 APPENDIX 69

Google Drive.. 69

LIST OF TABLES

Table 1A. Gene function and relevance to antidepressant response. 32

Table 1B. Gene function and relevance to antidepressant response 33

Table 1C. Gene function and relevance to antidepressant response 34

Table 2. Variants associated with metabolism and transport. 36

Table 3. Variants associated with neurotransmitter receptors and transporters. 37

Table 4. Variants associated with neuroplasticity and neurotrophic factors. 38

Table 5. Variants associated with stress response and hormonal regulation. 39

Table 6. Variants associated with synaptic function and signal transduction. 40

Table 7. Variants with other roles in antidepressant response. 41

Table 8. Identified SNPs in patient 1. 44

Table 9. Identified SNPs in patient 2. 48

Table 10. Identified SNPs in patient 3. 52

Table 11. Predictive Model Classification Report 56

LIST OF ABBREVIATIONS

ABCB1	ATP Binding Cassette Subfamily B Member 1
ADRA2A	Adrenergic Receptor Alpha 2A
API	Application Programming Interface
BDNF	Brain-Derived Neurotrophic Factor
CACNA1A	Calcium Voltage-Gated Channel Subunit Alpha1 A
COMT	Catechol-O-Methyltransferase
CREB	cAMP Responsive Element Binding Protein 1
CRT	Classification-Regression Trees
CYP1A2	Cytochrome P450 1A
CYP2D6	Cytochrome P450 2D6
FDA	Food and Drug Administration
FKBP5	FK506 Binding Protein 5
GDNF	Glial Cell-Derived Neurotrophic Factor
GSK3B	Glycogen Synthase Kinase 3 Beta
HAM-A	Hamilton Anxiety Rating Scale
HAMD	Hamilton Depression Rating Scale
MAO A	Monoamine Oxidase A
MADRS	Montgomery-Asberg Depression Rating Scale
MAPK1	Mitogen Activated Protein Kinase 1
MDD	Major Depressive Disorder
ML	Machine Learning
MTHFR	Methylenetetrahydrofolate Reductase

NESDA	Netherlands Study of Depression and Anxiety
NET G1287A	Norepinephrine Transporter Gene G1287A
NR3C1	Glucocorticoid Receptor Gene
NRXN1	Neurexin 1
NTRK2	Neurotrophic Receptor Tyrosine Kinase 2
NRI	Norepinephrine Reuptake Inhibitor
OPRM1	Opioid Receptor Mu 1
PHQ-9	Patient Health Questionnaire 9
rsIDs	Reference SNP Cluster IDs
SD	Sexual Dysfunction
SE	Side Effects
SERT	Sertraline
SLC6A4	Solute Carrier Family 6 Member 4
SNP	Single Nucleotide Polymorphism
SVM	Support Vector Machines
TESI	Treatment Emergent Suicidal Ideation
UCSC	University of California, Santa Cruz

CHAPTER 1 INTRODUCTION

Major Depressive Disorder (MDD) is a pervasive mental health condition afflicting approximately 6% of the global population, with a lifetime risk estimated at 16.6% (Flint, J., 2023). This debilitating disorder is characterized by persistent feelings of sadness, hopelessness, and a diminished interest in daily activities, significantly contributing to the overall global burden of disease. The challenges faced by individuals diagnosed with MDD are exacerbated by the limited efficacy of current antidepressant medications.

The landscape of existing antidepressant prescription medications is marked by challenges, as observed in the seemingly random and diverse responses experienced by Major Depressive Disorder (MDD) patients. With a substantial proportion, ranging from 30-50%, not responding optimally to conventional antidepressant medications, there is a pressing need for more effective and personalized treatment strategies (Flint, J., 2023). The proposed research aims to address this critical gap in treatment efficacy by unraveling the complexities of antidepressant response.

The primary objective of this project is to systematically annotate carefully selected genes associated with drug responses in the context of antidepressant treatment. By exploring published journal articles, the aim was to identify single nucleotide polymorphisms (SNPs) that have demonstrated associations with the response to treatment with antidepressant medications. This comprehensive investigation will cover various classes of antidepressants, ensuring a thorough understanding of genetic factors influencing treatment outcomes. The genes selected for analysis were chosen based on their relevance to neurotransmitter pathways, drug

metabolism, and any mechanisms impacting drug response. This targeted approach seeked to contribute valuable insights to the field of pharmacogenomics. Subsequently, the project utilized the identified SNPs in the development of a predictive algorithm, recommending personalized antidepressant treatment strategies based on individual genetic profiles. The overarching goal was to enhance treatment efficacy by providing tailored approaches informed by genetic factors.

Metabolism and Transport of Antidepressants

CYP2D6 (Cytochrome 450 2D6)

rs1065852

Cytochrome P450 enzyme activity, influenced by CYP genetic polymorphisms, affects the individual variability in the effectiveness and tolerability of antidepressants for major depressive disorder (MDD). Escitalopram is metabolized by *CYP2D6*, and recent studies have highlighted a correlation between clinical outcomes and *CYP2D6* genetic variants. This study conducted by Han, et al. (2013), aimed to investigate the relationship between the *CYP2D6 P34S* polymorphism (C188T, rs1065852) and the efficacy of escitalopram in Korean patients with MDD. Ninety-four patients diagnosed with MDD participated in this study, with their symptoms addressed using the 21-item Hamilton Depression Rating Scale (HAMD-21). The study examined the association between *CYP2D5* P34S polymorphism and clinical outcomes, such as the remission and response, after 1, 2, 4, 8, and 12 weeks of escitalopram treatment using multiple logistic regression analysis and the $\chi 2$ test. The percentage of P allele carriers (PP, PS) achieving remission was higher compared to S allele homozygotes (SS) after 8 and 12 weeks of escitalopram treatment. Similarly, P allele carriers showed a better treatment response after 8 and 12 weeks of escitalopram treatment than S allele homozygotes. This study formed the clinical variant guideline on Pharmacogenomics Knowledgebase (PharmGKB) stating that genotype *AA*

is associated with decreased response and remission rate when treated with escitalopram in people with Major Depressive Disorder as compared to genotypes *AG* and *GG*.

CYP1A2 (Cytochrome P450 1A2)

rs4646427

Allele *C* in variant rs4646427 is associated with increased metabolism of escitalopram in people with MDD as compared to those with allele *T*. Allele *C* is also associated with increased fatigue, nausea, and vomiting as compared to those with allele *T.* In a study conducted by Hsiang-Wei (2013), researchers explored the association between genetic polymorphisms in the *CYP1A2* gene and treatment response to the antidepressant escitalopram. Involving 158 patients, the research focused in ten single nucleotide polymorphisms (SNPs) in the *CYP1A2* gene, with serum levels of escitalopram and its metabolites measured by high performance liquid chromatography. Results showed the *CYP1A2* SNP rs4646427 variant significantly influenced the metabolic ratios of S-didesmethylcitalopram to S-desmethylcitalopram as week 2 of treatment. Patients with alleles linked to higher S-didesmethylcitalopram to S-desmethylcitalopram ratios experiences more severe side effects, indicating that *CYP1A2* genetic variants can predict escitalopram metabolism and early treatment adverse reactions.

Genotype *TT* is associated with slower response time when treated with paroxetine in people with MDD as compared to those with genotypes *CC* and *CT*. The study conducted by Lin et al. (2010) aimed to investigate whether genetic variations in the *CYP1A2* gene are linked to the efficacy and side effects of paroxetine treatment in patients with major depressive disorder. They recruited 241 MDD patients who had been taking paroxetine for 8 weeks, measuring their steady-state drug concentrations at weeks 2, 4 and 8. Nine selected SNPs from the *CYP1A2* gene were analyzed for associations with treatment outcomes. Results revealed significant associations between certain SNP allele types and paroxetine treatment response, particularly at

week 8. Response rates on depression and anxiety scales were linked to specific SNPs. In summary, the study revealed that *CYP1A2* functions as an additional metabolic enzyme for paroxetine, and proved rs4646427 was linked to a slower treatment response when treated with paroxetine.

rs4646425

The study conducted by Lin et al. (2010) also investigated SNP rs4646425 in the same study and found the presence of the minor *T* allele was associated with a higher proportion of non-remitters compared to remitters after paroxetine treatment at week 8. Additionally, individuals with the *CC* genotype of rs4646425 showed a slower response to treatment compared to other genotypes, as assessed by depression and anxiety rating using the HAM-D and HAM-A (Hamilton anxiety rating scale) scales.

Similar to rs4646427, rs4646425 was also investigated in the study conducted by Hsiang-Wei (2013). It was found that allele *T* was associated with increased metabolism of escitalopram in those with MDD. Therefore, this also led to increased adverse side effects including increased fatigue, nausea, and vomiting.

rs2069526

In rs2069526, allele *G* is associated with increased metabolism of escitalopram in people diagnosed with MDD as compared to allele *T.* This was demonstrated in the study from Hsiang-Wei et al as their findings suggested a modest but significant association with the metabolic ratio of S-desmethylcitalopram (S-DDCIT) to escitalopram (S-CIT) at week 8. Specifically, individuals carrying the G allele of rs2069526 exhibited higher metabolic ratios of S-DDCIT/S-CIT and S-DDCIT/S-didesmethylcitalopram (S-DCIT) compared to those carrying the *T* allele. This indicates that the enzymatic activities of cytochrome P450 1A2 (CYP1A2) may be higher in individuals with the *G* allele. Additionally, the *G* allele carriers had lower serum concentrations

of S-DCIT compared to their counterparts with the *T* allele. These results imply that rs2069526 may play a regulatory role in *CYP1A2* expression and enzymatic activity, particularly influencing S-CIT metabolism.

rs2470890

Allele *T* is associated with increased likelihood of remission when treated with paroxetine in people with MDD compared to allele C. This was proven in the Lin et al.(2010) study where the findings indicated that this SNP, located in exon 7 of the *CYP1A2* gene, showed significant associations with the treatment response to paroxetine in patients with MDD.

rs2472304

Similar to rs2470890, this SNP rs2472304 was also investigated in the Lin et al. (2010) study. It was proven that allele *A* is associated with increased likelihood of remission when treated with paroxetine in people with MDD as compared to those with allele *G*.

rs762551

Allele *A* is associated with increased dose of paroxetine in people with MDD compared to those with allele *C*. It is also associated with increased risk of fatigue when treated with paroxetine. This was supported in the Lin et al (2010) study as it displayed a noteworthy association with the paroxetine dose at week 4. Individuals with specific alleles exhibited a significant difference in paroxetine dosage requirement, indicating a potential role of this genetic variant in determining the appropriate treatment dosage for patients. Moreover, rs762551 demonstrated a significant association with fatigue symptoms as a side effect of paroxetine treatment.

ABCB1 (ATP Binding Cassette Subfamily B Member 1)

rs1045642

Genotype *AA* is associated with increased likelihood of orthostatic hypotension when treated with nortriptyline in people with MDD as compared to genotypes *AG* and *GG*. This was supported in a study conducted by Roberts et al. (2002), but it was later challenged in a study from Jensen et al (2012) who investigated the influence of *ABCB1* polymorphisms on nortriptyline pharmacokinetics and nortriptyline induced postural hypotension in healthy volunteers. The association between *ABCB1* polymorphisms and nortriptyline induced postural hypotension found in the previous study could not be confirmed. The result raised the possibility of a predisposition in heart rate response in the *TT* homozygotes rather than an effect of nortriptyline.

rs2032583

Allele *G* is associated with increased likelihood of adverse effects when treated with citalopram, fluvoxamine, paroxetine, sertraline, or venlafaxine in people with MDD as compared to allele *A*. This was supported by the de Klerk et al (2013) study which aimed to investigate the association between genetic variants in the *ABCB1* gene, which encodes for P-glycoprotein (P-gp) in the blood-brain barrier, and adverse effects of antidepressant drugs, particularly selective serotonin reuptake inhibitors (SSRIs). The researchers utilized data from the Netherlands Study of Depression and Anxiety (NESDA) study, involving 424 patients with major depressive disorder. Six *ABCB1* gene variants (1236T4C, 2677G4T/A, 3435T4C, rs2032583, rs2235040, and rs2235015) were selected, and their association with adverse drug effects was analyzed using multinomial regression analysis for both single variants and haplotypes. The results revealed a significant association between the number of adverse effects related to SSRIs and two variants, rs2032583 ($P = 0.001$) and rs2235040 ($P = 0.002$), as well as a haplotype ($P = 0.002$).

