Published in final edited form as:

Pharmacogenet Genomics. 2009 January ; 19(1): 1–10. doi:10.1097/FPC.0b013e3283163ecd.

Resequencing of serotonin-related genes and association of tagging SNPs to citalopram response

Eric J. Peters^a, Susan L. Slager^b, Greg D. Jenkins^b, Megan S. Reinalda^b, Holly A. Garriock^a, Stanley I. Shyn^a, Jeffrey B. Kraft^a, Patrick J. McGrath^c, and Steven P. Hamilton^a

^aDepartment of Psychiatry and Institute for Human Genetics, University of California, San Francisco, California

^bDepartment of Health Sciences Research, Mayo Clinic, Rochester, Minnesota

^cDepartment of Psychiatry, Columbia University College of Physicians and Surgeons, and New York State Psychiatric Institute, New York, USA

Abstract

Several reports have been published investigating the relationship between common variants in serotonin-related candidate genes and antidepressant response, and most of the results have been equivocal. We previously reported a significant association between variants in serotonin-related genes and response to the selective serotonin reuptake inhibitor fluoxetine. Here, we attempt to expand upon and replicate these results by (i) resequencing the exonic and putatively regulatory regions of five serotonin-related candidate genes (*HTR1A*, *HTR2A*, *TPH1*, *TPH2*, and *MAOA*) in our fluoxetine-treated sample to uncover novel variants; (ii) selecting tagging single nucleotide polymorphisms (SNPs) for these genes from the resequencing data; and (iii) evaluating these tagging SNPs for association with response to the selective serotonin reuptake inhibitor citalopram in an independent sample of participants who are enrolled in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) clinical study ($N=1953$). None of the variants associated previously with fluoxetine response were found to be associated with citalopram response in the STAR*D sample set. Nor were any of the additional tagging SNPs found to be associated with citalopram response. An additional SNP in *HTR2A* (rs7997012), previously reported to be associated with outcome of citalopram treatment in this sample, but not well tagged by any of the other SNPs we studied, was also genotyped, and was associated with citalopram response ($P=0.0002$), strongly supporting the previous observation in the same STAR*D sample. Our results suggest that resequencing the serotonin-related genes did not identify any additional common SNPs that have not been identified previously. It appears that genetic variation in these five genes has a marginal effect on response to citalopram, although a previously observed association was supported and awaits replication in an independent sample.

Keywords

association; citalopram; fluoxetine; pharmacogenetics; single nucleotide polymorphism

Correspondence to Steven P. Hamilton, MD, PhD, Department of Psychiatry, University of California, San Francisco, 401 Parnassus Avenue, Box NGL, San Francisco, CA 94143, USA; Tel: + 1 415 476 7889; fax: + 1 415 476 7800; SteveH@lppi.ucsf.edu.
Present address of Eric J. Peters is Illumina, Inc., San Diego, California, USA.

Supplementary data : Supplementary materials are available at *The Pharmacogenetics and Genomics Journal* Online (www.pharmacogeneticsandgenomics.com).

Introduction

Major depression is one of the most prevalent psychiatric disorders in the USA [1]. Owing, in part, to their favorable safety profiles, selective serotonin reuptake inhibitors (SSRIs) are the most commonly prescribed antidepressant medications. However, there is large variability in clinical response to treatment with SSRIs. Interindividual variation in serotonin pathway genes may influence response to antidepressant medication [2]. The molecular target of SSRIs, the serotonin transporter, is a clear candidate gene for antidepressant response. Several studies have investigated the influence of a repeat polymorphism (5HTTLPR) in the serotonin transporter gene on antidepressant response, and the results have been equivocal [3,4]. Even if a true association between the 5HTTLPR and antidepressant response exists, the small effect size still leaves a large portion of the variability in response unexplained. The delayed therapeutic onset of SSRIs suggests that antagonizing the serotonin transporter is not sufficient for antidepressant action. Downstream effects such as receptor sensitization or desensitization [5], transcriptional cascades involving BDNF and CREB1 [6], or alterations in neurogenesis could be responsible for the therapeutic actions of SSRIs [7]. There have been far fewer studies investigating the influence of other genes in the serotonin pathway on antidepressant response, and the majority of these studies have focused on a small number of variants per gene [2]. In this study, we investigate the influence of common genetic variation in five serotonin pathway genes (*HTR1A*, *HTR2A*, *TPH1*, *TPH2*, and *MAOA*) on clinical response to the SSRI citalopram. We chose serotonin receptors thought to be involved in antidepressant response (*HTR1A* and *HTR2A*), as well as genes whose products are involved in the catabolism (*TPH1* and *TPH2*) and metabolism (*MAOA*) of serotonin.

In an earlier indirect association study, we genotyped evenly spaced, largely noncoding, publicly available single nucleotide polymorphisms (SNPs) ($N=68$) in these five serotonin genes in a well-characterized sample ($N=96$) of depressed individuals taking the SSRI fluoxetine [8]. Several variants were significantly ($P < 0.05$) associated with fluoxetine response and response specificity, including variants in *HTR2A*, *TPH1*, and *TPH2*. These initial results prompted this study, which has two main goals. The first is to follow-up on the indirect association study by resequencing all exons, intron–exon boundaries, and 5' conserved noncoding sequence (CNS) of these five genes in the same depressed population taking fluoxetine. By doing this, we hoped to uncover potentially functional variants in coding and regulatory regions as well as any informative tagSNPs that were not genotyped in our initial study. The second goal of this study is to use this set of tagSNPs to explore the role of common variation in these five candidate genes on antidepressant response in a larger independent clinical sample of patients taking the SSRI citalopram. This assumes that different SSRIs share the same mechanism of action. Although this is unknown, as all SSRIs block the serotonin transporter, it seems reasonable to assume that there may be commonality in their therapeutic actions. For this purpose, we used several approaches for identifying potential tagging SNPs. We then genotyped a subset of patients ($N=1914$) from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) [9] study to evaluate association with treatment response.

