Pharmacogenomics of selective serotonin reuptake inhibitor treatment for major depressive disorder: genome-wide associations and functional genomics

Y Ji1,2, JM Bienacka2,7, S Hebbbring1, Y Chai1, GD Jenkins2, A Batzler2, KA Snyder3, MS Drews4, Z Desta5, D Flockhart5, T Mushiroda6, M Kubo6, Y Nakamura6, N Kamatani6, D Schaid2, RM Weinshilboum1 and DA Mrazek3

A genome-wide association (GWA) study of treatment outcomes (response and remission) of selective serotonin reuptake inhibitors (SSRIs) was conducted using 529 subjects with major depressive disorder. While no SNP associations reached the genome-wide level of significance, 14 SNPs of interest were identified for functional analysis. The rs11144870 SNP in the riboflavin kinase (RFK) gene on chromosome 9 was associated with 8-week treatment response (odds ratio (OR) = 0.42, P = 1.04 × 10^{-6}). The rs915120 SNP in the G protein-coupled receptor kinase 5 (GRK5) gene on chromosome 10 was associated with 8-week remission (OR = 0.50, P = 1.15 × 10^{-5}). Both SNPs were shown to influence transcription by a reporter gene assay and to alter nuclear protein binding using an electrophoretic mobility shift assay. This report represents an example of joining functional genomics with traditional GWA study results derived from a GWA analysis of SSRI treatment outcomes. The goal of this analytical strategy is to provide insights into the potential relevance of biologically plausible observed associations.

The Pharmacogenomics Journal (2013) 13, 456–463; doi:10.1038/tpj.2012.32; published online 21 August 2012

Keywords: selective serotonin reuptake inhibitors; SSRI; genome-wide association study; GWA; functional genomics; major depressive disorder

INTRODUCTION

Major depressive disorder (MDD) is a serious and prevalent psychiatric illness.1 Selective serotonin reuptake inhibitors (SSRIs) are an effective treatment for MDD, but treatment response is highly variable.2 Consequently, the identification of pharmacogenetic predictors of variable drug response phenotypes has become a major objective of MDD research.3,4,5

Genome-wide association (GWA) studies provide an agnostic approach which can identify novel genetic variants that may contribute to variation in drug response phenotypes.6 Three GWA studies of SSRI antidepressant outcomes have been reported, but none have identified associations of SNPs with treatment outcomes that reached genome-wide statistical significance or were validated in a separate population.7–9 This failure to demonstrate genome-wide significance may be related to (1) insufficient statistical power to detect small effect sizes for associated variants, (2) inconsistently defined phenotypes, (3) confounding effects from non-genetic factors or (4) inconsistencies in treatment protocols.10,11 However, the conduct of functional genomic studies of the SNPs identified by GWA analyses provides a novel strategy to identify genomic variants that may influence antidepressant treatment.

The present study is a report of GWA analyses of MDD patients who participated in a single-site pharmacogenomic study of SSRI therapy. Following the GWA analyses, reporter gene assays and electrophoretic mobility shift (EMS) assays were performed to determine whether any identified candidate SNPs were associated with a change in function.

MATERIALS AND METHODS

Study design and samples

The Mayo Clinic Pharmacogenomic Research Network Antidepressant Medication Pharmacogenomic Study (PGRN-AMPS) was supported by the NIGMS-Pharmacogenomics Research Network (PGRN), which has been described elsewhere.12,13 The PGRN-AMPS is an ongoing 8-week outpatient SSRI clinical trial that was performed at the Mayo Clinic in Rochester, MN, USA. Patients enrolled in the study met diagnostic criteria for MDD without psychosis or mania and had a 17-item Hamilton Depression Rating Scale (HAMD-17) score ≥ 14. The study was designed with inclusion and exclusion criteria similar to those used in the Sequenced Treatment Alternatives to Relieve Depression study (STAR*D).14 Potential study subjects taking an antidepressant, antipsychotic or mood-stabilizing medication were not eligible for enrollment. Patients with MDD initially received either 10 mg of escitalopram or 20 mg of citalopram. SSRI efficacy was determined using the 16-item Quick Inventory of Depressive Symptomatology (QIDS-C16) scores after 4-weeks and then 8-weeks of SSRI therapy. At 4-weeks after the initiation of treatment, the dose could be increased to 20 mg of escitalopram or 40 mg of citalopram after a clinical assessment of the subject. Unless there was a contraindication, dose was increased if the QIDS-C16 score at the follow-up visit was ≥ 9, and possibly following a clinical evaluation if the score was between 6 and 8. The dose could also be decreased, or treatment could be discontinued, if a patient

