Published in final edited form as:
J Psychiatr Res. 2013 September ; 47(9): 1157–1165. doi:10.1016/j.jpsychires.2013.05.002.

A Genome-wide Association Study of a Sustained Pattern of Antidepressant Response

Aimee M. Hunter, Ph.D.[1], **Andrew F. Leuchter, M.D.**[1], **Robert Power, MSc**[2], **Bengt Muthén, Ph.D.**[3], **Patrick J. McGrath, M.D.**[4], **Cathryn M. Lewis, Ph.D.**[2], **Ian A. Cook, M.D.**[1], **Holly A. Garriock, Ph.D.**[5,6], **Peter McGuffin, MB, Ph.D.**[2], **Rudolf Uher, Ph.D.**[2,7], and **Steven P. Hamilton, M.D., Ph.D.**[5]

[1] Department of Psychiatry and Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles

[2] Institute of Psychiatry, King's College London

[3] UCLA Graduate School of Education & Information Studies, University of California, San Francisco

[4] Department of Psychiatry, Columbia University College of Physicians and Surgeons and New York State Psychiatric Institute, University of California, San Francisco

[5] Department of Psychiatry and Institute for Human Genetics, University of California, San Francisco

[6] National Institute of Mental Health (NIMH), Division of Developmental Trajectories, Dalhousie University, Halifax, Nova Scotia

[7] Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia

Abstract

Genome-wide association studies (GWAS) have failed to replicate common genetic variants associated with antidepressant response, as defined using a single endpoint. Genetic influences may be discernible by examining individual variation between sustained versus unsustained patterns of response, which may distinguish medication effects from non-specific, or placebo responses to active medication. We conducted a GWAS among 1,116 subjects with Major Depressive Disorder from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial who were characterized using Growth Mixture Modeling as showing a sustained versus unsustained pattern of clinical response over 12 weeks of treatment with citalopram. Replication analyses examined 585 subjects from the Genome-based Therapeutic Drugs for Depression (GENDEP) trial. The strongest association with sustained as opposed to unsustained response in STAR*D involved a single nucleotide polymorphism (SNP; rs10492002) within the acyl-CoA synthetase short-chain family member 3 gene (*ACSS3*, p-value = 4.5×10^{-6}, odds ratio

Correspondence: Aimee M. Hunter, Ph.D. Assistant Professor of Psychiatry Semel Institute for Neuroscience and Human Behavior David Geffen School of Medicine at UCLA 760 Westwood Plaza, Rm. 57-455 Los Angeles, CA 90024-1759 310-206-2237 office amhunter@ucla.edu.

A portion of this paper was presented in poster format at the 11th Annual Pharmacogenetics in Psychiatry meeting March 31, 2012, New York, NY. The trajectories that form the basis for sustained and unsustained patterns of response in STAR*D were previously reported in Muthén et al., 2011.

= 0.61). No SNPs met our threshold for genome-wide significance. SNP data were available in GENDEP for 18 of the top 25 SNPs in STAR*D. The most replicable association was with SNP rs7816924 (p = 0.008, OR = 1.58); no SNP met the replication p-value threshold of 0.003. Joint analysis of these 18 SNPs resulted in the strongest signal coming from rs7816924 ($p = 2.11 \times 10^{-7}$), which resides in chondroitin sulfate N-acetylgalactosaminyltransferase 1 gene (*CSGALNACT1*). An exploratory genetic pathway analysis revealed evidence for an involvement of the KEGG pathway of long-term potentiation (FDR =.02). Results suggest novel genetic associations to sustained response.

Keywords

antidepressant; genetics; STAR*D; GENDEP; growth mixture modeling; citalopram

INTRODUCTION

Treatment with antidepressant medications is associated with significant improvements in clinical symptoms of Major Depressive Disorder (MDD), as well as improvements in functional status and quality of life. However, there is marked heterogeneity of outcomes including a subset of patients who show unsustained response (Muthén et al., 2011; Quitkin et al., 1984). Inter-individual variation in antidepressant response is under genetic influence (Tansey et al., 2012a), yet no genetic marker has shown a consistent association with clinical outcome (Tansey et al., 2012b; Uher et al., 2012). Limited progress in predicting drug efficacy may be in part due to heterogeneity in MDD related to complex gene-environment etiology (Keers & Aitchison, 2011), or the inability to separate specific response to antidepressants from naturalistic course or placebo response (Malhorta, 2010; Malhorta et al., 2012), among other factors.

The discovery of predictors of clinical response may also depend critically on the classification of outcomes. MDD trials commonly define outcomes using a predetermined cutoff score assessed at a single primary endpoint. This approach fails to account for patterns of change in clinical symptoms over time (Muthén et al., 2008) and may not reflect clinically or physiologically meaningful distinctions (Uher et al., 2010). Clinical changes over time are especially relevant when subjects exhibit alternating improvement and worsening of symptoms (Hunter et al., 2010) or a U-shaped pattern of outcome (Muthén et al., 2011; Quitkin et al., 1984). Unsustained response is clinically undesirable and may represent a “placebo” response rather than a “true drug” effect (Quitkin et al., 1984). Insofar as differences between sustained and unsustained response patterns may reflect a physiological substrate, it is of interest to examine these phenotypes for genetic association.