Methods

Fluoxetine sample

Our sample taking fluoxetine consisted of 96 participants enrolled in an NIMH-funded protocol (P.J.M., principal investigator) to assess relapse after fluoxetine discontinuation in depressed patients who had responded to fluoxetine. Inclusion in that clinical trial required participants aged 18–65 years to be in a current episode of major depression; there was no depression severity threshold for inclusion. The Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition* Axis I Disorders – Patient Edition (SCID-I/P) was used to establish all psychiatric diagnoses [10]. Patients were treated for 12 weeks

with open-label fluoxetine, and the majority (78%) of the patients were Caucasian. For a detailed description of this study population, see Peters *et al.* [8].

The Sequenced Treatment Alternatives to Relieve Depression sample

Of the 4041 participants, DNA was obtained from 1953 participants as part of the NIMH Human Genetics Initiative. The design of STAR*D was to enroll adults experiencing a major depressive episode (MDE) who exhibited neither an inadequate response nor intolerance to an adequate trial of any of the STAR*D protocol treatments during the current episode. The overall aim of STAR*D (A. John Rush, principal investigator) was to prospectively determine which of a number of treatments are beneficial for patients experiencing an unsatisfactory clinical outcome after treatment with citalopram. As the STAR*D design trial has been described extensively, we briefly summarize it here [9,11,12]. To make the findings as generalizable as possible, STAR*D used broad inclusion criteria and enrolled a diverse population, including good minority representation. Diagnoses were made using the Psychiatric Diagnostic Screening Questionnaire, whereas depressive symptoms were assessed with the 16-item Quick Inventory of Depressive Symptomatology collected at clinic visits [Self-Report (QIDS-SR) version] [13]. The QIDS-SR is highly correlated with the 17-item Hamilton Rating Scale for Depression (HRSD₁₇), and scores can be converted readily between the two instruments [14]. Participants meeting criteria and providing consent were administered citalopram as the initial treatment. The protocol encouraged 12 weeks of treatment with vigorous dosing of open-label citalopram (20–60 mg/day). The subsample of 1953 participants who consented to provide DNA samples was 61.8% females and 38.2% males, and with ethnic proportions of 78.1% Caucasians, 16.1% African-Americans, 3.5% multiracial, 1.1% Asian, 1.2% Pacific Islander/Native American, and 0.1% unspecified. Furthermore, 14.0% of the sample reported being Hispanic, and 43.5% of the sample came from primary care clinics, with the remaining 56.5% coming from specialty clinics. DNA from 1914 participants is currently available.

Access to the DNA samples and clinical data was approved by the STAR*D Ancillary Studies Committee, and clinical data were obtained from the Data Coordinating Center of STAR*D. Approval to carry out this work was obtained by the Committee on Human Research at the University of California, San Francisco.

Phenotypic definitions

For the citalopram treatment response analysis in the STAR*D sample, we define five interrelated response phenotype definitions of response (response, nonresponse, remission, persistent response, and nonpersistent response). The first two are responders and nonresponders: responders are patients who had at least 42 days of treatment and whose QIDS-SR on their final clinical visit shows $\geq 50\%$ reduction in score; the remaining patients, who had at least 42 days of treatment, were then considered nonresponders. The $\geq 50\%$ reduction in symptom severity on the HRSD₁₇ is the conventional definition of response in clinical trials. We used the QIDS-SR score to estimate severity as all participants had these data. We required this 42-day (or 6-week) threshold to ensure an adequate exposure to citalopram and to enhance the power to find associations between genotype and response by reducing potential heterogeneity [3]. Using this threshold, we found no statistical difference in the average total dosage of citalopram between those who were on the trial for at least 42 days (average total dosage=29.88 mg) and those who were not (average total dosage=30.43 mg). The 254 patients with less than 42 days of treatment were excluded from the analysis. The third phenotype definition is remission. Remission was defined as a QIDS-SR score ≤ 5 , which closely corresponds to the conventional definition of an HRSD score of ≤ 7 [14]. The final two phenotypes are based on our attempt to further reduce heterogeneity by attempting to separate placebo response from true drug response in antidepressant trials, as described above [15]. For these phenotypes in the STAR*D, a specific pattern of response was defined by persistence,

or the maintenance of response for the remainder of the study once it was attained. Earlier studies have defined 'specific' patterns to be further characterized by delayed response, that is, after the first 2 weeks [16,17]. We were unable to use this criterion because the STAR*D study design did not include ratings before week 2. We therefore defined 'persistent' responders as those patients who had a sustained response at all consecutive visits after the first visit with response, as measured by $\geq 50\%$ reduction in QIDS-SR scores. Those whose response occurred only at the last visit were removed from the analysis. In contrast, 'nonpersistent' responders were those participants who did not maintain their response after the first visit with a response. Note that persistent and nonpersistent responders are a subset of responders (as defined by the response phenotype above). Moreover, because visits were at least 2 weeks apart, we assumed that intervening weeks were characterized by the response defined by the previous visit. We compared persistent responders with nonresponders, allowing us to test the hypothesis that the persistent response to citalopram represented a more genetically homogenous group of persons taking citalopram. We also compared persistent responders with nonpersistent responders to test whether there are genetic differences between 'true drug' responders and 'placebo' responders, as suggested in our previous work [8].