1Division of Clinical Pharmacology, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN, USA; 2Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA; 3Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, USA; 4Department of Information Technology, Mayo Clinic, Rochester, MN, USA; 5Department of Medicine, Indiana University, Indianapolis West Lafayette, IN, USA and 6RIKEN Center for Genomic Medicine, Yokohama, Japan.

Correspondence: Dr DA Mrazek, Department of Psychiatry and Psychology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA.

E-mail: mrazek.david@mayo.edu

These authors contributed equally to this work

Received 3 January 2012; revised 26 June 2012; accepted 17 July 2012; published online 21 August 2012
developed consistent side effects. Blood samples were obtained at baseline for DNA extraction, and then again at weeks four and eight for assays of drug and metabolite levels. All patients provided written informed consent. The study protocol was approved by the Mayo Clinic Institutional Review Board.

**Treatment outcomes**

The two primary outcome phenotypes in the GWA analyses were ‘response’ (defined as ≥ 50% reduction in QIDS-C16 score from baseline to the last visit) and ‘remission’ (defined as a QIDS-C16 score of ≤ 5 at the last visit). For each of these two outcomes, analyses were performed using two strategies. The primary analyses included only subjects that were evaluated at the 8-week visit. The secondary analyses were performed with outcomes based on the final visit QIDS-C16 scores, referred to as the ‘last visit’ assessment. These analyses included subjects who had completed the full 8-week study, as well as those who dropped out of the study before the 8-week assessment.

**Genotyping and QC**

DNA from 529 patients was genotyped by the RIKEN Center for Genomic Medicine (Yokohama, Japan) using Illumina Human610-Quad BeadChips (Illumina, San Diego, CA, USA). Tagman genotyping assays (Applied Biosystems, Foster City, CA, USA) were used to perform genotyping for the replication study using STAR*D samples, and the Illumina Goldengate platform was used to genotype selected top SNPs identified in the published STAR*D GWAS using Mayo PGRN-AMPS DNA samples. Quality control assessments included overall genotype concordance rates based on duplicate sample genotyping and Mendelian inheritance checks based on genotyping of a CEPH trio of two parents and their child. For each SNP, the minor allele frequency, call rate, and departure from Hardy–Weinberg equilibrium were evaluated. Observed call rates, total heterozygosity and inbreeding coefficients were assessed for each subject using PLINK.15 Sex-checks based on X-chromosome heterozygosity were performed, and tests of identity-by-descent were used to identify potentially related subjects.

More than 97% of study participants were of white non-Hispanic (WNH) ancestry. Consequently, genotype–phenotype association analyses were restricted to the WNH subjects. A subset of 4855 independent SNPs with low local linkage disequilibrium (\(r^2 < 0.063\)) was used to verify self-reported ancestry of the subjects using the software STRUCTURE.16 In addition to the study population, 287 DNA samples from the ‘Human Variation Panel’ of lymphoblastoid cell lines (sample sets HD100CAU, HD100AA and HD100CHI) obtained from the Coriell Institute (Camden, NJ, USA) were included in this analysis. Probabilities of membership in each of the three known ancestral groups of the cell lines were calculated for each study subject. These analyses were used to verify self-reported race and to assign ancestry based on STRUCTURE analysis. Thus, 514 WNH subjects were included in the study.

Statistical analyses

Before genotype–phenotype association analyses, EigenStrat software was used to determine eigenvalues for the SNP correlation matrix.17 Eigenvalues that differed statistically from zero were determined based on Tracy–Widom test P-values, and the corresponding eigenvectors were used as covariates in the genetic analyses to account for differences in ancestry within the WNH set of subjects.