Specifically, serotonergic effects such as sleeplessness, gastrointestinal complaints, and sexual effects were significantly predicted by these variants and haplotype (P = 0.002/0.003). Therefore, the study concludes that genetic variants in *ABCB1* may contribute to adverse drug effects associated with SSRIs.

Neurotransmitter Receptors and Transporters

HTR2A (5-Hydroxytryptamine Receptor 2A)

rs6313

There was study conducted which aimed to investigate the relationship between several genetic polymorphisms, including the serotonin-2A receptor (*HTR2A*) gene -1438A/G and 102T/C polymorphisms, serotonin transporter gene (*SLC6A4*) 5-HTT-linked polymorphic region (5-HTTLPR) insertion/deletion variant, brain-derived neurotrophic factor (*BDNF*) gene Val66Met polymorphisms, and the occurrence of sexual dysfunction (SD) as an adverse effect of citalopram (CIT) or sertraline (SERT) treatment in patients with major depressive disorder (Oz et al., 2019). This study proved that those with genotype *AA* is associated with increased likelihood of psychological sexual dysfunction due to citalopram in people with MDD compared to genotypes *AG* and *GG*.

The meta-analysis study examined the relationship between hydroxytryptamine receptor 2A (*HTR2A*) gene polymorphisms (1438A/G, 102T/C, and rs7997012G/A) and the safety and efficacy of antidepressants in depression patients (Wan et al., 2020). Forty-two studies published before February 2020 were included in the analysis. The pooled analyses revealed significant associations between certain *HTR2A* gene polymorphisms and treatment outcomes. Specifically, the 1438A/G polymorphism was associated with a higher response rate to antidepressants under the dominant model. The rs7997012G/A polymorphism was linked to higher rates of remission,

with significant effects observed in various genetic models (dominant, recessive, and homozygote). Additionally, the 102T/C polymorphism was associated with a reduced risk of side effects under the recessive and homozygote models. This study was used to prove genotypes *AA* and *AG* is associated with increased response to antidepressants in people with MDD as compared to genotype *GG*. Additionally, it also supported the claim that genotypes *AA* and *AG* is associated with decreased risk of adverse events when treated with antidepressants compared to genotype *GG*.

Genotype *GG* is associated with increased risk of discontinuations due to adverse events and greater severity of side effects when treated with paroxetine. In a particular study, the authors aimed to identify genetic markers for antidepressant medication intolerance, including *HTR2A* rs6313 (Murphy et al., 2003). In a double-blind, randomized pharmacogenetic study, 246 elderly patients with major depression were treated with either paroxetine (a selective serotonin reuptake inhibitor) or mirtazapine (not an SSRI) for 8 weeks. Genotypes were determined for the 102 T/C single nucleotide polymorphism (SNP) in the serotonin 2A (5-HT(2A)) locus (*HTR2A*), which had been previously associated with psychotropic medication treatment outcomes. Clinical outcomes assessed included treatment discontinuations, adverse events, medication compliance, and changes in mood. The results of the study revealed that discontinuations resulting from paroxetine-induced side effects were notably linked with the *HTR2A C/C* genotype. Moreover, a significant linear relationship was observed between the number of *C* alleles and the likelihood of discontinuation. Additionally, paroxetine-treated patients with the *C/C* genotype experienced greater severity of side effects.

rs7997012

Genotypes *AG* and *GG* is associated with increased response to antidepressants in people with MDD. According to the meta-analysis study conducted by Lin, J-Y et al. (2014), the patients in the study who had these genotypes had an increased response to antidepressants.

rs6314

In a study, researchers examined the role of genetic variations in the serotonin-2A-receptor gene (5-*HT2A*) and the serotonin transporter gene (5-*HTTLPR*) in major depression (Minov et al., 2001). They analyzed two polymorphisms in the 5-HT2A gene (T102C and His452Tyr) and the insertion/deletion polymorphism in the promoter region of the 5-*HTTLPR* gene in a sample of 173 patients with major depression and 121 healthy controls. The study found no statistically significant differences in the distribution of genotypes or allele frequencies between patients and controls for any of the three investigated polymorphisms. However, there was a trend towards higher frequency of *S/S* (short allele) homozygotes in patients compared to controls for the 5-*HTTLPR* polymorphism, although this did not reach statistical significance. Interestingly, the study observed a different treatment response in patients with one or two *C*-alleles of the T102C polymorphism in the 5-*HT2A* gene, with a significantly higher decrease in Hamilton Depression Rating Scale (HAMD-17) scores after 4 weeks of antidepressant treatment. PharmGKB rewrote this association indicating that genotypes *AG* and *GG* are associated with increased response to antidepressants in people with MMD as compared to those with genotype *AA*.

rs2770296

Genotype *CC* is associated with increased response to bupropion in people with MDD ad compared to genotypes *CT* and *TT*. In this study, researchers investigated the genetic factors influencing response and remission to bupropion treatment in patients with major depressive

disorder (Tiwari et al., 2013). They examined 532 tagging single nucleotide polymorphisms (SNPs) in 34 candidate genes and analyzed their association with treatment outcomes in patients receiving either bupropion or placebo. The analysis revealed significant associations with remission following bupropion treatment for a SNP within the serotonin receptor 2A gene (*HTR2A* rs2770296) and response to bupropion treatment for a SNP in the dopamine transporter gene (rs6347). Additionally, among patients receiving placebo, marginal associations for remission were observed with SNPs in *HTR2A* (rs2296972) and the serotonin transporter gene (*5-HTT* or *SLC6A4* rs4251417). Placebo response was associated with SNPs in the glucocorticoid receptor gene (*NR3C1* rs1048261) and monoamine oxidase A gene (*MAOA* rs6609257). Although these associations were significant after gene-wide corrections, they would not remain significant after a more conservative study-wide correction for multiple tests. Nonetheless, the results suggest a potential role for *HTR2A* in remission to bupropion treatment, dopamine transporter in bupropion response, and *MAOA* gene in placebo response. These findings provide insights into the genetic factors underlying treatment response in major depressive disorder, particularly concerning bupropion pharmacology and placebo response mechanisms.

HTR1D (5-Hydroxytryptamine Receptor 1D)

rs6296

Genotype *CC* is associated with increased risk of agitation when treated with citalopram in children with MDD as compared to genotype *CG* and *GG*. The study from Amitai et al., (2016) aimed to investigate the association between genetic variations in genes related to the serotonergic system and citalopram side effects (SEs) in children and adolescents with major depressive disorder (MDD)/dysthymia and/or anxiety disorders. A total of 87 outpatients aged 7-

18 years received citalopram treatment in an 8-week open trial, and SEs were assessed using a specific questionnaire.

The study found that agitation was more prevalent in boys compared to girls. Additionally, subjects with the 5-*HTR1Dbeta CC* genotype exhibited higher rates of agitation compared to those with *CG* and *GG* genotypes. Specifically, individuals with the *CC* genotype reported more agitation symptoms. The study suggests that agitation may serve as an intermediate phenotype to suicidal behavior, indicating a potential role of the 5-*HTR1Dbeta* polymorphism in citalopram-related agitation among children and adolescents treated for depression and/or anxiety.

SLC6A4 (Solute Carrier Family 6 Member 4)

rs57098334

In a study involving 241 Korean patients with major depression, the association between genetic polymorphisms in serotonin and norepinephrine transporter genes and antidepressant response was investigated (Kim et al., 2006). Patients were treated with either a selective serotonin reuptake inhibitor (SSRI) or a norepinephrine reuptake inhibitor (NRI) for six weeks. Results showed that the response to NRIs was linked to a specific polymorphism in the norepinephrine transporter gene (*NET G1287A*), while SSRI response correlated with variations in the serotonin transporter gene (5-*HTT* intron 2 s/l variation and 5-*HTTLPR*). Interestingly, the favorable allele for SSRI response was different from that observed in white patients. Additionally, certain genotype combinations were associated with high response rates, while others correlated with low response rates. These findings suggest the potential for pharmacogenetic selection of antidepressant treatment based on monoamine transporter gene polymorphisms. It found that genotype (*CCCACCCGA)12/(CCCACCCGA)12* is associated with increased response to fluoxetine or sertraline in people with MDD.

OPRM1 (Opioid Receptor Mu 1)

rs1799971

Genotype *AA* is associated with increased risk of suicidal ideation due to tianeptine in people with Depressive Disorder, Major as compared to genotypes *AG* and *GG*. In a study involving 3566 adult depressed outpatients in France, researchers investigated the relationship between genetic polymorphisms in the opioidergic system and Treatment Emergent Suicidal Ideation (TESI) following the initiation of antidepressant treatment with tianeptine (Nobile et al., 2019). They focused on two specific genetic polymorphisms (SNPs) in the opioidergic system, rs1799971 from the *OPRM1* gene and rs105660 from the OPRK1 gene. Suicidal ideation was assessed at baseline and at 2, 4, and 6 weeks using the Montgomery-Asberg Depression Rating Scale. The study found a significant association between the *AA* genotype of rs1799971 and TESI, even after adjusting for potential confounders. However, there were no significant associations between the studied SNPs and worsening of suicidal ideation or lifetime suicide attempts. These findings suggest a potential involvement of the opioidergic system in TESI.

Neuroplasticity and Neurotrophic Factors

BDNF (Brain-Derived Neurotrophic Factor)

rs962369

Allele *C* is associated with increased likelihood of suicidal ideation when treated with escitalopram or nortriptyline in people with Depressive Disorder, Major as compared to allele *T*. In a study involving 796 adult patients with major depressive disorder treated with escitalopram or nortriptyline, researchers aimed to identify genetic predictors of increased suicidal ideation during antidepressant treatment (Perroud et al., 2009). They selected nine candidate genes associated with neurotrophic, serotonergic, and noradrenergic pathways. Comparing 123 polymorphisms in these genes between patients with and without increased suicidal ideation,

they found significant associations with polymorphisms in the *BDNF* gene, particularly rs962369. Additionally, an interaction between variants in *BDNF* and *NTRK2* genes was observed. Among men taking nortriptyline, suicidality was associated with a SNP in the *ADRA2A* gene. These findings suggest involvement of the neurotrophic system and potential interactions between *BDNF*, *NTRK2* (Neurotrophic Receptor Tyrosine Kinase 2), and *ADRA2A* (Adrenergic Receptor Alpha 2A) genes in vulnerability to suicidality during antidepressant treatment.

rs7103411

Genotype *TT* is associated with decreased response to citalopram in people with Depressive Disorder, Major as compared to genotype *CC*. This study investigated the association of three *BDNF* gene polymorphisms (rs7103411, Val66Met [Valine to Methionine substitution at codon 66 of the *BDNF* gene; rs6265], and rs7124442) with major depression and antidepressant treatment response in a sample of 268 German patients with major depression and 424 healthy controls (Domschke et al., 2009). Additionally, ten *BDNF* markers were tested for association with citalopram outcome in the STAR*D (Sequenced Treatment Alternatives to Relieve Depression) sample. Key findings include that *BDNF* was not associated with major depression as a categorical diagnosis. However, the *BDNF* rs7124442 TT genotype was significantly related to worse treatment outcomes over 6 weeks in major depression ($p = 0.01$), particularly in anxious depression ($p = 0.003$). The *BDNF* rs7103411 and rs6265 polymorphisms similarly predicted worse treatment responses over 6 weeks in clinical subtypes of depression such as melancholic depression (rs7103411: *TT* < *CC*, $p = 0.003$; rs6265: *GG* < *AA*, $p = 0.001$). All SNPs had main effects on antidepressant treatment response in ANOVA models when considering other SNPs as covariates. The STAR*D analyses did not yield significant results for any of the ten *BDNF*

markers. Post-hoc analyses suggest a potential minor role for *BDNF* genetic variation in antidepressant treatment outcomes, particularly in the context of melancholic depression.

rs7124442

Genotype *TT* is associated with decreased response to citalopram in people with Depressive Disorder, Major as compared to genotypes *CC* and *CT*. This result came from the same study as rs7103411 (Domschke et al., 2009).