Advanced statistical modeling techniques have identified various response patterns, including unsustained response during acute antidepressant treatment. Growth mixture modeling (GMM) is a systematic, data-driven approach that utilizes symptom severity measures from all available time points to identify distinct trajectories of response; cluster analytic features are incorporated into GMM to reveal latent “classes” or patterns of change in symptom severity over time (Muthén & Asparouhouv, 2009; Muthén & Shedden, 1999). Such techniques have been successfully applied to longitudinal data to identify response patterns of clinical relevance during pharmacotherapy interventions in MDD (Gueorguieva et al., 2011; Hunter et al., 2010; Muthén et al., 2011; Power et al., 2012; Uher et al., 2009; Uher et al., 2010; Uher et al., 2011).

GMM was recently applied to data from the Sequenced Treatment Alternatives to Relieve Depression trial (STAR*D) (Trivedi et al., 2006), a large open-label multi-site study that,

because of its size and inclusion of “real world” patients, is especially well suited to this technique. Analyses that examined all available scores on the 16-item clinician-rated Quick Inventory of Depressive Symptoms (QIDS-C) (Rush et al., 2003b) obtained at baseline and over 12 weeks during Level 1 treatment with citalopram yielded fundamental trajectory shapes providing evidence of four classes: ‘non-responders’; ‘partial improvers’; ‘sustained responders’ (SUS) showing monotonic improvement culminating in response at week 12; and ‘unsustained responders’ (UNS), showing U-shaped response-level improvement by week 6 but with a return of baseline-level symptoms by week 12 (Figure 1) (Muthén et al., 2011). SUS and UNS responder class sizes ranged from 32% to 45%, and 6% to19%, respectively, depending on the model (Muthén et al., 2011).

We hypothesize that sustained and unsustained response trajectories represent biologically distinct types of response to antidepressants. To test this hypothesis, we conducted a genomewide association study (GWAS) contrasting STAR*D subjects in SUS versus UNS response trajectory classes to determine whether common DNA variation determines durability of response to antidepressant treatment. Identification of individuals unlikely to sustain antidepressant response would have great clinical utility, providing incentive for aggressive optimization of treatment in susceptible individuals.

METHODS and MATERIALS

Overview

GWAS was conducted in the STAR*D dataset to test for association between single-locus SNP variants and durability of response (‘sustained’ versus ‘unsustained’ response class outcomes defined using GMM). SNPs with the strongest association were then examined prospectively for replication in subjects from the Genome-based Therapeutic Drugs for Depression (GENDEP) study. Secondary, gene-based analyses were conducted to determine the association between combined effects of SNPs within individual genes and response durability. A third, exploratory level of analysis examined: 1) the combined effects of SNPs in functionally related genes i.e., ‘gene set enrichment analysis,’ and 2) aggregate effects of SNPs (across genes) found in our STAR*D GWAS to have the strongest statistical association with the sustained-unsustained response phenotype, i.e., ‘SNP profile scoring’ analysis. This scoring algorithm was then tested in the GENDEP sample.

Subjects - STAR*D-based Analysis

STAR*D enrolled treatment-seeking adults from primary care and psychiatric outpatient settings across the United States, meeting DSM-IV criteria for non-psychotic MDD and having a score ≥ 14 on the 17-item Hamilton Depression Rating Scale (HAM-D). Included were subjects having psychiatric and other medical comorbidities other than those which were either contraindicated by the protocol medications (e.g. bulimia nervosa), or would specify alternative treatment (e.g. primary obsessive compulsive disorder). Enrollees had a mean entry score >21 on the HAM-D indicating moderate to severe depression (Trivedi et al., 2006). In Level 1, all subjects received flexible, manualized, measurement-based treatment with citalopram (60 mg./day maximum final dose) for up to 14 weeks based upon clinical response and side effects evaluated at weeks 2, 4, 6, 9, and 12. Patient care and evaluation were coordinated through investigators at 14 Regional Centers who provided protocol implementation oversight. Details of the STAR*D trial design and conduct (Fava et al., 2003; Rush et al., 2003a; Trivedi et al., 2006) have been described elsewhere. DNA samples were collected according to the STAR*D protocol as described previously (Kraft et al., 2007).

Response trajectory classes in STAR*D were generated using GMM analysis (Mplus version 5.21) applied to all available scores on the QIDS-C obtained at baseline and at weeks 2,4,6, 9, and 12 during Level 1 treatment with citalopram (Muthén et al., 2011). The NMAR models of Muthén et al. (2011) give different results from modeling under MAR. Whereas the NMAR models cannot be compared statistically, three of the four NMAR models give similar results, thereby supporting each other (Supplemental Table 1). The Muthén-Roy model is the preferred model due to its performance when using an auxiliary distal outcome (Muthén et al., 2011); however, the Roy model (Figure 1) is chosen here because it is easier to work with and gives results similar to those of the Muthén-Roy model. Of 1,491 genotyped subjects in STAR*D, the 4-class Roy model identified 869 (43%) with SUS response, and 247 (18%) with UNS response (Supplemental Table 2). Mplus scripts for this model are available at: http://www.statmodel.com/examples/stard/run5.out. Regarding self-determined ethnic/racial identity, we analyzed 774 Non-Hispanic Caucasians (69.4%), 122 Hispanic Caucasians (10.9%), 145 Non-Hispanic African-Americans (13.0%), 4 Hispanic African-Americans (0.4%), 16 Asians (1.4%), and 55 “other” (4.9%). The “other” category included subjects reporting “multi-racial” (n=39), as well as a small number reporting Native American, Pacific Islander, or unspecified race and ethnicity.