Of the 514 WNH subjects, 15 patients were found to have very low blood drug levels, suggesting non-adherence, and were excluded from analysis, resulting in 499 subjects who were included in the primary GWA analyses. Logistic regression was used to test for associations between the binary outcome variables, response and remission, and the genotype of each SNP. Genotypes were coded as the ‘dose’ of the minor allele. These analyses assume log-additive allele effects on response or remission. All GWA analyses were adjusted for eigenvectors that captured population stratification. Gender and age (categorized into four quartiles) were also considered as covariates for the GWA analyses. Age was not significantly associated with response or remission (\(P > 0.10\) for both outcomes). Gender was not associated with response (\(P > 0.10\) for both 8-week and final-visit response), but was marginally associated with remission (\(P = 0.057\) for 8-week remission, \(P = 0.036\) for final-visit remission). We therefore performed response and remission GWA analyses both with and without gender as a covariate. As the results were nearly identical, we report only findings from the analyses not adjusted for gender. Analyses were performed in R (http://www.R-project.org), SAS (SAS Institute) and PLINK.18 Odds ratios (ORs), 95% confidence intervals, and P-values were calculated based on the logistic regression models. P-values \(< 10^{-5}\) were considered statistically significant at the genome-wide level.

Before performing the GWA analyses, the power to detect SNP effects was estimated, assuming 500 subjects with 30% failing to achieve remission. It was determined that with a type 1 error rate of 10\(^{-4}\), the data would provide >80% power to detect ORs of at least 1.8 for common alleles (MAF 0.10–0.20).

**Functional genomics**

On review of the GWA analyses, 14 candidate SNPs were selected based on having some evidence of an association (\(P < 9.0 \times 10^{-5}\)) and being located within or near a candidate gene. SNPs in locations that could potentially influence gene transcription were preferentially selected. Reporter gene assays were used to demonstrate the effects of these SNPs on transcription and EMS assays were used to assess DNA–protein binding.

For the reporter gene assays, 200–300 bp DNA sequences that included a candidate SNP were cloned into a pgL3-promoter luciferase reporter vector that contained an SV40 promoter upstream of the luciferase gene (Promega Corporation, Fitchburg, WI, USA). One microgram of each reporter gene construct was cotransfected with 20 ng of the pRL-TK renilla luciferase vector as a control for transfection efficiency into a human neuroblastoma cell line, SK-N-BE(2), and two human glioblastoma cell lines, U-87 MG and U251 (American Type Culture Collection, ATCC, Manassas, VA), followed by dual-luciferase assays performed 24 h after transfection (Promega). Two independent transfections were performed for each reporter gene construct, with triplicate transfections for each construct in each experiment. Values for relative activity were expressed as percentage changes of the pgL3-promoter activity for vector without an insert. Comparisons were then made between pgL3 reporter gene constructs containing wild-type and variant nucleotides for each candidate SNP. DNA samples that were used to amplify SNP regions were selected from the Coriell ‘Human Variation Panel’ DNA samples (Camden).18 Sequences of primers used to amplify genomic regions containing the selected SNPs are listed in Supplemental Table 1.

EMS assays were used to determine whether a nucleotide change altered the ability of oligonucleotides to bind nuclear proteins. Nuclear extracts were prepared from SK-N-BE(2), U87 MG and U251 cells using the Nuclear Extract Kit (Active Motif, Carlsbad, CA, USA). Biotin-labeled and unlabeled oligonucleotides were synthesized by IDT Integrated DNA Technologies (Corvallis, IA, USA). EMS assays were performed using the LightShift Chemiluminescent EMSA Kit (PIERCE, Rockford, IL, USA). Oligonucleotide sequences used to perform the EMS assays are listed in Supplemental Table 2.