GDNF (Glial Cell-Derived Neurotrophic Factor)

rs2216711

Genotypes *AA* and *AG* is associated with decreased response to paroxetine in women with Depressive Disorder, Major as compared to genotype *GG*. This study aimed to evaluate the impact of specific genetic polymorphisms in the brain-derived neurotrophic factor (*BDNF*) and glial cell line-derived neurotrophic factor (*GDNF*) genes on the efficacy of paroxetine in treating major depressive disorder (MDD) (Wang et al., 2014). The investigation involved 298 patients with MDD who began treatment with 20 mg of paroxetine daily. After 6 weeks, plasma concentrations of paroxetine were measured, and changes in depression severity were assessed using the Hamilton Depression Rating Scale (HAM-D). Genotyping for the *BDNF* and *GDNF* polymorphisms was performed using the Sequenom MassArray system. At the 6-week follow-up, 73.5% of the patients responded to the treatment, while 26.5% did not. A lower threshold concentration of 50 ng/mL for paroxetine was identified for treatment response, with a linear relationship observed between paroxetine plasma concentration and clinical response. Significant associations were found between paroxetine treatment remission and specific SNPs, including rs6265 ($P < 0.001$), rs2973049 ($P = 0.005$), and rs2216711 ($P = 0.006$). The results suggest that genetic variants in the *BDNF* and *GDNF* regions, particularly the SNPs rs2973049 and

rs2216711 in the *GDNF* gene, may serve as indicators of treatment response to paroxetine in patients with MDD.

rs2973049

Genotypes *CT* and *TT* is associated with decreased response to paroxetine in people with Depressive Disorder, Major as compared to genotype *CC*. In this study, researchers investigated the impact of genetic variations, specifically the single nucleotide polymorphism (SNP) rs2973049, in the glial cell line-derived neurotrophic factor (GDNF) gene on the efficacy of paroxetine in treating major depressive disorder (MDD) (Wang et al., 2014). They conducted genotyping for rs2973049 and other polymorphisms in the brain-derived neurotrophic factor (*BDNF*) and *GDNF* genes in 298 MDD patients who initiated paroxetine treatment. Plasma concentrations of paroxetine were measured after 6 weeks, and changes in MDD severity were assessed using the Hamilton Depression Rating Scale (HAM-D). Results showed that 73.5% of the patients responded to paroxetine treatment, while 26.5% did not. A lower threshold concentration of 50 ng/mL for paroxetine was identified for treatment response, and a linear relationship was observed between paroxetine plasma concentration and clinical response. Notably, the allele types for rs2973049 demonstrated a significant association with paroxetine treatment remission at week 6 ($P = 0.005$), indicating that this genetic variant may be a potential indicator of treatment response to paroxetine in MDD patients.

CREB1 (cAMP Responsive Element Binding Protein 1)

rs889895

Genotype *GG* is associated with increased response to antidepressants in people with Depressive Disorder, Major as compared to genotypes *AA* and *AG*. This study aimed to explore the association between single nucleotide polymorphisms (SNPs) within the mitogen activated protein kinase 1 (*MAPK1*) and cyclic adenosine monophosphate (cAMP) responsive element

binding protein 1 (*CREB1*) genes and antidepressant treatment resistance in Major Depressive Disorder (MDD) patients (Calati et al., 2013). Among the investigated SNPs, rs889895 in the *CREB1* gene was assessed for its role in treatment resistance and other clinical features in MDD patients, along with a larger sample including bipolar disorder (BD) patients. The study found no significant association between rs889895 and treatment resistance or response in MDD patients. However, higher rates of the *GG* genotype of rs889895 were reported in MDD patients who achieved remission. Additionally, the rs889895 *GG* genotype in *CREB1* and rs8136867 *AG* genotype in *MAPK1* were associated with remission in the combined MDD and BD sample. These findings suggest a potential involvement of genetic polymorphisms in *CREB1* and *MAPK1* in treatment remission, highlighting their role in neuroplasticity and inflammatory processes in response to antidepressant treatment. Further research is warranted to confirm and expand upon these observations.

Stress Response and Hormonal Regulation

FKBP5 (FK506 Binding Protein 5)

rs17614642

Genotype *TT* is associated with decreased response to bupropion in people with Depressive Disorder, Major as compared to genotype *CT*. In a study investigating the response to bupropion and placebo in patients with major depressive disorder, researchers examined 532 single nucleotide polymorphisms (SNPs) in 34 candidate genes (Tiwari et al., 2013). They found a significant association between remission following bupropion treatment and a SNP in the serotonin receptor 2A gene (*HTR2A* rs2770296, p(corrected)=0.02). Response to bupropion was significantly associated with a SNP in the dopamine transporter gene (rs6347, p(corrected)=0.013). For patients receiving placebo, marginal associations for remission were

observed with SNPs in *HTR2A* (rs2296972, p(corrected)=0.055) and the serotonin transporter gene (5-*HTT* or *SLC6A4* rs4251417, p(corrected)=0.050). Additionally, placebo response was associated with SNPs in the glucocorticoid receptor gene (*NR3C1*; rs1048261, p(corrected)=0.040) and monoamine oxidase A gene (*MAOA*; rs6609257, p(corrected)=0.046). These findings suggest a potential role for *HTR2A* in remission to bupropion treatment, dopamine transporter in bupropion response, and *MAOA* gene in placebo response. However, further studies with more conservative corrections are needed to confirm these associations.

COMT (Catechol-O-Methyltransferase)

rs4680

Genotype *GG* is associated with decreased response to fluvoxamine in people with Depressive Disorder, Major as compared to allele *A*. In a study investigating the response to fluvoxamine antidepressant monotherapy, researchers examined the effect of the catechol-O-methyltransferase (*COMT*) Val(108/158)Met polymorphism (rs4680) (Benedetti et al., 2008). Forty-one inpatients with major depressive disorder were administered fluvoxamine for 6 weeks, and changes in depression severity were assessed weekly. The study found that rs4680 significantly interacted with time in affecting antidepressant response, with better effects observed in Met-carriers and worse effects in Val/Val homozygotes. This effect became significant by the fourth week of treatment and influenced final response rates. The findings suggest that rs4680 may play a role in shaping individual responses to fluvoxamine, indicating that factors affecting catecholaminergic neurotransmission could contribute to antidepressant response irrespective of their primary molecular target.

Allele *A* is associated with increased response to paroxetine in people with Depressive Disorder, Major as compared to allele *G*. In a study investigating the response to paroxetine antidepressant monotherapy, researchers examined the effect of the *COMT* Val (108/158) Met

polymorphism (rs4680) (Benedetti et al., 2008). Fifty-five outpatients with major depressive disorder were administered paroxetine for one month, and changes in depression severity were assessed weekly. The study found that rs4680 significantly interacted with time in affecting antidepressant response to paroxetine, with better effects observed in Met/Met homozygotes, worse effects in Val/Val homozygotes, and intermediate effects in heterozygotes. This effect became significant by the third week of treatment. The findings suggest that rs4680 may play a role in shaping individual responses to paroxetine, indicating that factors affecting catecholaminergic neurotransmission could contribute to antidepressant response.

Genotypes *AG* and *GG* is associated with increased response to bupropion in people with Depressive Disorder, Major as compared to genotype *AA*. In a study investigating the response to bupropion in major depressive disorder (MDD) patients, researchers examined the influence of *COMT* genotypes (Met/Met, Met/Val, and Val/Val) on depression scores (Fawver et al., 2020). A sample of 241 adult outpatients diagnosed with MDD completed genetic testing and the Patient Health Questionnaire (PHQ-9) at clinic visits over a period averaging 3.8 visits. Participants were prescribed bupropion or another antidepressant and remained adherent to treatment for over two months post-genetic testing. Results showed that for Val carriers, high doses of bupropion led to significantly lower PHQ-9 scores compared to no bupropion or low doses. Val carriers also responded differently to high-dose bupropion compared to Met/Met patients. These findings suggest that high-dose bupropion may be beneficial for MDD patients with Met/Val or Val/Val *COMT* genotypes but not for those with the Met/Met genotype. Prospective studies are needed to validate these findings and explore their clinical implications further.

Genotype *GG* is associated with increased response to venlafaxine in people with Depressive Disorder, Major as compared to genotypes *AA* and *AG*. In a randomized, double-blind, placebo-controlled clinical trial for major depressive disorder, responses to venlafaxine treatment were stratified by *COMT* genotypes, specifically the Val158Met polymorphism (rs4680) (Hopkins et al., 2013). The study found that improvements in depression scores were larger among subjects with Val/Val genotypes compared to those with Met/Met genotypes. This suggests that venlafaxine may have differential effects on noradrenergic flux based on *COMT* activity levels.

MTHFR (Methylenetetrahydrofolate Reductase)

rs1801131

Allele *G* is associated with increased response to Vitamin B-complex, Incl. Combinations in people with Depressive Disorder, Major. This study aimed to assess the efficacy and safety of reduced B vitamins as monotherapy in adults with major depressive disorder (MDD) who were positive for at least one methylenetetrahydrofolate reductase (*MTHFR*) polymorphism associated with depression (Mech & Farah, 2016). A total of 330 adult MDD patients positive for *MTHFR* C677T or A1298C polymorphism were enrolled, with 170 receiving reduced B vitamins and 160 receiving placeboes. Results showed that the active treatment group exhibited a significant reduction in homocysteine levels and demonstrated clinical improvement as measured by the Montgomery-Asberg Depression Rating Scale (MADRS), with 42% achieving remission by week 8. No significant side effects were observed, and there were no instances of mania. These findings support the therapeutic benefit of reduced B vitamins as monotherapy for MDD, particularly in patients with MTHFR polymorphism.

rs1801133

Allele *A* is associated with increased response to Vitamin B-complex, Incl. Combinations in people with Depressive Disorder, Major. This was found from Mech & Farah (2016) study.

Synaptic Function and Signal Transduction

NRXN1 (Neurexin 1)

rs4971678

Allele *T* is associated with decreased response to duloxetine in people with Depressive Disorder, Major as compared to allele *A*. In this study, machine learning (ML) models were used to predict treatment outcomes in major depressive disorder (MDD) patients receiving duloxetine (Maciukiewicz et al., 2018). A sample of 186 patients was categorized as "responders" or "remitters" based on MADRS scores. Genome-wide logistic regression and LASSO (Least Absolute Shrinkage and Selection Operator) regression were used to identify significant genetic variants related to duloxetine response/remission. Subsequently, classification-regression trees (CRT) and support vector machines (SVM) were applied to construct models. For response, none of the models performed significantly better than chance. However, for remission, SVM achieved moderate performance with an accuracy of 0.52, a sensitivity of 0.58, and a specificity of 0.46. The best performing SVM fold had an accuracy of 0.66, a sensitivity of 0.70, and a specificity of 0.61. While the models showed promising sensitivity, specificity remained moderate. This study provided proof that those with allele T in variant rs4971678 had a decreased response to duloxetine.

GSK3B (Glycogen Synthase Kinase 3 Beta)

rs334558

Genotype *GG* is associated with increased response to citalopram and fluoxetine in people with Depressive Disorder, Major as compared to allele *A*. The study investigated the

association between genetic variants in the *GSK3B* gene and major depressive disorder (MDD) as well as the therapeutic response to antidepressants (Tsai et al., 2008). Four polymorphisms of the *GSK3B* gene were genotyped in Chinese MDD patients and controls. While no significant association was found between these genetic variants and MDD susceptibility, three of the four polymorphisms were significantly associated with the therapeutic effect of antidepressants at 4 weeks. Specifically, rs334558 showed a significant association with the antidepressant therapeutic effect. The findings suggest that *GSK3B* genetic variants may influence the response to selective serotonin reuptake inhibitor (SSRI) antidepressant treatment, highlighting a potential role for drugs targeting GSK3B activity in MDD treatment strategies.

CACNA1A (Calcium Voltage-Gated Channel Subunit Alpha1 A)

rs2112460

Allele A is associated with increased response to antidepressants in people with Depressive Disorder, Major as compared to allele G. The study aimed to investigate the genetic factors associated with antidepressant response in Korean patients with major depressive disorder (MDD) using a genome-wide approach (Cocchi et al., 2016). A sample of 109 MDD patients of Korean origin undergoing antidepressant treatment was collected and phenotypes were assessed based on response and remission criteria. Genome-wide genotyping was performed, and analyses were conducted at the SNP, gene, and pathway levels. Among the replicated genes across samples, *CTNNA3* (Catenin Alpha 3) showed promise. Additionally, the inorganic cation transmembrane transporter activity pathway was associated with antidepressant response in both the Korean sample and an independent replication sample. This pathway included genes such as *CACNA1A*, *CACNA1C*, and *CACNB2*. The findings suggest the involvement of genes encoding subunits of L-type voltage-gated calcium channels in

antidepressant efficacy across different ethnicities, although replication of these findings is needed for confirmation.