Genotyping - STAR*D-based Analysis

STAR*D samples were genotyped and quality control procedures applied as described more fully elsewhere (Garriock et al., 2010). Two platforms were used for genotyping, the Affymetrix Human Mapping 500K Array Set (Affymetrix, South San Francisco, CA) and the Affymetrix Genome-Wide Human SNP Array (5.0). Samples run on the 500K Array were called using the Bayesian Robust Linear Model with Mahalanobis distance classifier (BRLMM) algorithm, while those analyzed on the 5.0 Array were called using the BRLMM-P algorithm. SNPs with minor allele frequencies less than 0.01 (n = 10,792) or with a call rate less than 95% (n = 42,908) were removed. After applying quality controls, 430,198 SNPs were subject to analysis. Individuals were excluded from analysis if they had more than 5% of their genotypes missing or were cryptically related to others in the sample, as determined using the IBS metric tabulated in PLINK v1.04 (Purcell et al., 2007).

Population Stratification in STAR*D

As in previous work with this dataset (Garriock et al., 2010), variation in subject ancestry was addressed using multidimensional scaling (MDS), a multivariate method to form a linear combination summary of rare SNP alleles. A total of 205,598 independent SNPs were used in the MDS analysis, with independence being determined with the “— indep” function in PLINK as previously described (Garriock et al., 2010). Each of 10 MDS values for each individual was tested for association to the tested phenotype; and only one was found to be associated, and thus retained as a covariate for analysis. The first two MDS vectors correlated with continental ancestry, as shown in Supplemental Figure 1.

Statistical Analyses- STAR*D Subjects

Genetic Association—Statistical analyses were conducted using PLINK v1.07.SNPs were tested for association with clinical outcome (SUS vs. UNS) using logistic regression. Covariates were medication tolerability (a STAR*D phenotype), gender, and an MDS-based measure of ancestry (Supplemental Table 2). The minor allele homozygous genotype served as the reference group, and each SNP was modeled in individuals as having a log-additive effect, after adjusting for the 3 covariates mentioned above. Genome-wide significance was set at 1.16×10^{-7}, representing a Bonferroni correction for 430,198 SNPs. Statistical power was estimated using Quanto (Gauderman, 2002) with parameters set at: alpha = 1.16×10^{-7}, risk of belonging to the UNS response (0.23) group, $1\text{-}\beta = 0.8$, log additive model, and

minor allele frequency of 0.25. For the SUS response phenotype, the genotypic relative risk that could be detected at 80% power was 1.97.

Subjects - GENDEP-based analysis

GENDEP incorporated treatment-seeking adults of white European ancestry with an ICD-10/DSM-IV diagnosis of major depressive disorder and currently in a moderate-to-severe depressive episode, treated across 9 European centres (in Belgium, Croatia, Denmark, Germany, Italy, Poland, Slovenia and the UK). Diagnosis was established using the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) interview and the study excluded those with a personal or family history of bipolar affective disorder, mood-incongruent psychotic symptoms or active substance dependence. Depression severity was assessed at recruitment and in weekly intervals over the 12 weeks of treatment with the 10-item Montgomery-Åsberg Depression Rating Scale (MADRS), rated by trained psychiatrists and psychologists with excellent inter-rater reliability (Uher et al., 2008). GENDEP was part-randomised with patients with no contraindications allocated randomly to receive either escitalopram or nortriptyline. If an individual had a known contraindication or a history of side effects to one drug, they were non-randomly allocated to the other. This resulted in 473 randomly allocated and 325 non-randomly allocated subjects (overall 56% on escitalopram). Full details of this study have been described elsewhere (Uher et al., 2009; Uher et al., 2010). The GENDEP project was approved by ethics boards of participating centers, and all participants provided written informed consent. The present replication sample consisted of 798 individuals (502 females) with post baseline data allowing for trajectory analysis.

Modeling of SUS vs. UNS response classes was performed within Mplus (version 6) using a procedure identical to that used in the STAR*D sample, with the Bayesian Information Criterion (BIC) used to establish the best fitting model (Muthén et al., 2011; Power et al., 2012). The results of trajectory modeling in GENDEP are summarized in Supplemental Table 3, Supplemental Figure 3, and are further described in a separate publication (Power et al., 2012). The 5 and 6 class Muthen-Roy models were found to have the best fits (difference in BIC score<2). The 6-class model was chosen for the replication analysis as it provided more distinct response patterns and was comparable to the STAR*D trajectories; this model identified 394 (49%) GENDEP subjects with SUS response, and 191 (24%) with UNS response.

Genotyping in GENDEP

Blood-derived DNA was genotyped at the Centre National de Genotypage (Evry Cedex, France) on the Illumina Human610-quad bead chip (Illumina, Inc., San Diego), as previously described (Uher et al., 2010).

Data analysis in GENDEP

Quality control procedures were carried out as previously reported (Uher et al., 2009). A minor allele frequency >0.01, individual genotype completeness above 95%, and SNP genotype completeness above 99% was required for inclusion. Association was corrected for the first five principal components to ensure that results were not confounded by population stratification. Analysis was restricted to the top 25 SNPs associated with transient responders in the STAR*D dataset. Using a reference sample of Europeans (CEU) from phase 3 of the HapMap Project (Altshuler et al., 2010), missing genotyped SNPs were imputed using BEAGLE 3.3. Analysis of association was carried out in PLINK (Purcell et al., 2007) comparing SUS to UNS responders. This used an additive genetic model, where the effect of a risk allele is presumed to increase in proportion to the number of risk alleles the individual has for that SNP. The first five ancestry-informative principal components

were used as covariates to address population stratification, as in the original analysis of this data set.