**RESULTS**

Demographic and clinical characteristics of the study subjects

Of the 529 participants in the Mayo PGRN-AMPS study, 15 subjects who were not identified as white and non-Hispanic and 15 non-adherent subjects were excluded, resulting in 499 WNH subjects for the secondary ‘last visit’ GWA analyses. The primary ‘eight week’ GWA analyses were conducted using a subset of 398 WNH subjects who were protocol adherent. Demographic information, baseline clinical characteristics and treatment outcomes for both the 8-week and the last-visit cohorts are summarized in Table 1.
Outcomes for medication-adherent white non-Hispanic subjects included in the last-visit and 8-week analyses

| Characteristic and measure | Last visit | Eight weeks |
|---------------------------|-----------|-------------|
| Age                       | (N = 499) | (N = 398)   |
| Education (years)         | 39.91 ± 13.8 | 40.64 ± 13.5 |
| Female gender             | 14.68 ± 2.4  | 14.91 ± 2.3   |
| QIDS-C baseline           | 312 (62.5%) | 255 (64.1%)  |
| QIDS-C four week          | 15.08 ± 3.47 | 14.99 ± 3.31  |
| QIDS-C week eight         | 8.46 ± 4.41  | 8.34 ± 4.41   |
| Age at onset of first episode of depression | 6.24 ± 4.05 | 6.24 ± 4.05 |
| Recurrent depression (> 1 episode) | 24.05 ± 13.82 | 24.70 ± 14.11 |
| Citalopram                | 483 (96.8%) | 386 (96.9%)  |
| Escitalopram              | 342 (68.8%) | 274 (68.7%)  |
| Remitter (QIDS ≤ 5)       | 206 (41.3%) | 198 (49.7%)  |
| Response (%) reduction QIDS ≥ 50% | 287 (57.5%) | 274 (68.8%)  |

Values are mean (± s.d.) or number of subjects, N (%).

Quality control analyses were performed for 592,236 genotyped SNPs. The genotypes had very high concordance rates in duplicate samples (> 99.99%) and very low rates of Mendelian inheritance errors (< 0.01%). Overall 30,700 SNPs failed genotyping. Another 28,610 SNPs had allele frequencies below the preset threshold of 0.01 and, consequently, were excluded from analysis due to low power. In addition, 49 SNPs were excluded owing to significant departures from Hardy–Weinberg equilibrium (P < 10^-8). The remaining SNPs (n = 532,877) had call rates exceeding 98%. One person who was reported to be female appeared to be genetically male based on non-heterozygosity of X-chromosome SNPs. Identity-by-descent analysis revealed several pairs of subjects that appear to be related (two likely parent–child pairs, one possible pair of siblings and three pairs that may be more distant relatives). Genome-wide heterozygosity was within the expected range.

Genetic association analysis of SSRI treatment outcomes

Separate GWA analyses based on the two primary outcomes of response and remission were conducted for both the 8-week protocol adherent sample (n = 398) and the last-visit sample (n = 499). Four eigenvectors identified through EigenStrat analysis were used as adjusting covariates in the analyses. Although none of these analyses demonstrated genome-wide significance (P < 10^-5), the 25 SNPs with most significant associations with the four outcome phenotypes are listed in Table 2. Manhattan plots of P-values from the analyses for 8-week outcomes are shown in Figure 1. Manhattan plots for last-visit outcomes are included in Supplementary Figure S1.

Eight-week response analyses

GWA analysis of the 8-week response outcome identified a region on chromosome 9 containing three interesting SNPs with P-values < 8.0 x 10^-6. The rs11144870 SNP is located in intron 2 of the riboflavin kinase (RFK) gene. The rs11144905 and rs785916 SNPs were mapped to introns in the adjacent glucosaminy1 N-acetyltransferase 1 (GCNT1) gene and were in linkage disequilibrium with rs11144870 (r^2 = 0.78 and 0.73, respectively).

Eight-week remission analyses

Three SNPs, rs1379887, rs7738598 and rs989040, were located in an intergenic region near the 5-hydroxytryptamine serotonin receptor 18 gene (HTR1B), were associated with 8-week remission at the P < 2 x 10^-5 level. The rs915120 and rs12116187 SNPs in the G protein-coupled receptor (GPCR) kinase gene 5 (GRK5) on chromosome 10 were also identified as candidate SNPs of interest.