DNA amplification

PCR reactions of 5 μ l containing 200 nmol/l of the forward and reverse primers, 10 ng genomic DNA template, 50 μ mol/l dNTPs (Roche, Indianapolis, Indiana, USA), 1 mol/l anhydrous betaine (Acros Organics, Geel, Belgium), 50 mmol/l KCl, 20 mmol/l Tris-HCl (pH 8.4), 2.5 mmol/l $MgCl_2$, and 0.25 U Platinum Taq DNA polymerase (Invitrogen, Carlsbad, California, USA). Samples were cycled using a touchdown protocol at 94°C for 3 min, followed by seven cycles of 94°C for 30 s, 65–59°C for 30 s (decreased by 1°C intervals per cycle), and 72°C for 30 s, followed by 38 cycles of 90°C for 30 s, 58°C for 30 s, and 72°C for 30 s, with a final 10 min at 72°C. The reactions were performed on an Applied Biosystems GeneAmp PCR System 9700 (Foster City, Indiana, USA) using 384-well plates (MJ Research, Waltham, Massachusetts, USA).

DNA resequencing and variant discovery

PCR primers were designed to amplify all exons and approximately 20 bp of flanking intronic sequence of the five genes (*MAOA*, *HTR1A*, *HTR2A*, *TPH1*, and *TPH2*). We also sequenced 5' proximal CNS identified using mammalian global sequence alignments. For this study, we used the VISTA genome browser and searched 5 kb of DNA 5' proximal of the transcriptional start site and defined CNS as sequence that had at least 70% identity in both human–mouse and human–rat comparisons (minimum window size=100 bp). Primers were designed using Primer3 software and were manufactured by Invitrogen (Carlsbad, California, USA). Sequencing primer sequences and amplicon information are shown in supplemental Table 1. After PCR, 1 μ l of PCR reaction was run on a 1% agarose gel to confirm the specificity of the reaction and that product was of expected size. Unsuccessful reactions were redesigned with different primers. PCR products were prepared for sequencing by incubating product with exonuclease I (0.5 U/sample) and shrimp alkaline phosphatase (2.5 U/sample) at 37°C for 90 min and inactivating the enzymes at 95°C for 15 min.

DNA sequencing of PCR product as template was performed using BigDye cycle sequencing on an ABI 3730x1 DNA analyzer. All amplicons in all samples ($N=95$) were sequenced in at least one direction, with the optimal strand choice based on sequencing both strands of a subset of these samples ($N=6$). Sequencing traces were analyzed for polymorphisms using Mutation Surveyor v2.4.1 (Soft Genetics, State College, Pennsylvania, USA) and confirmed by visual inspection. Variants identified on only one or two chromosomes in our population were confirmed by sequencing the opposite strand.

tagSNP selection

All Caucasian samples in our fluoxetine treatment population ($N=75$) were used in tagSNP selection, regardless of response phenotype. We sought to compare several different approaches for tagSNP selection. We used the method of Carlson *et al.* [18] ('ldselect.pl'), which is based on pairwise linkage disequilibrium (LD) between SNP markers and does not require haplotype block-like structure. We ran this algorithm to group SNPs into high LD 'bins', which contain SNPs with a minimum pairwise r^2 of 0.80. A single tagSNP from each 'bin' was then chosen to represent that group of SNPs in future studies (supplemental Table 3). We were also interested in testing for haplotype-specific effects, thus we wanted to explore how accurately the tagSNPs chosen using a pairwise r^2 threshold could reconstruct the underlying haplotypes. To limit haplotype diversity and facilitate haplotype inference, haplotype blocks were chosen using the program HaploBlockFinder [19]. We defined haplotype blocks as regions in which at least 80% of chromosomes assayed were represented by three or less common haplotypes. Within each haplotype block, we then used two haplotype tagging approaches. We used the method of Stram *et al.* [20] ('TagSNPs'), which selects haplotype tagging SNPs to optimize R_h^2 , which is the squared correlation between estimates of the number of copies of a particular haplotype h (inferred using only the tagSNPs) and the true number of copies of haplotype h carried by a subject (inferred using the entire SNP set), averaging over all possible genotype data under an assumption of Hardy-Weinberg equilibrium. We also selected haplotype tagging SNPs using the pattern recognition approach of Ke and Cardon [21], which does not account for the uncertainty of inferring haplotypes for genotypic data. Using these three sets of tagging SNPs, we then calculated the minimum R_h^2 for all common (<5%) haplotypes within each haplotype block to assess how well these sets predict the underlying haplotype structure.

tagSNP genotyping

Genotyping of tagSNPs was performed using either fluorescence polarization detection of template-directed dye-terminator incorporation (FP-TDI) or 5' exonuclease fluorescence assay (TaqMan) assays (supplemental Table 3). For tagSNPs genotyped by FP-TDI, after PCR, the excess primers, deoxynucleotides, and pyrophosphate in the PCR reaction were degraded by adding 0.1 μ l of 10 $\times$ PCR Clean-Up Reagent, containing a mixture of shrimp alkaline phosphatase and exonuclease I (PerkinElmer, Wellesley, Massachusetts, USA), 0.1 μ l of inorganic pyrophosphatase (Roche Applied Science, Indianapolis, Indiana, USA), and 0.8 μ l of PCR Clean-Up Dilution Buffer to each 5 μ l PCR reaction (PerkinElmer). The mixture was then incubated at 37°C for 60 min, followed by inactivation for 15 min at 80°C. The final step was the addition of a 4 μ l solution containing a final concentration of 0.5 μ mol/l TDI probe, 1 μ l of 10 $\times$ TDI Reaction Buffer, 0.5 μ l of AcycloTerminator Mix (containing R110 and TAMRA-labeled AcycloTerminators, corresponding to the polymorphic base), and 0.025 μ l of AcycloPol DNA polymerase (PerkinElmer). This mixture was cycled at 95°C for 2 min, followed by 25 cycles of 94°C for 15 s and 55°C for 30 s. After template-directed incorporation, fluorescence polarization was read using a VICTOR2 1420 Multilabel Counter (PerkinElmer), and genotypes were scored using custom software.