Replication and Joint analysis

In order to test for replication of the STAR*D findings, we identified 18 SNPs in the GENDEP dataset that were genotyped or imputed that corresponded with our top 25 STAR*D results. We considered a positive replication a finding in which the direction of effect between studies was identical, and where the GENDEP p-value met a p-value of 0.003, correcting for 18 tests. We carried a sign test to test the hypothesis that the direction of associations between STAR*D and GENDEP were concordant. We also sought to combine these analyses, and we examined these same 18 SNPs from the STAR*D GWAS and GENDEP replication and analyzed them in MetaP, a program that performs a weighted z meta-analysis by combining p-values from independent studies, while taking account of sample sizes and effect directions (Software: MetaP; Author: DongliangGe; URL: http://www.svaproject.org/metap.php). We utilized the Stouffer's z trend test, as it considers the p-values, sample sizes, and directions of effect for the analysis.

Gene-based analysis

We carried out a secondary analysis in STAR*D focusing on the gene as the basis of association. We implemented an approach where association between our response phenotype and all SNPs within each gene was examined. The program VEGAS (Liu et al., 2010) generates a test statistic that incorporates the effect of every genotyped SNP in a gene after adjusting for linkage disequilibrium (LD). SNPs are binned into as many as 17,787 gene sets when they are in genes or within 50kb of genes. LD information from HapMap samples is used for simulation of LD in the samples under study. Up to 1 million simulations are carried out adaptively from a multivariate normal distribution. A p-value threshold of 2.86×10^{-6} was used to determine significance.

Exploratory Methods

Gene enrichment analysis—To complement our analyses, we utilized a web-based program, i-GSEA4GWAS, accessible online at http://gsea4gwas.psych.ac.cn/ (Zhang et al., 2010), for identification of pathways associated with our phenotype in a gene set enrichment analysis that employs SNP label permutation and Kolmogorov-Smirnov-like statistical analysis. This in an empirical data-driven approach wherein each gene +/- the surrounding 100 kb is represented by its most significant SNP p-value and then ranked against all other genes from most to least significant. A ‘significance proportion-based enrichment score’ is then calculated based on the cumulative significance of groups of genes belonging to a common pathway. This algorithm considers proportions (rather than raw numbers) of genes in gene sets crossing a threshold for statistical significance to normalize for gene sets with disparate numbers of member genes. The algorithm generates an associated p-value and false discovery rate (FDR), with FDR values ≤0.05 denoting high confidence. We interrogated the canonical pathways, which are derived from the Molecular Signatures Database (MSigDB, v2.5) (Subramanian et al., 2005).

SNP profile scoring—In order to explore the hypothesis that our chosen phenotypes are influenced in aggregate by multiple common variants with weak effects, we used a SNP scoring routine in PLINK (‘--score’ command) to generate allelic scoring profiles for our response phenotype based on top results in the STAR*D analysis (i.e., SNPs with $p \leq 0.001$, 0.01, 0.1, 0.2, 0.3, 0.4, and 0.5). These profiles referenced a given allele and a $\log_{10}$ of the odds ratio (OR) for each of the top SNPs. Each individual for the corresponding analysis in GENDEP was then scored using these profiles. Scores were computed as sums across SNPs

of numbers of tested alleles present (0, 1, or 2) multiplied by the $\log_{10}$ OR for that SNP in the corresponding STAR*D dataset to weigh strength of the association. Only uncorrelated SNPs (intermarker $r^2 < 0.25$) were included. Logistic regression analyses were run in STATA v.9.2 (Statacorp, College Station, TX) using GENDEP response phenotype as the dependent variable and SNP score as the independent variable. The pseudo r-square from the logistic regression was used to estimate the proportion of additional variance explained when the polygenic scores were included into the model, compared to a model consisting only of covariates (the first five ancestry-informative principal components).

RESULTS

STAR*D Clinical and Demographic Features

A total of 1,116 subjects from STAR*D were analyzed (Supplemental Table 2). There were 869 subjects who were classified as SUS, while there were 247 subjects who were classified as UNS using our GMM algorithm. These latter subjects are described by a pattern of initial response to treatment with return of symptoms over time. Clinical and demographic variables were tested for association with the sustained response phenotype. Those that were associated with the sustained response pattern at a nominal level ($p<0.05$) were male gender, drug intolerance, and the 8th MDS vector. Interestingly, age, baseline severity, and race/ethnicity (as measured by the first several MDS vectors), were not associated with the response phenotype (Supplemental Table 2).

Genetic Association Results - STAR*D

Table 1 shows p-values, odds ratios (ORs) and 95% confidence intervals (CIs) for the top 25 SNPs in the GWAS analysis. Among these SNPs, 14 occurred in the intronic regions of 12 genes. The strongest finding involved intronic SNPs in the gene encoding acyl-CoA synthetase short-chain family member 3 (*ACSS3*), a mitochondrial enzyme predicted to generate acetyl-CoA for energy generation (Figure 2). The odds ratios for the associated SNPs in the area were 0.61-0.62, indicating that the minor allele increased risk of UNS pattern. No SNP met our threshold for genome-wide significance. Notable genes among the most associated regions include *SEMA5A* (rs448038, $p = 2.23 \times 10^{-5}$, odds ratio = 2.82), encoding semaphorin 5a, which was also found to be associated with autism in a GWAS (Weiss et al., 2009). The axonal guidance properties of semaphorin 5a are regulated in part by chondroitin sulfate proteoglycans (Kantor et al., 2004), whose synthesis are initiated by the enzyme encoded by *CSGALNACT1*, a gene also showing association in our sample (rs7816924, $p = 7.72 \times 10^{-6}$, odds ratio = 2.14). Other genes of interest include the thyroid stimulating hormone receptor (*TSHR*), thyrotropin-releasing hormone degrading enzyme (*TRHDE*), fibroblast growth factor 14 (*FGF14*), the SORCS 2 receptor (*SORCS2*), and the amyloid beta A4 protein isoform a precursor (*APP*).