Other Roles in Antidepressant Response

MC1R (Melanocortin 1 Receptor)

rs2228478

Allele *G* is associated with increased likelihood of remission when treated with desipramine in people with Depressive Disorder, Major as compared to allele *A*. The study investigated the association between the melanocortin 1 receptor (*MC1R*) gene and susceptibility to major depressive disorder (MDD) and response to antidepressant treatment (Wu et al., 2011). Among the identified single nucleotide polymorphisms (SNPs) within the MC1R gene, rs2228478 was found to be associated with remission with desipramine treatment. However, no associations were found for remission with fluoxetine treatment or with the combined sample treated with fluoxetine or desipramine. The frequency of one of the identified haplotypes was higher in depressed patients compared to controls. In-silico functional analysis indicated that rs2228478, along with other SNPs, had a significant impact on protein function. Overall, the findings suggest a potential association between the *MC1R* gene and both MDD susceptibility and treatment response to desipramine.

rs1801133

Allele A is associated with increased response to Vitamin B-complex, Incl. Combinations in people with Depressive Disorder, Major. This association was supported by the study mentioned previously with Vitamin-B complex that was conducted by Mech & Farah (2016).

CHAPTER 2 METHODOLOGY

Data Sources

Sharable Files

Any and all of the datasets, files, scripts, and sequences can be downloaded from the Google Drive link in the Appendix.

Gene DNA sequences

The genomic sequences were retrieved from Ensembl in FASTA format using a Jupyter Notebook. This process involved accessing the Ensembl REST API (Application Programming Interface), which provides programmatic access to a wide array of genomic data and associated resources. The ‘requests’ library was used to interact with the API, while ‘Bio’ from Biopython was used for sequence manipulation and processing (Reitz, 2023; Cock et al., 2009).

Variant Annotation Datasets

Clinical variant data were obtained from PharmGKB via their downloads page. This file includes a comprehensive list of variant-drug pairs along with their levels of evidence for all clinical annotations.

Additionally, another dataset was sourced by searching for 'Depressive Disorder, Major' on PharmGKB. Upon selecting 'Variant Annotations' in the sidebar, a list of all relevant annotations was displayed, and the corresponding table was downloaded. This dataset comprised 801 annotations. The dataset features included variant, literature, genes, association, significance, p-value, number of cases, number of controls, biogeographical groups, phenotype categories, pediatric information, additional details, and drugs. The processing and analysis of these datasets

were performed using the ‘pandas’ library (McKinney, 2010).

Gene Selection Process

Upon downloading the clinical variants data, it was uploaded to Jupyter Notebook for cleaning and processing. The cleaning process involved splitting the 'variant' column to ensure each variant belonged to its own instance and filtering the dataset by phenotype to retain only instances related to Major Depressive Disorder, the condition of focus for this study. Level 1A evidence was chosen for its clinically proven correlation between drug and variant, offering the most reliable data for our analysis.

The levels of evidence were categorized as 1A, 1B, 2A, 2B, 3, and 4. According to PharmGKB, 1A evidence pertains to variant-drug combinations with specific prescribing recommendations outlined in current clinical guidelines or FDA (Food and Drug Administration)-approved drug labels. These annotations must provide guidance for particular variants (e.g., *CYP2C9*3, *HLA-B*57:*01*) or establish a link between defined allele functions, diplotypes, and phenotypes, supported by at least one publication and clinical guidelines or drug labels offering variant-specific prescribing guidance. Filtering the dataset to include only 1A evidence variants resulted in 72 unique genes of interest. A list of corresponding variants was generated, with variants further separated into rsIDs and haplotypes.

The decision to focus on rsIDs over haplotypes was driven by the specific requirements of the UCSC Table Browser, which only accepts rsIDs (Reference SNP cluster IDs). rsIDs are unique identifiers for single nucleotide polymorphisms (SNPs), representing single point mutations in the genome. This specificity made them suitable for precise identification and

analysis. In contrast, haplotypes represent combinations of alleles or SNPs inherited together, covering a broader genomic region. Handling haplotypes required different approaches due to their composite nature. Therefore, the haplotypes were removed from the dataset, ensuring compatibility with the tools used in this study. The rsIDs list was saved as a .txt file to facilitate its subsequent upload to the UCSC Table Browser.

Bioinformatic Tools

The rsIDs .txt file was uploaded to the UCSC Table Browser, a bioinformatic tool that enables the retrieval of chromosomal data. By inputting the rsIDs, the Table Browser outputted the chromosomal start position, end position, and chromosome number of the SNPs. This output was saved as a .txt file to allow for seamless merging with the dataset in subsequent steps. Integrating the chromosomal location data with the variant annotations was crucial for accurately identifying the exact location of the SNPs within the genome, which is essential for downstream analysis and applications.

Processing MDD Variant Annotation Dataset

The Major Depressive Disorder Variant Annotations dataset was uploaded to Jupyter Notebook for further processing. It was filtered to retain only instances containing variants from the rsIDs .txt file, specifically focusing on 1A evidence variants. Since additional associations for these variants were present, the dataset was filtered to include only the significant instances. It was further refined by keeping only relevant columns necessary for the analysis.

After uploading the rsIDs .txt file to the UCSC (University of California, Santa Cruz) Table Browser, the chromosomal location data—comprising chromosomal start positions, end positions, and chromosome numbers—was retrieved. This data was crucial for accurately pinpointing the genomic locations of the SNPs relevant to our study.

Next, the retrieved chromosomal location data was merged with the variant annotations dataset using the 'Variant' column as the key identifier. This step ensured that each variant had associated chromosomal coordinates, facilitating precise genomic mapping.
Simultaneously, functions were developed to extract gene start and end positions from the downloaded gene FASTA files. These positions were then converted into data frames and integrated into the variant annotations dataset. This integration provided additional columns identifying the genomic coordinates (start and end positions) of each gene associated with the variants under study.

A new column, "SNP Loc on Gene," was created by subtracting the Gene Start Position from the chromStart and adding one, accounting for the gene index starting at zero. Another column, "Gene Length," was computed as the difference between the Gene End Position and Gene Start Position. SNP locations outside the gene's range were excluded to maintain consistency and accuracy.

Simulated Patient Profiles

To simulate patient profiles, a dataframe (snps_df) was created from the "Genes" and "SNP Loc on Gene" columns of the MDD variants dataset. This dataframe was grouped by the "Genes" column, aggregating SNP locations into lists and converting the grouped dataframe into a dictionary.

A list of selected mutations was generated by randomly selecting up to five SNP locations per gene. This number was chosen because single nucleotide polymorphisms (SNPs) are the most abundant genetic variations in the human genome, occurring on average once every 300 base pairs of sequence with a minor allele frequency (MAF) greater than 1% (Kruglyak & Nickerson, 2001; Stephens et al., 2001; Reich et al., 2003). Given the human genome contains

between 4–11.5 million SNPs, roughly one SNP every 1,000 nucleotides, selecting five SNPs provided a reasonable number to work with, reflecting the high prevalence and diversity of SNPs. There could be more SNPs present in a patient, but in this project, only up to five SNPs were selected as they matched the list of SNPs relevant to this study.

The selected gene-SNP pairs were used to modify gene sequences by inserting an 'X' at the mutation positions. This approach was chosen to mitigate logical errors encountered with direct nucleotide modifications. The modified sequences were saved into FASTA files, and their accuracy was validated by aligning and comparing them with reference sequences. A function was developed to replace the placeholder 'X' with a random nucleotide, ensuring accurate representation of mutations. Alignment with the reference sequence allowed for discovering where SNP was located.

Data Query and Analysis

A code was created to query the Major Depressive Disorder variant annotation dataset using specific gene names and SNP locations. Boolean indexing filtered rows based on both gene and SNP location, providing detailed information about the variant, including rsID, gene, association, p-value, significance, phenotype category, drug, chromosome number, chromosomal start and end positions, gene start and end positions, and gene length.

To ensure the reliability of variant-drug associations, the obtained association metrics and p-values were meticulously cross-referenced with existing literature findings. This validation step was crucial in affirming the clinical significance and statistical robustness of the identified relationships.

Subsequently, leveraging the comprehensive analysis of variant-drug associations, a tailored treatment plan was developed. This plan integrated specific prescribing

recommendations based on level 1A evidence from PharmGKB and other authoritative sources, which outline variant-specific drug guidelines. Phenotype categories linked to each variant guided the selection or adjustment of drugs to optimize therapeutic outcomes while minimizing adverse effects. The treatment strategy embraced a personalized medicine approach, customizing drug selections to individual genetic profiles to enhance efficacy and safety.

Predictive Algorithm

Based on the data query and analysis, treatment plans were meticulously crafted for each patient, focusing on their genetic SNPs. A comprehensive dataset was compiled from 30 patient profiles, detailing their patient IDs, genetic variants (variants 1-5), genotypes, and their primary antidepressant choices. This file can be found in the Appendix in the Google Drive link as patient_profiles.csv. To facilitate analysis, this dataset was encoded using Python's LabelEncoder from the sklearn.preprocessing module, ensuring each categorical variable was transformed into numerical equivalents (Pedregosa et al., 2011). This encoding process enabled the integration of new patient data seamlessly.

For the predictive phase, a logistic regression algorithm from the ‘sklearn.linear_model’ module was selected as the optimal model (Pedregosa et al., 2011). This choice was driven by its suitability for predicting categorical outcomes, specifically the primary antidepressant choice based on genetic SNPs. The algorithm was trained on the encoded dataset, learning patterns and associations between genetic variants and antidepressant responses.

When new patient data was introduced, the encoded values of their genetic variants were integrated into the dataset. Leveraging the trained logistic regression model, the algorithm then generated predictions for the best-suited antidepressant for each new patient, leveraging insights gleaned from the genetic profiles and treatment outcomes of the 30 patients. The classification

model and the accuracy score were both calculated using the ‘sklearn.metrics’ module (Pedregosa et al., 2011).

.

CHAPTER 3 RESULTS

Selected Genes and Antidepressant Relevance

The selection of genes for this study was guided by their known variants associated with antidepressant response mechanisms. Each gene was chosen based on its documented impact on drug metabolism, neurotransmitter signaling, or stress response pathways relevant to antidepressant efficacy. Cytochrome P450 2D6 (*CYP2D6*) and Cytochrome P450 1A2 (*CYP1A2*) were selected for their pivotal roles in metabolizing antidepressants, thereby influencing drug clearance rates and efficacy. Genes encoding serotonin receptors, such as 5-Hydroxytryptamine Receptor 2A (*HTR2A*) and 5-Hydroxytryptamine Receptor 1B (*HTR1D*), were included due to their central role in mediating the effects of serotonin-modulating antidepressants. Brain-Derived Neurotrophic Factor (*BDNF*) and FK506 Binding Protein 5 (*FKBP5*) were chosen for their influence on neuroplasticity and stress response, factors known to affect antidepressant treatment outcomes. Solute Carrier Family 6 Member 4 (*SLC6A4*), which encodes the serotonin transporter crucial for selective serotonin reuptake inhibitors (SSRIs), was also included based on its impact on serotonin reuptake efficiency.

The selection of these genes was further supported by extensive research validating their relevance in psychiatric pharmacogenomics. Studies have consistently linked variants in *CYP2D6* and *CYP1A2* to variations in antidepressant metabolism and clinical response profiles across different patient populations. Similarly, genetic variations in *HTR2A* and *HTR1D* have been associated with differential responses to serotonergic medications, highlighting their

significance in treatment outcomes. *BDNF* and *FKBP5* variants have been extensively studied in relation to stress response and neuroplasticity, influencing individual susceptibility to antidepressant therapies. Additionally, *SLC6A4* variants, including the well-known 5-*HTTLPR* polymorphism, have been implicated in altering serotonin transporter function, thereby affecting antidepressant efficacy and tolerability. Table 1 provides a detailed summary of these selected genes, presenting their gene symbols, names, functions, and established roles in antidepressant response mechanisms.