Analysis of secondary outcomes: last-visit response and remission

The most significant association in the last-visit response analysis was obtained for rs2248399, which is located in an intergenic region on chromosome 13 (P = 1.0 x 10^-7). This SNP is located ~ 500 kb from the D-amino acid oxidase activator gene (DAOA). The three SNPs near the 5-hydroxytryptamine (serotonin) receptor 18 gene (HTR1B), rs1379887, rs7738598 and rs989040, that were associated with remission in the 8-week analysis, were also identified in the last-visit remission analysis (P-values = 7.9 x 10^-6 1.7 x 10^-5 and 1.8 x 10^-5, respectively). The last-visit remission analysis also identified the rs7439567 (P = 3.1 x 10^-8) and the rs9761827 (P = 2.3 x 10^-5) SNPs in the protocadherin-18 gene (PCDH18) on chromosome 4.

Functional genomics

As described previously, 14 candidate SNPs were chosen for functional studies based on the significance level of their association with treatment outcomes and their location within or near a candidate gene of interest (Figure 2a). The reporter gene assays identified significant differences in the function of wild type (WT) and variant SNP sequences for five of these SNPs based on differences in luciferase activity between pGL3 constructs that contained the sequences in all three CNS-derived cell lines (Figure 2b). When cloned into the pGL3-Promoter vector, sequences containing the SNPs of interest appeared to function as either an enhancer element that significantly increased promoter activity when compared with activity for the pGL3-Promoter construct alone (P-values < 0.05 in both cell lines) or silencer element depending on the cell line used. Results for the neuroblastoma SK-N-BE(2) cells are shown in Figure 2b, and results for U-87 MG and U251 cell lines in Supplementary Figure S2a. The patterns of differences in luciferase activity between WT and variant alleles were consistent for SNPs rs11144870 (RFK), rs11144905 (GCNT1), rs2248399 (DAOA), rs915120 (GRK5) and rs12254134 (GRK5) in all three cell lines tested (P-values < 0.01).

EMS assays identified seven SNPs that altered the ability of these sequences to bind nuclear proteins in SK-N-BE(2) cells. Figure 2c shows the results for EMS assays for these seven SNPs, including rs11144870, rs785916 (GCNT1), rs2248399, rs2248714, rs1998560 (DAOA), rs915120 and rs12254134. All of these SNPs showed significant differences in nuclear protein binding patterns between WT and variant sequence oligonucleotides in SK-N-BE(2) cells. Similar patterns were observed for all of these SNPs in U-87 MG and U251 cells (Supplementary Figure S2b). HTR1B SNP rs7738598 displayed altered DNA–protein binding only in the two neuroblastoma cells (Supplementary Figure S2b), but not in the glioblastoma cells.

Four candidate SNPs, rs11144870 (RFK), rs915120 (GRK5), rs12254134 (GRK5) and rs2248399 (DAOA) showed differences...
### Table 2: Top 25 association results for each of the four outcomes: response at 8 weeks and last visit, and remission at 8 weeks and last visit (based on logistic regression adjusting for four eigenvectors)