For tagSNPs genotyped using 5' exonuclease fluorescence (Taqman) assays, 5 μ l reactions containing 10 ng of dried genomic DNA template, 2.5 μ l of Universal Taqman PCR Master Mix (Applied Biosystems), 0.085 μ l of 20 $\times$ Taqman assay probe (Applied Biosystems), and 2.42 μ l of sterile H₂O were cycled at 95°C for 10 min, followed by 40 cycles of 92°C for 15 s and 60°C for 1 m. Reaction fluorescence was read and genotypes were scored on an ABI 7900HT Sequence Detection System (Applied Biosystems). Genotype data have been deposited with the NIMH Center for Collaborative Genetic Studies on Mental Disorders (<http://nimhgenetics.org/>).

A subset of STAR*D samples ($N=90$) were genotyped twice to determine assay reproducibility. Out of a total of 2348 successful comparisons across the 28 tagSNPs, we observed 21 discrepant genotypes (99.1% reproducibility).

Statistical analysis

The frequency distributions of demographic and clinical variables were examined in the STAR*D sample and by the five phenotypes. To reduce the potential for population stratification, analyses were carried out using principal components analysis (PCA) using a set of ancestry informative markers [22], as described below. For these analyses, all samples used were self-identified as Caucasian or African-American. Other racial categories were not considered because of the small numbers of those samples. We tested for Hardy-Weinberg equilibrium within each of the Caucasian and African-American groups, and all participants from a stratum were used in the analysis as all participants had depression and the evaluated polymorphisms were not suspected to influence risk of depression.

We used unconditional logistic regression analysis to examine associations of the SNPs and each of the four phenotypic comparisons. These comparisons are (i) responder versus nonresponder, (ii) remitter versus nonresponders, (iii) persistent responders versus both nonresponders and nonpersistent responders, and (iv) persistent responders versus nonpersistent responders. Each polymorphism was modeled individually as gene-dosage effects in the regression models. This coding scheme was chosen because of its robustness to departure from the true additive genetic model [23]. Regression analyses were performed either unadjusted or adjusted for potential confounding effects, including sex, age, education (years of school), months in current MDE, and years since first MDE. We found that adjustment for potential confounders did not significantly influence the results, thus the values reported here are unadjusted. We estimated odds ratios (ORs) and 95% confidence intervals (CIs) for the carriers of the minor allele versus noncarriers of the minor allele. Owing to the large number of statistical tests, significance threshold was set at 0.01, and permutation tests were performed on any test that resulted in an asymptotic P value of 0.01 or less. This significance level was chosen as a way to balance wanting to reduce type I error in the face of multiple comparisons with investigating genes for which earlier evidence of association has been reported. To adjust for population stratification, covariates were created from SNPs using PCA. PCA is a multivariate method that summarizes all SNPs included in the PCA into several linear combinations of the count of rare alleles of SNPs (principal components or PCs). These PCs are linearly independent of each other, and are given consecutive numbering (i.e. PC#1, PC#2, etc.) based on how much variation in SNP rare allele count they explain. There were 124 SNPs included in the ancestral informative marker panel on 1804 individuals. One hundred and seventy-three individuals were eliminated from the analysis because of low call rates ($<90\%$), ending with 1631 people, five SNPs were then removed owing to low call rates ($<95\%$) to end at 119 SNPs used in the PCA (supplemental Methods, supplemental Table 4, and supplemental Figure 1). PCs were created using S-PLUS v8.0.1 (Tibco, Palo Alto, California, USA) using the princomp function. The first 10 PCs were taken to represent population stratification in analyses, and used as covariates in the association analysis.

Association between haplotypes within the haplotype blocks and citalopram response was calculated using a score test implemented in the computer program HAPLO.SCORE [24]. This test uses the expectation-maximization algorithm to estimate the posterior probability of each person's haplotype. These posterior probabilities are then used to calculate a person's expected haplotype score in the logistic regression analyses. All haplotypes with frequencies greater than 0.01 were simultaneously tested in the analysis. Global P values and individual haplotype P values were obtained. Statistical tests were performed in SAS version 8.2 or Splus version 6.2.1 statistical packages.