Table 1A. Gene function and relevance to antidepressant response.

Gene Symbol	Gene Name	Function and Relevance to Antidepressant Response
CYP2D6	Cytochrome P450 2D6	Metabolizes many antidepressants, such as SSRIs and TCAs, affecting drug levels and efficacy.
HTR2A	5-Hydroxytryptamine Receptor 2A	Encodes a serotonin receptor, targeted by many antidepressants. Variants can influence response to SSRIs and serotonergic drugs.
BDNF	Brain-Derived Neurotrophic Factor	Plays a crucial role in neuroplasticity, learning, and memory. Its variations may affect the efficacy of antidepressants, particularly SSRIs.
SLC6A4	Solute Carrier Family 6 Member 4	Encodes the serotonin transporter (SERT), targeted by SSRIs. Variants, like 5-HTTLPR, can affect response to antidepressants by altering serotonin reuptake efficiency.
ABCB1	ATP Binding Cassette Subfamily B Member 1	Encodes P-glycoprotein, affecting blood-brain barrier permeability. Genetic variations can influence drug levels in the brain and therapeutic response, including to antidepressants.
OPRM1	Opioid Receptor Mu 1	Involved in pain modulation and reward pathways. Variants may influence antidepressant response and side effects, especially with drugs affecting opioid pathways.
HTR1D	5-Hydroxytryptamine Receptor 1B	Encodes a serotonin receptor subtype implicated in mood regulation. Variants can affect serotonin signaling and response to serotonergic antidepressants.

Table 1B. Gene function and relevance to antidepressant response.

CYP1A2	Cytochrome P450 1A2	Metabolizes various drugs, including some antidepressants. Genetic variations can influence drug clearance rates and efficacy, similar to *CYP2D6*.
COMT	Catechol-O-Methyltransferase	Degrades catecholamines, including dopamine, affecting neurotransmitter levels. Variants may affect efficacy of antidepressants targeting these systems.
FKBP5	FK506 Binding Protein 5	Regulates the stress hormone system. Variants may affect HPA axis response to stress and influence efficacy of antidepressants, particularly those targeting stress response.
GDNF	Glial Cell-Derived Neurotrophic Factor	Supports survival of dopaminergic neurons. Variations may influence neuronal plasticity and response to certain antidepressants.
NRXN1	Neurexin 1	Involved in synaptic formation and function. Variants may impact synaptic signaling and efficacy of antidepressants.
MC1R	Melanocortin 1 Receptor	Involved in pigmentation and potentially pain and stress response pathways. Variations could influence mood and response to antidepressants.
GSK3B	Glycogen Synthase Kinase 3 Beta	Involved in neuroplasticity and inflammation. Genetic variations may affect response to mood stabilizers and antidepressants targeting *GSK3B* inhibition.
MTHFR	Methylenetetrahydrofolate Reductase	Involved in folate metabolism and homocysteine regulation. Variants may influence folate levels linked to mood regulation and efficacy of certain antidepressants.

Table 1C. Gene function and relevance to antidepressant response.

CREB1	cAMP Responsive Element Binding Protein 1	Involved in transcription of genes associated with neuroplasticity and cellular resilience. Variants may influence effectiveness of antidepressants modulating intracellular signaling pathways.
CACNA1A	Calcium Voltage-Gated Channel Subunit Alpha1 A	Encodes a subunit of voltage-dependent calcium channel, involved in neurotransmitter release and neuronal excitability. Variants may affect neuronal activity and response to antidepressants.

Identified Variants

These selected variants across Tables 2 to 7 were chosen based on their documented associations with antidepressant response in Major Depressive Disorder (MDD). They represented genetic factors influencing metabolism and transport, neurotransmitter receptors and transporters, neuroplasticity and neurotrophic factors, stress response, hormonal regulation, synaptic function, and signal transduction. These associations were critical for studying personalized treatment plans aimed at optimizing antidepressant efficacy and tailoring therapies to individual genetic profiles in patients with MDD. Understanding these genetic variations could aid in predicting treatment outcomes and potentially guide clinical decisions for more effective management of depression.

Table 2. Variants associated with metabolism and transport.

Gene	Variant	Antidepressant Relevance
CYP2D6	rs1065852	Influences the metabolism of antidepressants like escitalopram, with P allele carriers showing higher remission rates in MDD patients (Han et al., 2013).
CYP1A2	rs4646427	Variant *C* is associated with increased escitalopram metabolism and potential adverse effects (Hsiang-Wei et al., 2013).
CYP1A2	rs4646425	*T* allele linked to slower response to paroxetine and higher non-remission rates (Lin et al., 2010).
CYP1A2	rs2069526	Allele *G* influences escitalopram metabolism and treatment outcomes in MDD patients (Hsiang-Wei et al., 2013).
CYP1A2	rs2470890	*T* allele associated with improved remission rates with paroxetine treatment (Lin et al., 2010).
CYP1A2	rs2472304	Allele *A* correlates with increased remission rates with paroxetine therapy (Lin et al., 2010).
CYP1A2	rs762551	Genetic variant linked to altered paroxetine dosage requirements and increased fatigue risk (Lin et al., 2010).
ABCB1	rs1045642	Conflicting evidence on its impact on nortriptyline-induced orthostatic hypotension in MDD patients (Roberts et al., 2002; Jensen et al., 2012).
ABCB1	rs2032583	Variant *G* associated with higher risk of adverse effects with SSRIs like citalopram and fluvoxamine (de Klerk et al., 2013).

Table 3. Variants associated with neurotransmitter receptors and transporters.

Gene	Variant	Antidepressant Relevance
HTR2A	rs6313	Genotype *AA* associated with increased risk of sexual dysfunction (SD) due to citalopram in MDD patients compared to *AG* and *GG* genotypes (Oz et al., 2019).
HTR2A	rs7997012	Genotypes *AG* and *GG* linked to increased antidepressant response and decreased risk of adverse events in MDD patients (Lin et al., 2014).
HTR2A	rs6314	No significant differences found in genotype distribution between MDD patients and controls, but certain genotypes associated with better treatment response based on depression rating scale scores (Minov et al., 2001).
HTR2A	rs2770296	Genotype *CC* associated with increased response to bupropion in MDD patients compared to *CT* and *TT* genotypes (Tiwari et al., 2013).
HTR1D	rs6296	Genotype *CC* associated with increased risk of agitation when treated with citalopram in children and adolescents with MDD compared to *CG* and *GG* genotypes (Amitai et al., 2016).
SLC6A4	rs57098334	Certain genotype combinations associated with increased response to fluoxetine or sertraline in MDD patients (Kim et al., 2006).
OPRM1	rs1799971	Genotype *AA* associated with increased risk of suicidal ideation due to tianeptine in adults with Major Depressive Disorder compared to *AG* + *GG* genotypes (Nobile et al., 2019).

Table 4. Variants associated with neuroplasticity and neurotrophic factors.

Gene	Variant	Antidepressant Relevance
BDNF	rs962369	Allele *C* associated with increased likelihood of suicidal ideation during escitalopram or nortriptyline treatment in Major Depressive Disorder (MDD) compared to allele *T* (Perroud et al., 2009).
BDNF	rs6265	Variant associated with worse treatment outcomes in major depression, particularly in anxious and melancholic subtypes, when treated with citalopram (Domschke et al., 2009).
BDNF	rs7103411	Genotype *TT* associated with decreased response to citalopram in MDD compared to genotype *CC* (Domschke et al., 2009).
BDNF	rs7124442	Genotype *TT* associated with decreased response to citalopram in MDD compared to genotypes *CC* + *CT* (Domschke et al., 2009).
GDNF	rs2216711	Genotypes *AA* + *AG* associated with decreased response to paroxetine in women with MDD compared to genotype *GG* (Wang et al., 2014).
GDNF	rs2973049	Genotypes *CT* + *TT* associated with decreased response to paroxetine in MDD compared to genotype *CC* (Wang et al., 2014).
CREB1	rs889895	Genotype *GG* associated with increased response to antidepressants in MDD patients compared to genotypes *AA* + *AG* (Calati et al., 2013).

Table 5. Variants associated with stress response and hormonal regulation.

Gene	Variant	Antidepressant Relevance
FKBP5	rs17614642	Genotype *TT* associated with decreased response to bupropion in Major Depressive Disorder (MDD) compared to genotype *CT* (Tiwari et al., 2013).
COMT	rs4680	Genotype *GG* associated with decreased response to fluvoxamine in MDD compared to allele *A* (Benedetti et al., 2008). Allele A associated with increased response to paroxetine in MDD compared to allele *G* (Benedetti et al., 2008). Genotypes *AG* + *GG* associated with increased response to bupropion in MDD compared to genotype *AA* (Fawver et al., 2020). Genotype GG associated with increased response to venlafaxine in MDD compared to genotypes *AA* + *AG* (Hopkins et al., 2013).
MTHFR	rs1801131	Allele *G* associated with increased response to Vitamin B-complex in MDD (Mech & Farah, 2016).
MTHFR	rs1801133	Allele *A* associated with increased response to Vitamin B-complex in MDD (Mech & Farah, 2016).

Table 6. Variants associated with synaptic function and signal transduction.

Gene	Variant	Antidepressant Relevance
NRXN1	rs4971678	Allele *T* associated with decreased response to duloxetine in Major Depressive Disorder (MDD) compared to allele *A* (Maciukiewicz et al., 2018).
GSK3B	rs334558	Genotype *GG* associated with increased response to citalopram and fluoxetine in MDD compared to allele *A* (Tsai et al., 2008).
CACNA1A	rs2112460	Allele *A* associated with increased response to antidepressants in MDD compared to allele *G* (Cocchi et al., 2016).

Table 7. Variants with other roles in antidepressant response.

Gene	Variant	Relevance to antidepressants
MC1R	rs2228478	Allele *G* associated with increased likelihood of remission with desipramine treatment in Major Depressive Disorder (MDD) compared to allele *A* (Wu et al., 2011).
MC1R	rs2228479	Specific relevance not provided in the current literature; further study needed for comprehensive understanding of its role in antidepressant response in MDD patients.

Patient 1 Results

Identified Variants in Patient 1

Variant rs762551 was identified in patient 1. Variant rs762551 is located within the *CYP1A2* gene on chromosome 15, specifically positioned at 74,749,576 base pairs. Within the gene sequence, this SNP occupies the 731st nucleotide. Originally characterized by the wildtype nucleotide C, patient 1 exhibits a mutation to nucleotide A at this position on one of their chromosome 15 alleles. Therefore, the genotype for this SNP in the patient is *AC*. Due to the presence of the *A* allele, the patient had an increased risk of fatigue ($p = 0.0055$), along with an increased dose requirement ($p = 0.0028$) when treated with paroxetine.

Variant rs4680 was also identified in patient 1, attaining a genotype of *AG*. The presence of allele A in variant rs4680 resulted in risk of efficacy in regard to multiple antidepressants. Genotype *AG* suggested an intermediate response compared to genotype *GG*, which is associated with decreased response to fluvoxamine. Allele *A* was also associated with response to paroxetine compared to allele *G*. Genotype *AG* was also associated with increased response to bupropion compared to genotype *AA*. Genotype *GG* was associated with increased response to venlafaxine compared to genotype *AG* and *AA*. Therefore, patient 1 may have a response that is intermediate.

Additionally, variant rs889895 was identified in this patient. Variant rs889895 is located on chromosome 2 on the *CREB1* gene. Patient 1 had the genotype *AG* for this variant, indicating that the patient could have a less favorable response to antidepressants in comparison to those individuals who possess the *GG* genotype. Their response might be better than a patient with the *AA* genotype, but just not as favorable as the *GG*.

Variant rs2216711 on gene *GDNF*, located on chromosome 5, was another identified variant, attaining the genotype *AG*. With a p-value of 0.013, the association stated that genotypes with *AA* and *AG* are associated with decreased response to paroxetine in women with MDD compared to those with genotype *GG*.

In the gene *CYP2D6*, the variant rs1065852 was identified in the patient. With a p-value of 2.0E-16, allele *A* was found to be associated with plasma concentration of S-didesmethylcitalopram when treated with citalopram or escitalopram in people with MDD compared to those with allele *G*. All of the identified variants are summarized in Table 8.

Table 8. Identified SNPs in patient 1.