| Rank | SNP | Chr. | Position | MAF | Responder/ remitter | Non-responder/ non-remitter | OR | P-value |
|------|-----|------|----------|-----|---------------------|-------------------------------|----|---------|
|      |     |      |          |     |                     |                               |    |         |
| 1    | rs1379887* | 6  | 78407561 | 0.27 | 0.42 | 0.49 | 9.05E-06 | HR1B | flanking_5'UTR |
| 2    | rs915120*  | 6  | 12180103 | 0.32 | 0.48 | 0.58 | 1.15E-05 | GRK5 | flanking_5'UTR |
| 3    | rs10899655 | 11 | 79894667 | 0.11 | 0.23 | 0.41 | 1.62E-05 | NAR2 | flanking_5'UTR |
| 4    | rs773589*  | 6  | 78671581 | 0.52 | 0.36 | 1.87 | 1.67E-05 | HR1B | flanking_3'UTR |
| 5    | rs59840*   | 6  | 78408635 | 0.27 | 0.41 | 0.51 | 1.69E-05 | HR1B | flanking_3'UTR |
| 6    | rs978814   | 6  | 81016631 | 0.37 | 0.52 | 0.54 | 3.82E-05 | BCKDHB | flanking_5'UTR |
| 7    | rs194530   | 12 | 91053739 | 0.37 | 0.52 | 0.47 | 2.93E-05 | RFK | flanking_5'UTR |
| 8    | rs1640617  | 4  | 76814320 | 0.40 | 0.26 | 1.91 | 4.05E-05 | SOX52 | flanking_5'UTR |
| 9    | rs1225413* | 10 | 12116187 | 0.35 | 0.49 | 0.52 | 4.24E-05 | GRK5 | flanking_5'UTR |
| 10   | rs12963758 | 12 | 7814091  | 0.10 | 0.03 | 4.35 | 4.44E-05 | PTPRM | flanking_5'UTR |
| 11   | rs1849899  | 12 | 71909005 | 0.29 | 0.44 | 0.54 | 4.67E-05 | FLJ2264 | flanking_5'UTR |
| 12   | rs1359873  | 13 | 74194810 | 0.36 | 0.26 | 2.01 | 4.76E-05 | LOC40015 | flanking_5'UTR |
| 13   | rs943679   | 1  | 68340006 | 0.16 | 0.28 | 0.47 | 4.81E-05 | GPR177 | flanking_5'UTR |
| 14   | rs318211   | 6  | 8102546  | 0.37 | 0.52 | 0.55 | 3.94E-05 | BCKDHB | flanking_3'UTR |
| 15   | rs415435   | 2  | 96958205 | 0.30 | 0.27 | 0.51 | 3.08E-05 | DHH1 | flanking_5'UTR |
| 16   | rs934638   | 6  | 78438553 | 0.29 | 0.43 | 0.53 | 4.99E-05 | HTR1B | flanking_5'UTR |

### Logistic regression adjusting for four eigenvectors

- **Response at 8 weeks**
- **Response at last visit**
- **Remission at 8 weeks**
- **Remission at last visit**

- **OR** represents the odds ratio for the association.
- **P-value** indicates the statistical significance of the association.

### Additional Information

- **MAF** represents the minor allele frequency.
- **Gene region** indicates the genomic region in which the SNP is located.

---

**Legend:**
- **OR** (Odds Ratio)
- **P-value** (Probability value for the association test)
between WT and variant SNP sequences in both reporter gene and EMS assays performed with all three cell lines.

Replication

The SNPs selected for functional studies based on GWA analyses of the PGRN-AMPS outcomes data were also genotyped using DNA from the STAR*D study. Association analyses were performed for the same four treatment outcomes with 12 of the 14 SNPs, as two SNPs failed genotyping, using DNA from WNH STAR*D subjects who had HAMD-17 scores >14 before SSRI therapy. None of the associations for these SNPs were statistically significant in the STAR*D samples (Supplemental Table S7). The top 25 SNPs reported for the STAR*D GWA analyses were also evaluated for possible association with treatment response in the PGRN-AMPS samples. None of these associations were statistically significant (Supplemental Table S8).

DISCUSSION

Four GWA studies of antidepressant treatment outcomes have been conducted including the PGRN-AMPS study.6–10 None of these GWA analyses were able to identify associations that reached genome-wide statistical significance. Furthermore, no specific SNPs or genes have been associated with outcomes across any two of these studies. Given that replication has not been reported, functional validation studies provide an alternative strategy to take the next step toward deciphering the underlying biology of candidate SNPs that might influence antidepressant response. The association of SNP rs11144870 in RCK1 with SSRI outcomes is intriguing as riboflavin kinase is an essential enzyme that catalyzes the phosphorylation of riboflavin (vitamin B2) to form flavin mononucleotide. This is a critical step for both vitamin B2 metabolism and flavin cofactor synthesis. Riboflavin and its flavin cofactors influence the folate and methionine cycles as riboflavin functions as a cofactor for methylene tetrahydrofolate reductase. Methylene tetrahydrofolate reductase is the enzyme that converts 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, which might indirectly influence the intensity of depressive symptoms and the effects of SSRI therapy.