GENDEP Clinical and Demographic Features

Of the 798 individuals with post-baseline data, 192 were classified as UNS while 394 were classified as SUS (see Supplementary Figure 2). Demographic and clinical variables were tested for association with response class. SUS and UNS subjects did not differ on age or gender. UNS was more common during treatment with escitalopram than with nortriptyline and in subjects with more severe depression at baseline (both $p<0.01$), as has been described previously (Power et al., 2012).

Genetic Association Replication in GENDEP

Of the top 25 SNPs in the STAR*D analysis, 18 had genotyped or imputed data that passed quality control requirements in GENDEP. Five SNPs could not be assessed due to

frequencies of 1 or less (likely due to greater historical mixed ancestry in the US compared to Europe), and two could not be imputed at the threshold level we used for analysis. Two SNPs in the replication sample were associated with unsustained vs. sustained response at nominal significance (see Table 2). The strongest association was with rs7816924 on chromosome 8; this association was in the same direction as that in STAR*D with an OR of 1.58 (p=0.008), just below the level needed when correction for multiple testing was applied (p=0.003).

Pooled Results

We carried out a combined analysis of the 18 SNPs in STAR*D and GENDEP using p-values from independent studies that account for the impact of sample sizes and directions of effect. None of the variants met genome-wide significance. In this joint analysis, the strongest genetic association with sustained versus unsustained response was with the minor allele of SNP rs7816924 in the *CSGALNACT1* gene ($p = 2.11 \times 10^{-7}$) (Table 2). We tested whether the findings for the two studies were concordant in direction of effect more often than expected by chance. A sign test suggested a significant difference (15 concordant and 3 non-concordant pairs, $p = 0.008$).

Gene-based analysis

We sought to determine if there were genes showing excess association signals even when taking into account the level of inter-SNP LD and gene size. In our gene-based test using VEGAS software, seven genes in three regions showed genome-wide significance (Supplemental Table 5; Supplemental Figure 3). Because the genes in each of the regions were in close proximity to another, it is possible that some SNPs (which were counted if they were within 50kb of a gene) were being used for more than one gene to assess the gene-based statistic.

Exploratory Results

Genetic pathway analysis—In this and the following section, we carried out additional analyses to expand the scope of our GWAS and gene-based analysis. Because typical GWAS and gene-based analyses emphasize the most statistically significant individual variants, or genes, respectively, there is often less focus in complex traits on the combined effects of SNPs in functionally related genes. Recognizing this limitation, we sought to identify sets of genes and functional pathways enriched for stronger association signals than one would expect by chance. We utilized the web-based utility i-GSEA4GWAS for a gene set enrichment analysis and filtered the results for items with a false-discovery rate (FDR) ≤ 0.05. The canonical pathways which performed most strongly in the GWAS of unsustained responders versus sustained responders included: Alzheimer's disease, type I diabetes mellitus, the tumor necrosis factor pathway, antigen processing and presentation, long term potentiation, the mPR pathway, WNT signaling, and the GAQ pathway. Genes near SNPs with nominal association in the GWAS are shown in Supplemental Table 4 For example, we found that SNPs were nominally associated with our phenotype in 32 of 54 of the genes with testable variants in the KEGG long term potentiation pathway, including CACNA1C, GRM1, GRM5, GRIA1, GRIN2A, GRIN2B, GRIN2C, PRKACG, and *PRKCA* (FDR = 0.02).

Polygenic profile analysis—The risk score models in STAR*D included: 84,186 SNPs with p≤ 0.5, 70,373 SNPs with p≤ 0.4, 55,143 SNPs with p≤ 0.3,38,521 SNPs with p≤ 0.2, 20,213 SNPs with p ≤ 0.1, 2,091 SNPs with p ≤ 0.01, and 182 SNPs with p ≤ 0.001. Risk score models developed in STAR*D did not significantly predict the sustained response phenotype in GENDEP (Table 3).

DISCUSSION

We carried out a GWAS to address the hypothesis that DNA variation influences a pattern of unsustained as compared to sustained antidepressant response in individuals with unipolar major depression in the STAR*D sample. Our strongest finding involved *ACSS3*, a gene not previously linked with antidepressant biology. The risk allele increased the likelihood of the unsustained response. There is no overlap between the top findings from this analysis and our previous GWAS of citalopram response in the STAR*D sample (Garriock et al., 2010), suggesting that our approach is not simply differentiating responders from non-responders as identified using a primary clinical endpoint. Instead, by comparing response trajectory phenotypes, we may be identifying genetic determinants of a subset of drug response that is unsustained over the first several months of treatment. There were no strong clinical or demographic predictors separating sustained responders from unsustained responders.

The results summarized here may identify new genes or pathways related to clinically useful patterns of response, although the lack of genome-wide significance renders any speculation premature. Our attempt to replicate the very top findings from our GWAS in an independent sample yielded no significant finding given the number of SNPs tested in the GENDEP dataset. A joint analysis of the two samples using 18 of the top STAR*D SNPs that could be genotyped or imputed in GENDEP provided modest additional support for the *ACSS3* SNPs, while showing increasing support for a SNP in *CSGALNACT1*. Regarding *CSGALNACT1*, it may be of interest to note that in a previous GWAS in GENDEP, the strongest reported association with antidepressant response (to nortriptyline) was with a SNP in another chondroitin sulfate related gene, uronyl 2-sulfotransferase (UST) (Uher et al., 2010). While chondroitin sulfate has a well established role in the skeletal system (cf Vangsness 2009 PMID 19111223), there is evidence that it may also play a regulatory role in neuroplasticity, regeneration, and brain development processes (Galtrey and Fawcett, 2007; Orlando et al., 2012). In a murine knock-out model, the absence of *CSGALNACT1* has been reported to alter cortical thickness, potentially by altering cell migration patterns (Onaga et al., 2010). Similarly, variants in the KEGG long-term potentiation pathway have been linked to neurodegenerative disorders where depressive symptoms are common (Ramanan et al., 2012; Botta-Orfila et al., 2012) and preliminarily to primary affective disorders (Kao et al., 2012).