Gene	**SNP Location on Gene**	**Original Nucleotide**	**Mutant Nucleotide**	**Genotype**	**Variant**
CYP1A2	731	*C*	*A*	*A/C*	rs762551
COMT	22377	*G*	*A*	*A/G*	rs4680
CREB1	4468	*A*	*G*	*A/G*	rs889895
GDNF	16065	*A*	*G*	*A/G*	rs22126711
CYP2D6	4193	*T*	*A*	*A/T*	rs1065852

Treatment Plan for Patient 1 Based on Genetic Profile

Primary Recommendation: Bupropion

Based on the patient's genotype, bupropion was recommended as the initial treatment option. The *COMT* rs4680 variant with the *AG* genotype was associated with an increased response to bupropion compared to the *AA* genotype. Once patient was started on bupropion, close monitoring for both efficacy and potential side effects was recommended to ensure optimal treatment outcomes.

Secondary Option: Fluvoxamine

If bupropion did not yield the desired results or caused adverse side effects, fluvoxamine was considered an alternative. The patient's *COMT* rs4680 *AG* genotype suggested an intermediate response compared to the *GG* genotype. Fluvoxamine might have been viable, but still required regular assessment to ensure it was effective.

Medications to Use with Caution

Paroxetine

Due to the potential decreased response in women (rs22126711), increased risk of fatigue (rs762551), and increased dose requirement (rs762551), paroxetine was not the optimal choice for this patient.

Citalopram/Escitalopram

The *AA* genotype in variant rs1065852 was linked with higher plasma concentrations of S-didesmethylcitalopram, potentially affecting drug metabolism and efficacy. These medications were recommended to used cautiously, with possible dosage adjustments and close monitoring to avoid adverse side effects due to altered drug metabolism.

General Considerations

The *AG* genotype in variant rs889895 indicated a less favorable response to antidepressants. This was considered when evaluating the overall effectiveness of any chosen antidepressant.

Patient 2 Results

Identified Variants in Patient 2

Based on the genetic test results for SNP rs2470890 located in the *CYP1A2* gene, it was found that patient 2 had the *CT* genotype. This genetic variant had been associated with an increased likelihood of achieving remission when treated with paroxetine for Major Depressive Disorder, compared to individuals with the *CC* genotype.

Furthermore, based on the genetic test results for SNP rs4646427 located in the *CYP1A2* gene, patient 2 had the *CT* genotype. This genetic variant had been associated with a slower response time when treated with paroxetine for Major Depressive Disorder, compared to individuals with the *CC* and *CT* genotypes. Additionally, allele *C* had been linked with increased metabolism of escitalopram, as well as increased fatigue and nausea/vomiting when treated with escitalopram, compared to allele *T*. These associations suggested that, based on patient 2's genetic profile, paroxetine might have led to a slower response in treating depressive symptoms.

Based on the genetic test result for SNP rs334558 located in the *GSK3B* gene, patient 2 was identified to have the *AG* genotype. This genetic variant indicated that genotype *GG* had been associated with increased response to citalopram and fluoxetine for Major Depressive Disorder compared to allele *A* carriers, suggesting that individuals with the *GG* genotype might have experienced better outcomes with these medications.

Based on the genetic test result for SNP rs889895 located in the *CREB1* gene, patient 2 was identified to have the *AG* genotype. This genetic variant indicated that genotype *GG* had been associated with increased response to antidepressants for Major Depressive Disorder compared to genotypes *AA* and *AG*. This association suggested that individuals with the *GG* genotype might have experienced better response to antidepressant treatment. While patient 2 carried the *AG* genotype, which typically indicated a response profile less favorable than *GG* but potentially better than *AA*, this genetic information provided valuable insights into treatment efficacy considerations.

Based on the genetic test results for SNP rs4646425 located in the *CYP1A2* gene, patient 2 was identified to have the *CT* genotype. This genetic variant has shown associations where genotype *CC* was linked with a slower response time when treated with paroxetine for Major Depressive Disorder, compared to genotypes *CT* and *TT*. Additionally, allele *T* has been associated with increased metabolism of escitalopram and increased incidence of fatigue and nausea/vomiting when treated with escitalopram, compared to allele *C*. Furthermore, allele *T* was also associated with a decreased likelihood of achieving remission when treated with paroxetine, compared to allele *C*. These associations provide insight into the potential effects of patient 2's genetic profile on treatment responses. All of the identified variants are summarized in Table 9.

Table 9. Identified SNPs in patient 2.

Gene	SNP Location on Gene	Original Nucleotide	Mutant Nucleotide	Genotype	Variant
CYP1A2	6240	*T*	*C*	*C/T*	rs2470890
CYP1A2	4506	*T*	*C*	*C/T*	rs4646427
GSK3B	273114	*G*	*A*	*A/G*	rs334558
CREB1	4468	*A*	*G*	*A/G*	rs889895
CYP1A2	2095	*C*	*T*	*C/T*	rs64646425

Treatment Plan for Patient 2 Based on Genetic Profile

Primary Recommendation: Citalopram or Fluoxetine

Patient 2's genetic profile suggested that the *AG* genotype for SNP rs334558 in the *GSK3B* gene might indicate an intermediate response to citalopram and fluoxetine compared to the *GG* genotype, which was associated with an increased response. Starting with citalopram or fluoxetine was recommended, with close monitoring to assess efficacy and side effects.

Secondary Option: Bupropion

If citalopram or fluoxetine did not provide the desired response or caused adverse effects, bupropion was considered an alternative. This recommendation was based on general antidepressant efficacy and the patient's genetic profile, which suggested variability in response to SSRIs.

Medications to Use with Caution

Paroxetine

CT genotype in variant rs2470890 was associated with an increased likelihood of achieving remission compared to *CC* genotype. However, caution was advised due to potential slower response times associated with rs4646427 (*CYP1A2*).

Escitalopram

CT genotype in rs4646427 was associated with slower response times, increased metabolism, and higher incidence of fatigue and nausea/vomiting compared to *CC* and *TT* genotypes.

General Considerations

Genotype *AG* in variant rs88989 suggested a response to antidepressants less favorable than *GG* but potentially better than *AA*. Monitoring antidepressant response closely was recommended to optimize treatment outcomes.

Patient 3 Results

Identified Variants in Patient 3

Based on the sequencing analysis, patient 3 was identified to have the *AG* genotype for SNP rs1065852 located in the *CYP2D6* gene. This genetic variant indicates that allele *A* is associated with plasma concentration of S-didesmethyl-citalopram when treated with citalopram or escitalopram for Major Depressive Disorder, compared to allele *G*. This association suggests that individuals with the *AG* genotype may experience differences in plasma concentration levels compared to those with the *GG* genotype.

Using the alignment of the mutated sequence with the reference sequence, it was found that patient 3 had the *CT* genotype for SNP rs6265 located in the *BDNF* gene. This genetic variant indicated that genotype *CT* was associated with increased response to antidepressants for Major Depressive Disorder, compared to genotypes *CC* and *TT*. Additionally, genotype *CC* was linked with decreased response to citalopram compared to genotype *TT*, suggesting that individuals with the *CT* genotype might exhibit a response profile intermediate between CC and *TT* genotypes. Furthermore, genotype *TT* was associated with a decreased response to paroxetine compared to genotypes *CC* and *CT*.

Based on the genetic analysis, patient 3 was found to have the *AG* genotype for SNP rs4680 located in the *COMT* gene. This genetic variant indicated that genotypes *AG* and *GG* were associated with increased response to bupropion for Major Depressive Disorder, compared

to genotype *AA*. Additionally, allele *A* was associated with increased response to paroxetine, suggesting that individuals with the *AG* genotype might have experienced a more favorable outcome compared to those with allele *G.* Moreover, genotype *GG* was linked with increased response to venlafaxine compared to genotypes *AA* and *AG*.

Based on the genetic results for SNP rs4646427 located in the *CYP1A2* gene, patient 3 had the *CT* genotype. This genetic variant had been associated with a slower response time when treated with paroxetine for Major Depressive Disorder, compared to individuals with the *CC* and *CT* genotypes. Additionally, allele *C* had been linked with increased metabolism of escitalopram, as well as increased fatigue and nausea/vomiting when treated with escitalopram, compared to allele *T*. These associations suggested that, based on patient 3's genetic profile, paroxetine might have led to a slower response in treating depressive symptoms.

The genetic analysis revealed that patient 3 had the *AG* genotype for SNP rs6314 located in the *HTR2A* gene. This genetic variant indicated that genotypes *AG* and *GG* were associated with increased response to antidepressants for Major Depressive Disorder, compared to genotype *AA*. This association suggested that individuals with the *AG* genotype might have experienced a more favorable response to antidepressant treatment compared to those with genotype *AA*. All of the identified variants are summarized in Table 10.

Table 10. Identified SNPs in patient 3.

Gene	SNP Location on Gene	Original Nucleotide	Mutant Nucleotide	Genotype	Variant
CYP2D6	4193	*G*	*A*	*A/G*	rs1065852
BDNF	3476	*C*	*T*	*C/T*	rs6265
COMT	22377	*G*	*A*	*A/G*	rs4680
CYP1A2	4506	*T*	*C*	*C/T*	rs4646427
HTR2A	3353	*G*	*A*	*A/G*	rs6314

Treatment Plan for Patient 3 Based on Genetic Profile

Primary Recommendation: Bupropion

Patient 3's genetic profile suggested a potential favorable response to bupropion due to their *AG* genotype for SNP rs4680 in the *COMT* gene. This genotype was associated with increased response to bupropion for Major Depressive Disorder compared to genotype *AA*. Bupropion was recommended as the initial treatment option, with close monitoring for efficacy and side effects to optimize treatment outcomes.

Secondary Option: Fluvoxamine

If bupropion did not yield desired results or was not tolerated well, fluvoxamine was considered as an alternative. Patient 3's *AG* genotype for rs4680 in *COMT* suggested an intermediate response profile compared to *GG*, indicating potential effectiveness with fluvoxamine, albeit less than with bupropion.

Medications to Use with Caution

Paroxetine

Given patient 3's *CT* genotype for SNP rs4646427 in the *CYP1A2* gene, which was associated with a slower response time to paroxetine, this medication was recommended to be used cautiously. The genotype *CT* suggested that paroxetine might lead to a slower improvement in depressive symptoms.

Citalopram/Escitalopram

The association of allele *C* with increased metabolism of escitalopram and potential for increased fatigue and nausea/vomiting (rs4646427) prompted careful consideration. Dosage adjustments and close monitoring were crucial to manage potential side effects.

General Considerations

Considering patient 3's *AG* genotype for SNP rs6314 in the *HTR2A* gene, which was associated with an increased response to antidepressants, selecting medications that targeted serotonin pathways, such as bupropion or fluvoxamine, was particularly beneficial.

Predictive Model Results

A new patient, patient 31, was identified with the genetic variants rs2112460 (genotype *AG*) and rs7103411 (genotype *CT*). Upon inputting these variants into the logistic regression algorithm developed using the patient profile treatment data, escitalopram was predicted as the most suitable antidepressant for this patient. Additionally, patient 31 was found to have the variants rs4971678 (genotype *AT*) and rs889895 (genotype *AG*). The predicted antidepressant for this patient was escitalopram.

The classification report in Table 11 offered a comprehensive assessment of the model's efficacy predicting antidepressant choices based on genetic profiles. Precision, which measures the accuracy of positive predictions, varied across different antidepressants. For example, the model achieved a precision of 0.83 for predicting bupropion, indicating that it accurately predicted this antidepressant 83% of the time. Conversely, citalopram had a precision of 0.00, implying the model did not correctly predict this antidepressant in any instance. When the primary antidepressant choice was 'citalopram or fluoxetine', 'desipramine', and 'desipramine or bupropion', the model achieved perfect precision with a score of 1.00, indicating it had flawless predictions when these antidepressants were the true label.

Recall, also known as sensitivity, gauges the model's ability to correctly identify actual positives. Escitalopram exhibited a recall of 0.89, indicating the model accurately identified 89% of instances where escitalopram was the true label. The F1-score, which balances precision and

recall, highlights the harmonic mean of these metrics. Paroxetine, for instance, achieved an F1-score of 0.80, suggesting a well-balanced performance between precision and recall for this category.

Support denotes the actual number of occurrences of each antidepressant in the test dataset. For instance, there were 9 instances of escitalopram in the dataset, reflecting its prevalence among the predictions made by the model.