The association of SNP rs11144870 in RFK with SSRI outcomes is intriguing as riboflavin kinase is an essential enzyme that catalyzes the phosphorylation of riboflavin (vitamin B2) to form flavin mononucleotide. This is a critical step for both vitamin B2 metabolism and flavin cofactor synthesis. Riboflavin and its flavin cofactors influence the folate and methionine cycles as riboflavin functions as a cofactor for methylene tetrahydrofolate reductase. Methylene tetrahydrofolate reductase is the enzyme that converts 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate.19 Previous studies have suggested that components of the folate and methionine cycles may be involved in increasing the risk for developing MDD and might influence treatment outcomes.20–24

### Table 2 (Continued)

| Rank | SNP     | Chr. | Position | MAF         | Association with response/ remission | Gene | Gene region | Distance to gene (bp) |
|------|---------|------|----------|-------------|--------------------------------------|------|-------------|----------------------|
|      |         |      |          | OR P-value  |                                      |      |             |                      |
| 19   | rs757240 | 4    | 7658657  | 0.39 0.39 0.26 0.39 1.89 5.33E–05 |  SORS2        | intron | –32296      |
| 20   | rs2252523| 16   | 1937615  | 0.25 0.25 0.39 0.39 1.89 5.49E–05 |  RPL3L        | intron | –233        |
| 21   | rs9542990| 13   | 7199327  | 0.22 0.22 0.35 0.35 1.89 6.20E–05 |  FLG2264      | flanking, 5'UTR | –206771     |
| 22   | rs9352817| 6    | 8108298  | 0.36 0.36 0.36 0.36 1.89 6.78E–05 |  BCKDHB       | intron | –27502      |
| 23   | rs1936158| 6    | 78241094 | 0.48 0.48 0.35 0.35 1.89 6.89E–05 |  HTR1B        | flanking, 5'UTR | –11255      |
| 24   | rs351294 | 5    | 154428952| 0.05 0.05 0.13 0.13 1.89 7.40E–05 |  MRPL22       | flanking, 3'UTR | –102208     |
| 25   | rs1777770| 7    | 7855148  | 0.15 0.15 0.27 0.27 1.89 7.62E–05 |  HTR1B        | flanking, 5'UTR | –321644     |

Abbreviations: bp, base pair; Chr., chromosome; MAF, minor allele frequency; NA, not available; OR, odds ratio; SNP, single-nucleotide polymorphism.4aSNPs selected for functional genomic studies.
variants in GCNT1 would be of interest (see the Nervous System database, http://www.itb.cnr.it/gncdb/).

A SNP in an intergenic region near the DAOA gene (rs2248399) was identified by these analyses and may be potentially functional significant based on the reporter gene assay. This was identified by these analyses and may be potentially functional significant based on the reporter gene assay. This SNP and two other DAOA SNPs were also shown to have the potential to affect the binding of nuclear proteins. None of these SNPs have been included in previous candidate gene studies of DAOA for schizophrenia or bipolar disease.26–31

Six SNPs that mapped to an intergenic region ~150–500 kb distant from the HTR1B gene were associated with 8–week and last-visit remission (see Table 2 and Supplemental Tables S5 and S6). The HTR1B gene has been reported to be associated with psychiatric phenotypes, as well as response to SSRI treatment.32–35 EMS assays were performed with these three SNPs and with rs7738598 were found to display a difference between WT and variant sequences in nuclear protein binding in the two glioblastoma cells that were tested.

Two of the intronic SNPs in GRK5 (rs915120 and rs12254134) that were associated with remission also altered function. A different member of the G protein-coupled receptor kinase family, GRK2, has been found to be upregulated during antidepressant treatment36 and GRK5 has been shown to regulate GPCRs receptors such as the β1-adrenergic receptor37 and the dopamine D1A receptor.38 GRKS is highly expressed in many tissues, including human heart and brain.39,40 A single functional GRK5 polymorphism, rs17098707, that results in a Gln41Leu change in amino-acid sequence has been reported to regulate cardiac function.41,42 The intronic SNPs identified in this GWA study were not in linkage disequilibrium with the Gln41Leu polymorphism, suggesting that these novel GRKS SNPs might function independently from the Gln41Leu polymorphism. As >90% of GPCRs are expressed in the brain, the identification of functional GRKS SNPs may provide novel directions for future studies of variation in antidepressant response.