Limitations of using the STAR*D sample for genetic studies have been extensively discussed (Garriock et al., 2010; Kraft et al., 2007; Laje et al., 2009), with the most relevant involving the lack of a placebo arm, inadequate measures of drug adherence, and stratification due to the multiethnic nature of the sample. We have addressed the last issue by controlling for population structure by incorporating covariates in the analysis derived from estimates of genetic ancestry. Interestingly, major ancestry vectors were not associated with the response phenotype, unlike in our previous studies of general response or remission on citalopram (Garriock et al., 2010). Our genomic inflation factor (1.0, mean chi-squared = 0.96) suggests no systematic inflation of statistics that would be expected by hidden population structure. Post-hoc analysis of the largest sub-group, non-Hispanic Whites (n=774), did not provide stronger results.

Our sample was only large enough to detect genotypic relative risks in the range of 2.0 or greater, and we observed findings with odds ratios in the range of 0.20-0.64 for protective alleles and 1.62-2.82 for risk alleles. Given the limited power, further meta- or mega-analysis (Ripke et al., 2012) of all available samples with similar pharmacogenetic phenotypes will be required before any definitive conclusions can be generated. Similarly, while our data could be construed as not supporting the role of common genetic variation in antidepressant response durability, this conclusion may be premature. We found that the

direction of effect was concordant between STAR*D and GENDEP significantly more often than expected, suggesting that our lack of genome-wide significance may be related to the limited sample sizes. It is possible that larger samples may detect true risk alleles of small effect, but still of scientific interest. It is also possible that higher density genotyping of common alleles not covered by our genotyping platform would uncover missed associations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

A portion of this paper was presented in poster format at the 11th Annual Pharmacogenetics in Psychiatry meeting March 31, 2012, New York, NY. The trajectories that form the basis for sustained and unsustained patterns of response in STAR*D were previously reported in Muthén et al., 2011.

Support: Clinical data used in the GWAS were obtained from the limited access datasets distributed from the NIMH-supported “Sequenced Treatment Alternatives to Relieve Depression” (STAR*D) (Contract # N01MH90003 to the University of Texas Southwestern Medical Center). Genotyping of the STAR*D sample was supported by NIMH grant MH072802 to S.P. Hamilton. Analysis of STAR*D clinical data and response trajectories was supported by NIMH grant 1R34MH085933-01 to A.M. Hunter. This manuscript reflects the views of the authors and may not reflect the opinions or views of the STAR*D Study Investigators or the NIH.

Data for replication analyses were obtained from the GENDEP project which was funded by the European Commission Framework 6 grant, EC Contract Ref.: LSHB-CT-2003-503428. H. Lundbeck provided nortriptyline and escitalopram for the GENDEP study. GlaxoSmithKline and the UK National Institute for Health Research of the Department of Health contributed to the funding of the sample collection at the Institute of Psychiatry, London. GENDEP genotyping was funded by a joint grant from the U.K. Medical research council and GlaxoSmithKline (G0701420) and the European Commission Innovative Medicine Initiative Joint Undertaking (IMI-JU) grant n^{O} 115008. R. Uher is supported by the Canada Research Chairs program (http://www.chairs-chaires.gc.ca/) and the European Commission Innovative Medicine Initiative Joint Undertaking (IMI-JU) grant n^{O} 115008.

REFERENCES

Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Peltonen L, Dermitzakis E, Bonnen PE, Altshuler DM, Gibbs RA, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, Jia X, Palotie A, Parkin M, Whittaker P, Yu F, Chang K, Hawes A, Lewis LR, Ren Y, Wheeler D, Gibbs RA, Muzny DM, Barnes C, Darvishi K, Hurles M, Korn JM, Kristiansson K, Lee C, McCarrol SA, Nemesh J, Dermitzakis E, Keinan A, Montgomery SB, Pollack S, Price AL, Soranzo N, Bonnen PE, Gibbs RA, Gonzaga-Jauregui C, Keinan A, Price AL, Yu F, Anttila V, Brodeur W, Daly MJ, Leslie S, McVean G, Moutsianas L, Nguyen H, Schaffner SF, Zhang Q, Ghori MJ, McGinnis R, McLaren W, Pollack S, Price AL, Schaffner SF, Takeuchi F, Grossman SR, Shlyakhter I, Hostetter EB, Sabeti PC, Adebamowo CA, Foster MW, Gordon DR, Licinio J, Manca MC, Marshall PA, Matsuda I, Ngare D, Wang VO, Reddy D, Rotimi CN, Royal CD, Sharp RR, Zeng C, Brooks LD, McEwen JE. Integrating common and rare genetic variation in diverse human populations. Nature. 2010; 467:52–58. [PubMed: 20811451]

Botta-Orfila T, Tolosa E, Gelpi E, Sànchez-Pla A, Martí MJ, Valldeoriola F, Fernández M, Carmona F, Ezquerra M. Microarray expression analysis in idiopathic and LRRK2-associated Parkinson's disease. Neurobiology of Disease. 2012; 45:462–468. [PubMed: 21946334]