Results

DNA resequencing

To expand on the results of our previous association study using evenly spaced publicly available SNP markers, we sequenced the coding region, intron–exon boundaries, and 5' flanking CNS of five serotonin pathway genes (*HTR1A*, *HTR2A*, *TPH1*, *TPH2*, and *MAOA*) in a subset ($N=95$) of the clinical sample treated with fluoxetine. This sample set consisted of participants of primarily Caucasian descent, as identified by self-report. With this sample set, and assuming a binomial distribution, we had an approximately 85% likelihood to detect autosomal (*HTR1A*, *HTR2A*, *TPH1*, and *TPH2*) gene variants at 1% minor allele frequency (MAF) and 77% likelihood to detect X-linked (*MAOA*) gene variants at 1% MAF. We had greater than 99.9% likelihood to detect variants with 5% MAF in all genes investigated. We sequenced approximately 25 kb of DNA per participant and identified 93 variants (supplemental Table 2). Most of the variants identified were rare, in fact 40 were singletons, where the minor allele was seen on only a single chromosome. We identified 36 SNPs that were novel and not in public databases; however, all of these were uncommon ($<10\%$ MAF) in our population. A total of 57 SNPs are currently in database SNP, with an average MAF in our sample of 0.172. We observed only one common amino-acid altering variant (His452Tyr in *HTR2A*, rs6314, MAF=0.11) and six out of the nine nonsynonymous variants we observed were seen as singletons. No nonsynonymous variants were detected in the *TPH1* gene and only one was observed in the *TPH2* gene (Ser41Tyr) as a singleton. We did not observe the Arg441His variant in *TPH2* reported by another group to be common in depression [25], which we had greater than 99% power to detect at least once.

Through resequencing these candidate genes, we genotyped by direct sequencing 16 SNPs that were initially genotyped using FP-TDI in our previous study of publicly available SNPs in these genes [8]. We found two out of 1520 matched calls were inconsistent between genotyping using FP-TDI and sequencing, yielding a genotyping accuracy of 99.87%. This value is reassuring given the substantial effect genotyping errors can have on haplotype inference [26].

tagSNP selection

Although it is common to select tagging SNPs from HapMap data, we sought to empirically examine patterns of variation in our fluoxetine clinical sample. We used the combined genotype dataset of genetic variant information from our current resequencing and previous genotyping efforts to select maximally informative tagSNPs for genotyping in our STAR*D clinical sample. The combined genotype data consisted of 145 SNPs; 77 from resequencing, 52 from our previous genotyping effort, and 16 that were assayed using both methods. For tagSNP selection, we only used data from the Caucasian individuals ($N=75$) in our patient population taking fluoxetine because they consisted of 78% of the sample. We found that tagging rare SNPs resulted in the majority of these rare SNPs being selected as tagSNPs (Fig. 1). The inclusion of rare ($<10\%$ MAF) SNPs almost doubled the number of tagging SNPs selected (28 for common SNPs vs. 51 for all SNPs). Similarly, for haplotype tagging methods such as the approach of Stram *et al.*, tagging rare recombinant haplotypes ($<5\%$ population frequency) required large increases in number of tagging SNPs. In several instances, rare haplotypes were not tagged accurately (i.e. $R_h^2 < 0.8$) even when all SNPs within a block were selected as tagging SNPs (data not shown). We thus limited our analysis of tagSNP performance to the tagging of common SNPs and common haplotypes ($>10\%$ population frequency, $N=71$ SNPs).

In addition to serving as proxies for other SNPs, we sought to investigate how accurately these tagSNPs would infer the underlying haplotype structure. To aid haplotype tagging SNP selection, we selected regions of limited haplotype diversity, also known as haplotype blocks.

In total, eight haplotype blocks were identified, with an average length of 33 kb, using the chromosomal coverage criteria of Zhang *et al.* The *HTR2A* gene was composed of three blocks, whereas the *TPH2* gene had two blocks and the three other genes were all captured by a single haplotype block.

As shown in Fig. 2, tagSNPs selected using the r^2 criteria of Carlson *et al.* predicted all the common haplotypes well in all blocks investigated (> 0.75 average minimum R_h^2). In contrast, tagSNPs chosen using the pattern recognition approach of Ke *et al.* do not predict all common haplotypes well for four of the eight blocks (< 0.60 average minimum R_h^2). The three methods varied in their efficiency in reducing number of tagSNPs to be genotyped. The pattern recognition approach of Ke *et al.* was the most efficient, and selected 16 tagSNPs for the five genes out of a total of 71 common SNPs. The Stram *et al.*'s method selected 21 tagSNPs, and the Carlson *et al.*'s method selected 28 tagSNPs overall. Given that using a pairwise r^2 criteria is analytically straightforward and that the tagSNPs chosen will be sufficient in accurately reconstructing the underlying common haplotypes, we used the 28 tagSNPs selected by the method of Carlson *et al.* for genotyping in our STAR*D clinical population taking citalopram (supplemental Table 3).

tagSNP association to citalopram response

We sought to test two main hypotheses with the tagSNPs selected in these candidate genes using the response phenotypes described above. We first tested whether these variants were associated with overall response to citalopram by comparing responders to nonresponders and remitters to nonresponders. We also tested whether these variants affect persistent drug response by comparing persistent responders to nonresponders and persistent responders to nonpersistent responders. All analyses were corrected for population stratification using PCA based on genotype data from 119 ancestry informative markers. Finally, we tested two additional SNPs, rs7997012 and rs4570625, not for their tagging properties but because of earlier reports associating them with relevant phenotypes.

Table 1 shows the single locus association results with P value of less than 0.01 for our primary phenotypic comparisons; remission and response versus nonresponse. The only tagSNP to meet this threshold was an intronic SNP in the *HTR2A* gene (rs1923884) in the remitter versus nonresponder phenotype ($P=0.01$, OR: =0.72, 95% CI: 0.55–0.95). An SNP subsequently chosen for its previously reported association with citalopram in the same dataset, rs7997012, also showed significant association in our analysis (Table 1 and supplemental Table 5). No tagSNPs met our threshold for significance in our subtype drug response comparisons; persistent responders versus nonresponders, nor in our persistent responders versus nonpersistent responders (supplemental Table 5) comparison.