In the present study, the selection of the SNPs for functional assessment was limited to those identified during our GWA analyses. Although other experimental approaches are available to assess the functional consequences of genetic variants, the application of these two commonly used in vitro functional assays has highlighted nine candidate SNPs that may be worthy of further mechanistic pursuit using alternative methods. Although many GWA studies have been performed with psychiatric phenotypes, few have identified genomic loci that were replicated and could successfully be used as robust ‘biomarkers’ in clinical psychiatric practice. A lack of reliable model systems for functional genomic studies of the biological mechanism underlying the association is among the factors that have prevented the translation of genomic research to psychiatric practice. The use of induced pluripotent stem (iPS) cells represents a novel and promising tool for functional validation and mechanistic studies of genomic loci identified through GWA studies.43

Limitations of our study also include the fact that detailed information on certain clinical factors, such as comorbid psychiatric diagnoses, was not available. In addition, the influence of potentially important covariates such as drug dose and blood levels has not yet been fully explored. However, important associations between genetic variants and clinical outcomes can be missed by adjusting for covariates such as blood drug levels that serve as intermediate factors. Subsequent analyses will be focused on blood drug levels and their association with genetic variation and treatment outcomes. Our top findings were not replicated by an analysis of samples from the STAR*D study, and we also failed to replicate the best association findings from the STAR*D GWA analyses. Similarly, the PGRN-AMPS analyses did not replicate the rs1126757 in the IL11 gene that was reported to have some association with escitalopram response in the GENDEP project. Finally, the association of the rs6989467 SNP reported for the MARS GWA study was not replicated in the PGRN-AMPS analyses. Several important differences between the studies may have contributed to the lack of replication, including differences in baseline clinical characteristics of participants in the two studies.

GWA studies have been shown to be a powerful approach for the identification of novel genomic markers for disease risk and pharmacogenomic phenotypes.44–46 However, GWA studies have not been as informative when applied to psychiatric phenotypes. None of them has demonstrated genome-wide significant associations nor have these studies replicated each other. The PGRN-AMPS is unique in that it is the only analysis that used blood drug assays to evaluate treatment adherence and is the only study that analyzed the functional significance of identified SNPs of interest. Given the underlying phenotypic heterogeneity (which is typical of studies of psychiatric illnesses), it is particularly challenging to identify homogenous treatment samples. A possible solution is to subclassify patients using biomarkers such as pharmacometabolomic characteristics to identify more biologically similar patients.12

© 2013 Macmillan Publishers Limited

The Pharmacogenomics Journal (2013), 456–463
**CONFLICT OF INTEREST**

Dr Mrazek has developed intellectual property that has been licensed by AssureRx Health, which has been subsequently incorporated into a physician decision support software product. He has also received research funding from AssureRx Health to create and maintain a bibliographic system designed to monitor the scientific literature related to studies addressing psychiatric pharmacogenomic relationships. No other author has any potential conflict of interest to report.

**ACKNOWLEDGEMENTS**

This research was supported, in part, by NIH grants RO1 GM28157, U19 GM61388 (The Pharmacogenomics Research Network), P20 1P20AA017830-01 (The Mayo Clinic Center for Individualized Treatment of Alcohol Dependence), and a PhRMA Foundation Center of Excellence in Individualized Medicine. This publication was supported by NIH/NCRR/NCATS Grant Number UL1 RR024150. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. We thank the staff of the Mayo Clinic Rochester Department of Psychiatry and Psychology for their effort in recruiting the patients to the Mayo PGRN-AMPS, Dr Ryan Abo for his assistance in gene annotation, and Lori Solmonson for her assistance with the preparation of the paper. The authors also wish to thank the STAR*D investigators for both their effort and generosity in sharing the clinical data and DNA samples.

**REFERENCES**

1. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003; 289: 3095–3105.