Fava M, Rush AJ, Trivedi MH, Nierenberg AA, Thase ME, Sackeim HA, Quitkin FM, Wisniewski S, Lavori PW, Rosenbaum JF, Kupfer DJ. Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study. Psychiatric Clinics of North America. 2003; 26:457–494. [PubMed: 12778843]

Galtrey CM, Fawcett JW. The role of chondroitin sulfate proteoglycans in regeneration and plasticity in the central nervous system. Brain Research Reviews. 2007; 54:1–18. [PubMed: 17222456]

Garriock HA, Kraft JB, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD, Reinalda MS, Slager SL, McGrath PJ, Hamilton SP. A Genomewide Association Study of Citalopram Response in Major Depressive Disorder. Biological Psychiatry. 2010; 67:133–138. [PubMed: 19846067]

Gauderman WJ. Sample size requirements for association studies of gene-gene interaction. The American Journal of Epidemiology. 2002; 155:478–484.

Gueorguieva R, Mallinckrodt C, Krystal JH. Trajectories of depression severity in clinical trials of duloxetine: insights into antidepressant and placebo responses. Archives of General Psychiatry. 2011; 68(12):1227–37. [PubMed: 22147842]

Hunter AM, Muthén BO, Cook IA, Leuchter AF. Antidepressant response trajectories and quantitative electroencephalography (QEEG) biomarkers in major depressive disorder. Journal of Psychiatric Research. 2010; 44(2):90–98. [PubMed: 19631948]

Kantor DB, Chivatakarn O, Peer KL, Oster SF, Inatani M, Hansen MJ, Flanagan JG, Yamaguchi Y, Sretavan DW, Giger RJ, Kolodkin AL. Semaphorin 5A Is a Bifunctional Axon Guidance Cue Regulated by Heparan and Chondroitin Sulfate Proteoglycans. Neuron. 2004; 44:961–975. [PubMed: 15603739]

Kao CF, Jia P, Zhao Z, Kuo PH. Enriched pathways for major depressive disorder identified from a genome-wide association study. International Journal of Neuropsychopharmacology. 2012; 15:1401–1411. [PubMed: 22243633]

Keers R, Aitchison KJ. Pharmacogenetics of antidepressant response. Expert Review of Neurotherapeutics. 2011; 11(1):101–125. [PubMed: 21158559]

Kraft JB, Peters EJ, Slager SL, Jenkins GD, Reinalda MS, McGrath PJ, Hamilton SP. Analysis of Association Between the Serotonin Transporter and Antidepressant Response in a Large Clinical Sample. Biological Psychiatry. 2007; 61:734–742. [PubMed: 17123473]

Laje G, Perlis RH, Rush AJ, McMahon FJ. Pharmacogenetics Studies in STAR*D: Strengths, Limitations, and Results. Psychiatric Services. 2009; 60:1446–1457. [PubMed: 19880459]

Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, Hayward NK, Montgomery GW, Visscher PM, Martin NG, Macgregor S. A Versatile Gene-Based Test for Genome-wide Association Studies. American Journal of Human Genetics. 2010; 87:139–145. [PubMed: 20598278]

Malhotra AK. The Pharmacogenetics of Depression: Enter the GWAS. American Journal of Psychiatry. 2010; 167:493–495. [PubMed: 20439394]

Malhotra AK, Zhang JP, Lencz T. Pharmacogenetics in psychiatry: translating research into clinical practice. Molecular Psychiatry. 2012; 17(8):760–769. [PubMed: 22083729]

Muthén, B.; Asparouhov, T. Growth mixture modeling: Analysis with non-Gaussian random effects.. In: Fitzmaurice, G.; Davidian, M.; Verbeke, G.; Molenberghs, G., editors. Longitudinal Data Analysis. Chapman & Hall/CRC Press; Boca Raton, FL: 2009. p. 143-165.

Muthén B, Asparouhov T, Hunter AM, Leuchter AF. Growth modeling with nonignorable dropout: Alternative analyses of the STAR*D antidepressant trial. Psychological Methods. 2011; 16(1):17–33. [PubMed: 21381817]

Muthén, B.; Brown, H.; Leuchter, AF.; Hunter, AM. General approaches to analysis of course: Applying growth mixture modeling to randomized trials of depression medication.. In: Shrout, P.; Keyes, K.; Ornstein, K., editors. Causality and Psychopathology: Finding the Determinants of Disorders and their Cures. American Psychiatric Publishing; Washington D.C.: 2008. p. 159-178.

Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics. 1999; 55:463–469. [PubMed: 11318201]

Onaga SH, Takeuchi K, Watanabe Y, Komuta Y, Izumikawa T, Kitagawa H, Igarashi M. Neuroscience Research. 2010; 68:e250.