As expected, based on our tagSNP ascertainment, allelic association between the variants was low in our citalopram population, with no pairwise r^2 value greater than 0.8. We inferred haplotypes within each haplotype block in the five candidate genes and tested them globally for association to citalopram response. We identified a single block in *HTR1A*, three blocks in *HTR2A*, one block in *TPH1*, and two blocks in *TPH2*. The most significant haplotypic association occurred in *HTR2A*, involving a five-SNP block including rs7997012 (global $P=0.0004$, remission phenotype, supplemental Table 6).

Discussion

In this study, we sought to test whether DNA variation in five serotonergic candidate genes is associated with clinical response to SSRI treatment. No variants met our significance threshold in our phenotypic comparisons investigating association with categorical response or persistent

drug response. Similar null results were obtained using inferred haplotypes within these genes. Given that we had adequate power to detect reasonable effect sizes, it appears that variation in these five genes does not significantly influence patient response to citalopram. We chose the most intuitive and clinically meaningful phenotypes for analysis, namely categorical response and remission. There are multitudinous potential analyses that can be carried out given the richness of the phenotypic data. The availability of our genotype data at the NIMH Center for Collaborative Genetic Studies on Mental Disorders facilitates additional exploratory hypothesis testing.

We had previously reported an association to fluoxetine response for several of these variants, but none of them were significantly associated with citalopram response in our current STAR*D sample. Several factors could explain this difference, including different underlying mechanisms of action for the two drugs, differences in patient ascertainment and response definition between the two studies, cryptic population stratification, or simply type I error. We attempted to control for population stratification in this study by analyzing the data within self-identified ethnic groups, as this has been shown to correlate well with marker allele frequencies [27].

One strength of this study is the extensive approach we took to selecting tagSNPs within these candidate genes using dense marker data in our fluoxetine sample. Resequencing of these genes did not identify additional common SNPs that were not in public databases, which indicates that the variant databases are now complete enough to make resequencing to discover new common tagSNPs in the exonic regions of these five serotonin genes unnecessary in Caucasian populations. Thorough resequencing in diverse populations would likely add previously unknown variants, and LD patterns for each ethnic group studied still need to be determined empirically before selecting tagSNPs. Of the 93 SNPs we encountered by resequencing, 35 have been genotyped in the HapMap Project, and we found high correlation between allele frequencies between populations African and European ancestry ($r^2=0.78$, average difference in allele frequency=0.12, supplemental Table 7). Naturally, relevant variation could exist in deep intronic regions or in distal intergenic regions that we did not sequence. Our LD-based coverage of variants within African-American samples in this study was limited, owing to the fact that our discovery sample was enriched with Caucasian samples, and thus our tagSNPs were biased for coverage in Caucasian samples. Furthermore, we did not attempt to tag or genotype less common ($MAF<0.1$) variants in the STAR*D population and are assuming a common phenotype common variant framework, and it remains to be determined if the uncommon, but possibly functional variants, that we describe here have a role in treatment response. Practically, we have limited statistical power to detect statistically significant associations with rare variants even in the large ($N>1900$) STAR*D sample. Some studies, however, suggest that rare variants may play a role in complex phenotypes [28,29].

We did not resequence all intronic or intergenic regions of these genes, and thus cannot rule out variant associations in these regions that were not captured by our tagSNPs. By focusing our 5' proximal putative promoter resequencing on only CNS regions with greater than 70% identity between human and rodent species, we assumed that human promoter regions are necessarily conserved across species, which may not be the case for human-specific regulation. Furthermore, this creates gaps in our resequencing coverage of the 5' proximal regions that may or may not be tagged by other variants. For instance, a variant in the putative promoter region of TPH2 (rs4570625) [30,31] has been shown to affect gene function. This variant, however, is not in a human-rodent conserved sequence. Owing to the putative function of this variant, it was genotyped in the STAR*D sample and we did not detect associations with citalopram response or tolerance ($P>0.05$, supplemental Table 5).

Our comparison of haplotype and SNP tagging methods revealed that the methods vary in the number of tagSNPs selected and their ability to accurately reconstruct common haplotypes. The pattern recognition approach to selecting haplotype tagging SNPs was the most efficient at reducing genotyping costs, however, within several blocks the tagSNPs could not accurately reconstruct the underlying haplotypes. This is not surprising given that this method does not take into account the uncertainty involved in predicting haplotypes from genotypic data. The R_h^2 method of Stram *et al.* accounts for this uncertainty, and selected a greater number of tagSNPs to be genotyped. The pairwise r^2 SNP tagging method of Carlson *et al.* was the least efficient at reducing genotyping costs, although this approach has become the standard approach for selecting tagging SNPs. We were, however, able to show that the tagSNPs selected using pairwise r^2 measures were sufficient to tag the common haplotypes as well. Given the computational simplicity of this method and as it allows us to test for both single locus and haplotypic effects, we chose to use the tagSNPs selected by this method in our STAR*D sample.