The model demonstrated an accuracy of 0.70, indicating that it correctly predicted the antidepressant for 70% of the patients in the test set. The macro average precision, recall, and F1-score were 0.78, 0.73, and 0.73 respectively, providing an aggregated view across all antidepressants without considering class imbalance. In contrast, the weighted average precision, recall, and F1-score were 0.72, 0.70, and 0.68 respectively, offering a more comprehensive evaluation of the model's performance across all classes, considering varying frequencies of antidepressant prescriptions.

Table 11. Predictive Model Classification Report

Antidepressant	**Precision**	**Recall**	**F1-Score**	**Support**
Bupropion	0.83	0.71	0.77	7
Citalopram	0.33	0.33	0.33	3
Citalopram or Fluoxetine	1.00	1.00	1.00	1
Desipramine	1.00	1.00	1.00	1
Desipramine or Bupropion	1.00	1.00	1.00	1
Escitalopram	0.67	0.89	0.76	9
Fluoxetine	0.50	0.33	0.40	3
Fluoxetine or Sertraline	1.00	0.33	0.50	3
Paroxetine	0.67	1.00	0.80	2
Accuracy			0.70	30
Macro Average	0.78	0.73	0.71	30
Weighted Average	0.72	0.70	0.68	30

CHAPTER 4 DISCUSSION

In recent years, advancements in genetic sequencing have transformed the approach to treating Major Depressive Disorder (MDD), introducing a personalized medicine paradigm centered around individual genetic profiles. This study used patient DNA sequences to pinpoint critical single nucleotide polymorphisms (SNPs) linked to how patients respond to antidepressant medications and their associated side effects. These genetic variants were meticulously compared against reference sequences to tailor precise treatment plans aimed at optimizing therapeutic outcomes.

Subsequently, these tailored treatment plans served as the foundation for developing a predictive algorithm. This algorithm was designed to forecast the most suitable antidepressant medication based on each patient's identified genetic SNPs. By integrating genetic insights into clinical decision-making processes, the research aimed to enhance the efficacy and safety of antidepressant therapies for individuals diagnosed with MDD.

Overall, this study represented a pivotal step towards advancing precision psychiatry by leveraging genetic data to inform personalized treatment strategies, ultimately aiming to improve patient outcomes and quality of life in managing Major Depressive Disorder.

Results Linked to Literature Evidence

The results from the genetic profiling of 30 patients revealed significant associations between specific genetic variants and treatment responses, aligning with findings from existing literature.

Patient 1 exhibited the rs762551 variant in the *CYP1A2* gene, indicating increased fatigue risk and higher paroxetine dose requirements, aligning with previous studies highlighting its influence on drug metabolism and side effects (Lin et al., 2010). The rs4680 variant in *COMT* suggested an intermediate response to antidepressants, influencing treatment recommendations towards bupropion, which was supported by studies demonstrating its efficacy in patients with similar genotypes (Fawver et al., 2020). Similarly, rs889895 in *CREB1* suggested a less favorable response to antidepressants compared to the *GG* genotype, emphasizing the importance of genotype-specific treatment considerations (Calati et al., 2013).

In patient 2, the genetic profile indicated variants associated with differential responses to antidepressants. Notably, the rs2470890 and rs4646427 variants in *CYP1A2* suggested slower response times to paroxetine, aligning with literature suggesting genotype *CC* correlates with delayed efficacy (Han et al., 2013). The rs334558 variant in *GSK3B* indicated an intermediate response to citalopram and fluoxetine, influencing initial treatment choices towards these SSRIs. These findings underscored the utility of genetic insights in optimizing antidepressant selection for individual patients.

Patient 3's genetic analysis highlighted the rs4680 variant in *COMT*, guiding treatment towards bupropion due to its association with improved response in patients with *AG* genotype, a finding consistent with previous research (Benedetti et al., 2008). Additionally, rs4646427 in *CYP1A2* indicated a slower response to paroxetine, necessitating cautious use of this medication. The rs6314 variant in *HTR2A* suggested an enhanced response to antidepressants, supporting serotonin-targeting therapies like fluvoxamine or bupropion (Serretti et al., 2007).

Limitations

While the findings provided valuable insights, several limitations warranted consideration to contextualize the results effectively. The primary constraint of this study was the use of simulated DNA profiles rather than real patient data. While simulated data allowed controlled experimentation, it may not have fully captured the complexities and variations present in actual human genomes. Future studies should prioritize the inclusion of real patient DNA sequences to validate findings and enhance the clinical relevance of genotype-based treatment recommendations.

Another limitation pertained to the diversity of ethnicities represented in the genetic studies referenced. Genetic variations and their implications for drug response can vary significantly across different ethnic groups. The study relied on literature that often encompassed only one ethnicity of patients, and ethnicities varied between studies. This restricted the generalizability of findings to diverse populations. Future research efforts should prioritize including diverse ethnic cohorts to better understand genetic influences on antidepressant response across different demographic groups.

Moreover, while genomic sequencing technologies advanced, they have not yet become universally accessible or routine in clinical practice at the time of this study. The complexity and cost associated with genetic testing could pose barriers to widespread adoption in healthcare settings. As technologies continued to evolve and became more affordable, integrating genetic testing into routine clinical practice for personalized medicine approaches will become increasingly feasible.

In discussing the limitations of this study, it was crucial to acknowledge the scope of genetic variants examined. This research primarily focused on single nucleotide polymorphisms

(SNPs) identified by rsIDs, while excluding other types of genetic variations such as insertions, deletions, and structural variants that also played roles in influencing drug response. This narrow focus restricted the comprehensive understanding of genetic determinants in pharmacogenomics, potentially overlooking important contributors to antidepressant efficacy and side effects.

Furthermore, the study did not fully explore the complex genetic interactions within and between genes, which could significantly modulate the effects of SNPs on drug response. These interactions were pivotal for refining predictions based on genetic profiles and advancing personalized medicine approaches effectively. Future research efforts should aim to broaden the scope to include a wider array of genetic variations beyond SNPs, thereby capturing the full spectrum of genetic influences on antidepressant treatment outcomes.

It was important to note that while technological advancements facilitated the identification and study of genetic variants, there remained limited evidence and understanding regarding the impact of insertions, deletions, and structural variants on antidepressant response. Addressing these gaps through rigorous research and comprehensive data integration would be essential for enhancing the accuracy and applicability of genetic insights in clinical settings. This approach held promise for optimizing antidepressant therapy based on individual genetic profiles, ultimately improving treatment efficacy and patient outcomes across diverse populations.

Another critical factor to consider regarding the predictive model was the limited size of the patient profile dataset, which included genetic variants and primary drug choices. Using a smaller dataset for logistic regression can lead to detrimental effects that impact the model's robustness and reliability. Logistic regression models trained on small datasets are susceptible to overfitting, where the model captures noise or random fluctuations rather than genuine

underlying patterns. Additionally, small datasets tend to exhibit greater variability in outcomes due to the limited number of data points available for training. These challenges highlight the importance of addressing dataset size when developing predictive models to ensure more accurate and dependable results in clinical applications.

Extensions

In extending this project, several innovative directions could significantly enhance the understanding and application of pharmacogenomics in antidepressant treatment. Longitudinal studies tracking genetic variants over time would provide insights into how these variants evolve and their sustained influence on treatment outcomes. Integrating multi-omics data could uncover comprehensive molecular mechanisms underlying antidepressant response, potentially identifying new biomarkers and therapeutic targets. Advanced machine learning and artificial intelligence techniques could refine predictive models for personalized therapy, improving treatment efficacy based on genetic profiles.

Expanding the study to encompass larger, diverse population cohort studies would enhance the generalizability of findings across different ethnicities and backgrounds. Additionally, conducting functional validation studies using experimental models would validate genetic variants' biological effects on drug metabolism and response, strengthening the clinical relevance of genetic insights. Educational initiatives for patients and healthcare providers on genetics in antidepressant therapy would promote informed decision-making and treatment adherence.

Furthermore, addressing ethical, legal, and social implications associated with genetic testing in psychiatry is crucial. This includes privacy concerns, insurance coverage implications, and ensuring equitable access to genetic testing and personalized medicine approaches.

Designing prospective clinical trials and implementation studies would bridge research findings with clinical practice, validating genotype-guided therapy and guiding evidence-based treatment strategies.

One notable limitation of the predictive model was its reliance on a relatively small dataset. To address this, an extension of this project could involve generating additional simulated patient profiles, constructing corresponding treatment plans, and integrating this new data into the existing dataset used for training the algorithm. By expanding the dataset in this manner, the model would benefit from a more diverse and comprehensive representation of genetic profiles and treatment responses, thereby enhancing its robustness and capacity to generate reliable predictions for personalized antidepressant therapy. This approach aims to mitigate the limitations posed by the initial dataset size, enabling more accurate and effective clinical decision-making based on genetic insights.

By pursuing these extensions, future research can advance the field of pharmacogenomics in psychiatry, ultimately aiming to personalize antidepressant therapy based on individual genetic profiles and improve outcomes for patients with major depressive disorder. These efforts will contribute to the evolution of precision medicine in psychiatry, enhancing therapeutic efficacy and patient care.

Practical Benefits of Precision Medicine

In addition to exploring potential extensions, it was crucial to underscore the practical benefits of this study's methodology. By concentrating on clinically proven associations sourced from PharmGKB and implementing a program to align mutated versus reference sequences for identifying SNPs and their precise locations, this approach significantly enhanced data accessibility and relevance. Traditionally, prescribers had grappled with the arduous task of

navigating through extensive datasets that contained clinical annotations, often comprising over 5000 rows. This process was time-consuming and may not have consistently yielded pertinent information tailored to individual patients. In contrast, this method also offered a distinct advantage over trial-and-error approaches, where treatment adjustments were based solely on observed patient responses without the insights provided by genetic data.

The developed program facilitated streamlined query inputs, allowing prescribers to obtain targeted outputs specific to individual patient profiles. This capability was pivotal in clinical settings, where swift and accurate identification of genetic variants influencing antidepressant response could significantly inform personalized treatment decisions. By reducing the complexity and time required for interpreting data, this methodology empowered prescribers to make informed clinical decisions efficiently.

Moreover, the use of the predictive algorithm derived from patient treatment plans added another layer of benefit. It enabled automated analysis of genetic data to predict optimal antidepressant choices based on individual genetic profiles. This predictive capability not only enhanced precision medicine in psychiatry but also supported clinicians in selecting the most suitable treatments promptly and effectively. Ultimately, this integrated approach ensured that genetic insights translated effectively into improved patient care outcomes, potentially leading to cost savings, reduced adverse events, and improved patient satisfaction in clinical settings.

Conclusion

Advancements in genetic sequencing have transformed Major Depressive Disorder (MDD) treatment, emphasizing personalized medicine based on genetic profiles. This study utilized patient DNA sequences to identify significant SNPs influencing antidepressant responses and side effects, aligning them with reference sequences to optimize treatment strategies.

Integrating genetic data into clinical decision-making represents a pivotal step towards precision psychiatry in managing depressive disorders.

Genetic profiling of the patients identified key associations between specific variants and treatment outcomes, consistent with existing literature. For instance, variants like rs762551 in *CYP1A2* correlated with increased fatigue risk and altered medication requirements.

In conclusion, this study underscored the potential of pharmacogenomics to refine antidepressant therapy through individualized genetic profiles. Addressing its limitations and advancing research into broader genetic landscapes will further propel precision medicine in psychiatry, optimizing therapeutic interventions and patient care outcomes.

CHAPTER 5 LITERATURE CITED

Amitai, M., Kronenberg, S., Carmel, M., Michaelovsky, E., Frisch, A., Brent, D., Apter, A., Chen, A., Weizman, A., & Fennig, S. (2016). Pharmacogenetics of citalopram-related side effects in children with depression and/or anxiety disorders. *Journal of Neural Transmission*, *123*(11), 1347–1354. https://doi.org/10.1007/s00702-016-1585-7

Benedetti, F., Colombo, C., Pirovano, A., Marino, E., & Smeraldi, E. (2008). The catechol-O-methyltransferase Val(108/158)Met polymorphism affects antidepressant response to paroxetine in a naturalistic setting. *Psychopharmacology*, *203*(1), 155–160. https://doi.org/10.1007/s00213-008-1381-7

Benedetti, F., Dallaspezia, S., Colombo, C., Lorenzi, C., Pirovano, A., & Smeraldi, E. (2010). Effect of catechol-O-methyltransferase Val(108/158)Met polymorphism on antidepressant efficacy of fluvoxamine. *European Psychiatry*, *25*(8), 476–478. https://doi.org/10.1016/j.eurpsy.2009.12.007

Calati, R., Crisafulli, C., Balestri, M., Serretti, A., Spina, E., Calabrò, M., Sidoti, A., Albani, D., Massat, I., Höfer, P., Amital, D., Juven-Wetzler, A., Kasper, S., Zohar, J., Souery, D., Montgomery, S., & Mendlewicz, J. (2013). Evaluation of the role of MAPK1 and CREB1 polymorphisms on treatment resistance, response and remission in mood disorder patients. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *44*, 271–278. https://doi.org/10.1016/j.pnpbp.2013.03.005

Cocchi, E., Fabbri, C., Han, C., Soo Jung Lee, Patkar, A. A., Masand, P. S., Chi Un Pae, & Alessandro Serretti. (2016). Genome-wide association study of antidepressant response: involvement of the inorganic cation transmembrane transporter activity pathway. *BMC Psychiatry*, *16*(1). https://doi.org/10.1186/s12888-016-0813-x

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., & de Hoon, M. J. L. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics, 25(11), 1422-1423. https://doi.org/10.1093/bioinformatics/btp163

de Klerk, O. L., Nolte, I. M., Bet, P. M., Bosker, F. J., Snieder, H., den Boer, J. A., Bruggeman, R., Hoogendijk, W. J., & Penninx, B. W. (2012). ABCB1 gene variants influence tolerance to selective serotonin reuptake inhibitors in a large sample of Dutch cases with major depressive disorder. *The Pharmacogenomics Journal*, *13*(4), 349–353. https://doi.org/10.1038/tpj.2012.16

Domschke, K., Lawford, B., Laje, G., Berger, K., Young, R., Morris, P., Deckert, J., Arolt, V., McMahon, F. J., & Baune, B. T. (2009). Brain-derived neurotrophic factor (BDNF)

gene: no major impact on antidepressant treatment response. *The International Journal of Neuropsychopharmacology*, *13*(01), 93. https://doi.org/10.1017/s1461145709000030

Fawver, J., Flanagan, M., Smith, T., Drouin, M., & Mirro, M. (2020). The association of COMT genotype with buproprion treatment response in the treatment of major depressive disorder. *Brain and Behavior*, *10*(7). https://doi.org/10.1002/brb3.1692

Flint, Jonathan. "The Genetic Basis of Major Depressive Disorder." *Molecular Psychiatry*, 26 Jan. 2023, https://doi.org/10.1038/s41380-023-01957-9.

Han, K.-M., Chang, H. S., Choi, I.-K., Ham, B.-J., & Lee, M.-S. (2013). CYP2D6 P34S Polymorphism and Outcomes of Escitalopram Treatment in Koreans with Major Depression. *Psychiatry Investigation*, *10*(3), 286–293. https://doi.org/10.4306/pi.2013.10.3.286

Hopkins, S. C., Reasner, D. S., & Koblan, K. S. (2013). Catechol-O-methyltransferase genotype as modifier of superior responses to venlafaxine treatment in major depressive disorder. *Psychiatry Research*, *208*(3), 285–287. https://doi.org/10.1016/j.psychres.2013.04.021

Jensen, B. P., Roberts, R. L., Vyas, R., Bonke, G., Jardine, D. L., & Begg, E. J. (2012). Influence of ABCB1 (P-glycoprotein) haplotypes on nortriptyline pharmacokinetics and nortriptyline-induced postural hypotension in healthy volunteers. *British Journal of Clinical Pharmacology*, *73*(4), 619–628. https://doi.org/10.1111/j.1365-2125.2011.04126.x

Kruglyak, L., & Nickerson, D. A. (2001). Variation is the spice of life. *Nature Genetics*, *27*(3), 234–236. https://doi.org/10.1038/85776

Kim, H., Lim, S.-W., Kim, S., Kim, J.-W., Chang, Y. H., Carroll, B. J., & Kim, D. K. (2006). Monoamine transporter gene polymorphisms and antidepressant response in koreans with late-life depression. *JAMA*, *296*(13), 1609–1618. https://doi.org/10.1001/jama.296.13.1609

Kuo, H.-W., Liu, S. C., Tsou, H.-H., Liu, S.-W., Lin, K.-M., Lu, S.-C., Hsiao, M.-C., Hsiao, C.-F., Liu, C.-Y., Chen, C.-H., Lu, M.-L., Shen, W. W., Tang, H.-S., Liu, S.-I., Chang, L.-H., Wu, H.-Y., Chang, Y.-S., Yeh, T.-K., Chen, A. C., & Liu, Y.-L. (2013). CYP1A2 genetic polymorphisms are associated with early antidepressant escitalopram metabolism and adverse reactions. *Pharmacogenomics*, *14*(10), 1191–1201. https://doi.org/10.2217/pgs.13.105

Lin, J.-Y., Jiang, M.-Y., Kan, Z.-M., & Chu, Y. (2014). Influence of 5-HTR2A genetic polymorphisms on the efficacy of antidepressants in the treatment of major depressive disorder: A meta-analysis. *Journal of Affective Disorders*, *168*, 430–438. https://doi.org/10.1016/j.jad.2014.06.012

Lin, K.-M., Tsou, H.-H., Tsai, I-Ju., Hsiao, M.-C., Hsiao, C.-F., Liu, C.-Y., Shen, W. W., Tang, H.-S., Fang, C.-K., Wu, C.-S., Lu, S.-C., Kuo, H.-W., Liu, S. C., Chan, H.-W., Hsu, Y.-T., Tian, J.-N., & Liu, Y.-L. (2010). CYP1A2genetic polymorphisms are associated with treatment response to the antidepressant paroxetine. *Pharmacogenomics*, *11*(11), 1535–1543. https://doi.org/10.2217/pgs.10.128

Maciukiewicz, M., Marshe, V. S., Hauschild, A.-C., Foster, J. A., Rotzinger, S., Kennedy, J. L., Kennedy, S. H., Müller, D. J., & Geraci, J. (2018). GWAS-based machine learning approach to predict duloxetine response in major depressive disorder. *Journal of Psychiatric Research*, *99*, 62–68. https://doi.org/10.1016/j.jpsychires.2017.12.009

McKinney, W. (2010). Data structures for statistical computing in Python. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 51-56). Austin, TX: SciPy. https://doi.org/10.25080/Majora-92bf1922-00a

Mech, A. W., & Farah, A. (2016). Correlation of Clinical Response With Homocysteine Reduction During Therapy With Reduced B Vitamins in Patients With MDD Who Are Positive for MTHFR C677T or A1298C Polymorphism. *The Journal of Clinical Psychiatry*, *77*(05), 668–671. https://doi.org/10.4088/jcp.15m10166

Minov, C., Baghai, T. C., Schüle, C., Zwanzger, P., Schwarz, M. J., Zill, P., Rupprecht, R., & Bondy, B. (2001). Serotonin-2A-receptor and -transporter polymorphisms: lack of association in patients with major depression. *Neuroscience Letters*, *303*(2), 119–122. https://doi.org/10.1016/s0304-3940(01)01704-9

Murphy, G. M., Kremer, C., Rodrigues, H. E., & Schatzberg, A. F. (2003). Pharmacogenetics of Antidepressant Medication Intolerance. *American Journal of Psychiatry*, *160*(10), 1830–1835. https://doi.org/10.1176/appi.ajp.160.10.1830

Nobile, B., Ramoz, N., Jaussent, I., Gorwood, P., Olié, E., Castroman, J. L., Guillaume, S., & Courtet, P. (2019). Polymorphism A118G of opioid receptor mu 1 (OPRM1) is associated with emergence of suicidal ideation at antidepressant onset in a large naturalistic cohort of depressed outpatients. *Scientific Reports*, *9*, 2569. https://doi.org/10.1038/s41598-019-39622-3

Oz, M. D., Baskak, B., Uckun, Z., Artun, N. Y., Ozdemir, H., Ozel, T. K., Ozguven, H. D., & Suzen, H. S. (2019). Association between serotonin 2A receptor (HTR2A), serotonin transporter (SLC6A4) and brain-derived neurotrophic factor (BDNF) gene polymorphisms and citalopram/sertraline induced sexual dysfunction in MDD patients. *The Pharmacogenomics Journal*, *20*(3), 443–450. https://doi.org/10.1038/s41397-019-0127-8

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825-2830. https://jmlr.org/papers/v12/pedregosa11a.html

Perroud, N., Aitchison, K. J., Uher, R., Smith, R., Huezo-Diaz, P., Marusic, A., Maier, W., Mors, O., Placentino, A., Henigsberg, N., Rietschel, M., Hauser, J., Souery, D., Kapelski, P., Bonvicini, C., Zobel, A., Jorgensen, L., Petrovic, A., Kalember, P., & Schulze, T. G. (2009). Genetic Predictors of Increase in Suicidal Ideation During Antidepressant Treatment in the GENDEP Project. *Neuropsychopharmacology*, *34*(12), 2517–2528. https://doi.org/10.1038/npp.2009.81

Reich, D., Gabriel, S., & Altshuler, D. (2003). Quality and completeness of SNP databases. *Nature Genetics*, *33*(4), 457–458. https://doi.org/10.1038/ng1133

Reitz, K. (2023). Requests: HTTP for Humans. https://docs.python-requests.org/en/latest/

Roberts, R. L., Joyce, P. R., Mulder, R. T., Begg, E. J., & Kennedy, M. A. (2002). A common P-glycoprotein polymorphism is associated with nortriptyline-induced postural hypotension in patients treated for major depression. *The Pharmacogenomics Journal*, *2*(3), 191–196. https://doi.org/10.1038/sj.tpj.6500099

Stephens, J. C. (2001). Haplotype Variation and Linkage Disequilibrium in 313 Human Genes. *Science*, *293*(5529), 489–493. https://doi.org/10.1126/science.1059431

Tiwari, A. K., Zai, C. C., Sajeev, G., Arenovich, T., Müller, D. J., & Kennedy, J. L. (2013). Analysis of 34 candidate genes in bupropion and placebo remission. *International Journal of Neuropsychopharmacology*, *16*(4), 771–781. https://doi.org/10.1017/S1461145712000843

Tsai, S.-J. ., Liou, Y.-J. ., Hong, C.-J. ., Yu, Y. W.-Y. ., & Chen, T.-J. . (2008). Glycogen synthase kinase-3beta gene is associated with antidepressant treatment response in Chinese major depressive disorder. *The Pharmacogenomics Journal*, *8*(6), 384–390. https://doi.org/10.1038/sj.tpj.6500486

Wan, Y., Zhai, X., Tan, H., Ai, Y., & Zhao, L. (2020). Associations between the 1438A/G, 102T/C, and rs7997012G/A polymorphisms of HTR2A and the safety and efficacy of antidepressants in depression: a meta-analysis. *The Pharmacogenomics Journal*, *21*(2), 200–215. https://doi.org/10.1038/s41397-020-00197-2

Wang, X.-C., Xu, D.-J., Chen, G.-H., Xia, Q., & Liu, L.-N. (2014). Association of 2 Neurotrophic Factor Polymorphisms With Efficacy of Paroxetine in Patients With Major Depressive Disorder in a Chinese Population. *Therapeutic Drug Monitoring*, *36*(5), 612–617. https://doi.org/10.1097/ftd.0000000000000062

Wu, G.-S., Luo, H.-R., Dong, C., Mastronardi, C., Licinio, J., & Wong, M.-L. (2011). Sequence polymorphisms of MC1R gene and their association with depression and antidepressant response. *Psychiatric Genetics*, *21*(1), 14–18. https://doi.org/10.1097/ypg.0b013e32834133d2

CHAPTER 6 APPENDIX

Google Drive

The following link is for a shared google drive containing all of the datasets that were utilized for the project. It also consists of any files created during the process, gene DNA reference sequences, modified sequences from patient 1, 2, and 3. The drive also consists of the Python scripts used to conduct this project. Additionally, this drive contains the csv file of the patient profile training data used to build the predictive algorithm.

https://drive.google.com/drive/folders/1uYIGPWHdOvpNjyeIdH5IZf8Pdkc9j32H?usp=drive_link