Although we have attempted to capture the majority of common variation within these genes, current genotyping costs prohibit complete ascertainment of all variants. Interestingly, a recent study by McMahon *et al.* that used the same STAR*D clinical population and genotyped 768 variants in 68 genes including 37 SNPs in the five genes reported here, with an overlap of seven of the 27 SNPs used in our study. Interestingly, their and our results are largely congruent, in that none of these genes were associated with citalopram response, with one exception. The authors reported highly significant associations between a SNP in the *HTR2A* gene (rs7997012) and several phenotypes, including citalopram response, citalopram remission, and a continuous measure of symptom reduction [32]. This intronic SNP was not initially included in our study, and was not in significant LD with any of our tagSNPs within this gene (max $r^2=0.19$). In fact, phase 2 HapMap data show that only one variant (rs9567732, 11kb 3' of rs7997012) within 1 Mb of this SNP has moderate ($r^2 > 0.80$) LD with it. Analysis of the rs7997012 SNP using our phenotypic definitions and statistical methods also showed moderate association with citalopram response ($P<0.0002$, OR: 1.43, 95% CI: 1.13–1.81), similar to the level of significance reported in the other study for response (reported as $P=0.00004$ in the entire sample). Interestingly, response and remission rates and the minor allele frequency of rs7997012 are strongly correlated with self-reported race in the STAR*D sample. Nevertheless, despite differences in experimental design, it is very promising that in the same sample, the same SNP in *HTR2A* shows a positive association with citalopram response. Our remission phenotype also showed strong association at this marker. As always, replication of this finding in an independent sample is critical, and as the effect size of the variant was small, additional risk variants are required to use genetic information in clinical decision making.

In summary, our study has attempted to broadly investigate these five serotonin genes for association to citalopram response in a large patient sample. Using both single locus and haplotype tests, none of the polymorphisms we interrogated appear to be strongly associated with citalopram response or response specificity/persistence, although our work provides some support for a previous finding in the STAR*D sample for a single SNP in *HTR2A* [32]. Thus far, genetic variation has not been shown to be a consistent clinical predictor of SSRI response and it is possible that many genes of small effect may be involved. However, given that little is known about exactly how SSRIs exert their antidepressant effects *in vivo*, interrogation of DNA variation in other neuronal pathways [33] or across the entire genome may be required to gain further insight.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institute of Mental Health (MH072802) and NARSAD Young Investigator Award (S.P.H.), HHMI Pre-Doctoral Fellowship (E.J.P.), and by the State of New York, which provided partial support to P.J.M. for this work, and grant CA 94919 (S.L.S.) from the National Cancer Institute. The authors appreciate the efforts of the STAR*D Investigator Team for acquiring, compiling, and sharing the STAR*D clinical dataset. STAR*D was funded by NIMH through contract (N01MH90003) to the University of Texas Southwestern Medical Center at Dallas (A. John Rush, P.I.). The authors also thank Stephen Wisniewski, Ph.D., Director, STAR*D Data Coordinating Center, University of Pittsburgh for demographic data. They also would like to acknowledge Dr Francis McMahon for helpful discussions.

References

1. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry* 2005;62:617–627. [PubMed: 15939839]
2. Serretti A, Benedetti F, Zanardi R, Smeraldi E. The influence of Serotonin Transporter Promoter Polymorphism (SERTPR) and other polymorphisms of the serotonin pathway on the efficacy of antidepressant treatments. *Prog Neuropsychopharmacol Biol Psychiatry* 2005;29:1074–1084. [PubMed: 15939518]
3. Kraft JB, Peters EJ, Slager SL, Jenkins GD, Reinalda MS, McGrath PJ, et al. Analysis of association between the serotonin transporter and antidepressant response in a large clinical sample. *Biol Psychiatry* 2007;61:734–742. [PubMed: 17123473]
4. Smits KM, Smits LJ, Schouten JS, Stelma FF, Nelemans P, Prins MH. Influence of SERTPR and STin2 in the serotonin transporter gene on the effect of selective serotonin reuptake inhibitors in depression: a systematic review. *Mol Psychiatry* 2004;9:433–441. [PubMed: 15037864]
5. Blier P, DE MC, Chaput Y. A role for the serotonin system in the mechanism of action of antidepressant treatments: preclinical evidence. *J Clin Psychiatry* 1990;51(Suppl):14–20. [PubMed: 2157700]
6. Shirayama Y, Chen AC, Nakagawa S, Russell DS, Duman RS. Brain-derived neurotrophic factor produces antidepressant effects in behavioral models of depression. *J Neurosci* 2002;22:3251–3261. [PubMed: 11943826]
7. Santarelli L, Saxe M, Gross C, Surget A, Battaglia F, Dulawa S, et al. Requirement of hippocampal neurogenesis for the behavioral effects of antidepressants. *Science* 2003;301:805–809. [PubMed: 12907793]
8. Peters EJ, Slager SL, McGrath PJ, Knowles JA, Hamilton SP. Investigation of serotonin-related genes in antidepressant response. *Mol Psychiatry* 2004;9:879–889. [PubMed: 15052272]
9. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry* 2006;163:28–40. [PubMed: 16390886]
10. First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. Structured Clinical Interview for DSM-IV Axis I Disorders-Patient Edition (SCID-I/P). Version 2.0 edition. New York: Biometrics Department, New York State Psychiatric Institute; 1995.
11. Fava M, Rush AJ, Trivedi MH, Nierenberg AA, Thase ME, Sackeim HA, et al. Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study. *Psychiatr Clin North Am* 2003;26:457–494. [PubMed: 12778843]
12. Rush AJ, Fava M, Wisniewski SR, Lavori PW, Trivedi MH, Sackeim HA, et al. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Controlled Clinical Trials* 2004;25:119–142. [PubMed: 15061154]
13. Zimmerman M, Mattia JI. The reliability and validity of a screening Questionnaire for 13 DSM-IV Axis I disorders (the Psychiatric Diagnostic Screening Questionnaire) in psychiatric outpatients. *J Clin Psychiatry* 1999;60:677–683. [PubMed: 10549684]
14. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry* 2003;54:573–583. [PubMed: 12946886]

15. Ross DC, Quitkin FM, Klein DF. A typological model for estimation of drug and placebo effects in depression. *J Clin Psychopharmacol* 2002;22:414–418. [PubMed: 12172342]
16. Quitkin FM, Rabkin JD, Markowitz JM, Stewart JW, McGrath PJ, Weisstaub N, et al. Use of pattern analysis to identify true drug response. A replication. *Arch Gen Psychiatry* 1987;44:259–264. [PubMed: 3548638]
17. Stewart JW, Quitkin FM, McGrath PJ, Amsterdam J, Fava M, Fawcett J, et al. Use of pattern analysis to predict differential relapse of remitted patients with major depression during 1 year of treatment with fluoxetine or placebo. *Arch Gen Psychiatry* 1998;55:334–343. [PubMed: 9554429]
18. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 2004;74:106–120. [PubMed: 14681826]
19. Zhang K, Jin L. HaploBlockFinder: haplotype block analyses. *Bioinformatics* 2003;19:1300–1301. [PubMed: 12835279]
20. Stram DO, Haiman CA, Hirschhorn JN, Altshuler D, Kolonel LN, Henderson BE, et al. Choosing haplotype-tagging SNPs based on unphased genotype data using a preliminary sample of unrelated subjects with an example from the Multiethnic Cohort Study. *Hum Hered* 2003;55:27–36. [PubMed: 12890923]
21. Ke X, Cardon LR. Efficient selective screening of haplotype tag SNPs. *Bioinformatics* 2003;19:287–288. [PubMed: 12538253]
22. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38:904–909. [PubMed: 16862161]
23. Freidlin B, Zheng G, Li Z, Gastwirth JL. Trend tests for case-control studies of genetic markers: power, sample size and robustness. *Hum Hered* 2002;53:146–152. [PubMed: 12145550]
24. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 2002;70:425–434. [PubMed: 11791212]
25. Zhang X, Gainetdinov RR, Beaulieu JM, Sotnikova TD, Burch LH, Williams RB, et al. Loss-of-function mutation in tryptophan hydroxylase-2 identified in unipolar major depression. *Neuron* 2005;45:11–16. [PubMed: 15629698]
26. Kirk KM, Cardon LR. The impact of genotyping error on haplotype reconstruction and frequency estimation. *Eur J Hum Genet* 2002;10:616–622. [PubMed: 12357332]
27. Tang H, Quertermous T, Rodriguez B, Kardia SL, Zhu X, Brown A, et al. Genetic structure, self-identified race/ethnicity, and confounding in case-control association studies. *Am J Hum Genet* 2005;76:268–275. [PubMed: 15625622]
28. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, et al. Strong association of de novo copy number mutations with autism. *Science* 2007;316:445–449. [PubMed: 17363630]
29. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008;320:539–543. [PubMed: 18369103]
30. Lin YM, Chao SC, Chen TM, Lai TJ, Chen JS, Sun HS. Association of functional polymorphisms of the human tryptophan hydroxylase 2 gene with risk for bipolar disorder in Han Chinese. *Arch Gen Psychiatry* 2007;64:1015–1024. [PubMed: 17768266]
31. Zhou Z, Roy A, Lipsky R, Kuchipudi K, Zhu G, Taubman J, et al. Haplotype-based linkage of tryptophan hydroxylase 2 to suicide attempt, major depression, and cerebrospinal fluid 5-hydroxyindoleacetic acid in 4 populations. *Arch Gen Psychiatry* 2005;62:1109–1118. [PubMed: 16203956]
32. McMahon FJ, Buervenich S, Charney D, Lipsky R, Rush AJ, Wilson AF, et al. Variation in the gene encoding the serotonin 2a receptor is associated with outcome of antidepressant treatment. *Am J Hum Genet* 2006;78:804–814. [PubMed: 16642436]
33. Paddock S, Laje G, Charney D, Rush AJ, Wilson AF, Sorant AJM, et al. Association of GRIK4 with outcome of antidepressant treatment in the STAR*D cohort. *Am J Psychiatry* 2007;164:1181–1188. [PubMed: 17671280]

Fig. 1. Effect of minor allele frequency on tagging single nucleotide polymorphism (SNP) selection. The above histogram displays the number of SNPs ($N=145$ total) at each minor allele frequency bin (solid black bars). The striped bars indicate the number of SNPs, in each minor allele frequency bin, which were selected as tagSNPs using the method of Carlson *et al.* [18].

Fig. 2.

Accuracy of haplotype inference for three sets of tagging single nucleotide polymorphism (SNPs). Each of the eight haplotype blocks is shown on the x -axis. The minimum haplotype R_h^2 value for all common haplotypes within each block, inferred using the tagSNP set, is shown on the y -axis. Three different methods for selection of tagSNPs are shown (see Methods). Owing to the fact that methods of Carlson *et al.* [18] and Ke and Cardon [21] do not output a unique set of tagSNPs, average ($\pm$ SD) minimum R_h^2 of all unique combinations of tagSNPs are shown.

Table 1
Tagging SNP association results for the citalopram response and remission and response phenotype comparisons

SNP	Gene	Remit			Resp42		
		P value	OR (95% CI)	P value	OR (95% CI)	P value	OR (95% CI)
rs1923884	HTR2A	0.01	0.72 (0.55,0.95)	0.02	0.75 (0.58,0.97)		
rs7997012	HTR2A	0.00003	1.52 (1.20,1.95)	0.0002	1.43 (1.13,1.81)		

Results from 33 SNPs with additive *P* values < 0.05 are shown.
CI, confidence interval; OR, odds ratio; SNP, single nucleotide polymorphism.