Orlando C, Ster J, Gerber U, Fawcett JW, Raineteau O. Perisynaptic chondroitin sulfate proteoglycans restrict structural plasticity in an integrin-dependent manner. The Journal of Neuroscience. 2012; 32:18009–17. (2012). [PubMed: 23238717]

Power RA, Muthén B, Henigsberg N, Placentino A, Mendlewicz J, Maier W, McGuffin P, Lewis CM, Uher R. Non-random dropout and the relative efficacy of escitalopram and nortriptyline in treating major depressive disorder. Journal of Psychiatric Research. 2012; 46(10):1333–1338. [PubMed: 22770672]

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: A tool set for whole-genome association and population based linkage analyses. American Journal of Human Genetics. 2007; 81:559–575. [PubMed: 17701901]

Quitkin FM, Rabkin JG, Ross D, Stewart JW. Identification of true drug response to antidepressants. Use of pattern analysis. Archives of General Psychiatry. 1984; 41:782–786. [PubMed: 6378117]

Ramanan VK, Kim S, Holohan K, Shen L, Nho K, Risacher SL, Foroud TM, Mukherjee S, Crane PK, Aisen PS, Petersen RC, Weiner MW, Saykin AJ, Alzheimer's Disease Neuroimaging Initiative (ADNI). Genome-wide pathway analysis of memory impairment in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort implicates gene candidates, canonical pathways, and networks. Brain Imaging and Behavior. 2012; 6:634–648. [PubMed: 22865056]

Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G, Byrne EM, Blackwood DH, Boomsma DI, Cichon S, Heath AC, Holsboer F, Lucae S, Madden PA, Martin NG, McGuffin P, Muglia P, Noethen MM, Penninx BP, Pergadia ML, Potash JB, Rietschel M, Lin D, Müller-Myhsok B, Shi J, Steinberg S, Grabe HJ, Lichtenstein P, Magnusson P, Perlis RH, Preisig M, Smoller JW, Stefansson K, Uher R, Kutalik Z, Tansey KE, Teumer A, Viktorin A, Barnes MR, Bettecken T, Binder EB, Breuer R, Castro VM, Churchill SE, Coryell WH, Craddock N, Craig IW, Czamara D, De Geus EJ, Degenhardt F, Farmer AE, Fava M, Frank J, Gainer VS, Gallagher PJ, Gordon SD, Goryachev S, Gross M, Guipponi M, Henders AK, Herms S, Hickie IB, Hoefels S, Hoogendijk W, Hottenga JJ, Iosifescu DV, Ising M, Jones I, Jones L, Jung-Ying T, Knowles JA, Kohane IS, Kohli MA, Korszun A, Landen M, Lawson WB, Lewis G, Macintyre D, Maier W, Mattheisen M, McGrath PJ, McIntosh A, McLean A, Middeldorp CM, Middleton L, Montgomery GM, Murphy SN, Nauck M, Nolen WA, Nyholt DR, O'Donovan M, Oskarsson H, Pedersen N, Scheftner WA, Schulz A, Schulze TG, Shyn SI, Sigurdsson E, Slager SL, Smit JH, Stefansson H, Steffens M, Thorgeirsson T, Tozzi F, Treutlein J, Uhr M, van den Oord EJ, Van Grootheest G, Völzke H, Weilburg JB, Willemsen G, Zitman FG, Neale B, Daly M, Levinson DF, Sullivan PF. A mega-analysis of genome-wide association studies for major depressive disorder. Molecular Psychiatry. 2012 doi: 10.1038/mp.2012.21.

Rush AJ, Trivedi M, Fava M. Depression. IV: STAR*D treatment trial for depression. American Journal of Psychiatry. 2003; 160:237. [PubMed: 12562566]

Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, Markowitz JC, Ninan PT, Kornstein S, Manber R, Thase ME, Kocsis JH, Keller MB. 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. Biological Psychiatry. 2003; 54:573–583. [PubMed: 12946886]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:15545–15550. [PubMed: 16199517]

Tansey KE, Guipponi M, Hu X, Domenici E, Lewis G, Malafosse A, Wendland JR, Lewis CM, McGuffin P, Uher R. Heritability of antidepressant response estimated from genome-wide data. Genome Medicine. 2012; 4(6):52. [PubMed: 22738351]

Tansey KE, Guipponi M, Perroud N, Bondolfi G, Domenici E, Evans D, Hall SK, Hauser J, Henigsberg N, Hu X, Jerman B, Maier W, Mors O, O'Donovan M, Peters TJ, Placentino A, Rietschel M, Souery D, Aitchison KJ, Craig I, Farmer A, Wendland JR, Malafosse A, Holmans P, Lewis G, Lewis CM, Stensbol TB, Kapur S, McGuffin P, Uher R. Genetic Predictors of Response to Serotonergic and Noradrenergic Antidepressants in Major Depressive Disorder: A Genome-Wide Analysis of Individual-Level Data and a Meta-Analysis. PLOS Medicine. 2012; 9(10):e1001326. doi:10.1371/journal.pmed.1001326. [PubMed: 23091423]

Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, Norquist G, Howland RH, Lebowitz B, McGrath PJ, Shores-Wilson K, Biggs MM, Balasubramani GK, Fava M, STAR*D Study Team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. American Journal of Psychiatry. 2006; 163:28–40. [PubMed: 16390886]

Uher R, Farmer A, Maier W, Rietschel M, Hauser J, Marusic A, Mors O, Elkin A, Williamson RJ, Schmael C, Henigsberg N, Perez J, Mendlewicz J, Janzing JG, Zobel A, Skibinska M, Kozel D, Stamp AS, Bajs M, Placentino A, Barreto M, McGuffin P, Aitchison KJ. Measuring depression:

comparison and integration of three scales in the GENDEP study. Psychological Medicine. 2008; 38(2):289–300. [PubMed: 17922940]

Uher R, Maier W, Hauser J, Marusic A, Schmael C, Mors O, Henigsberg N, Souery D, Placentino A, Rietschel M, Zobel A, Dmitrzak-Weglarz M, Petrovic A, Jorgensen L, Kalember P, Giovannini C, Barreto M, Elkin A, Landau S, Farmer A, Aitchison KJ, McGuffin P. Differential efficacy of escitalopram and nortriptyline on dimensional measures of depression. British Journal of Psychiatry. 2009; 194(3):252–259. [PubMed: 19252156]

Uher R, Mors O, Rietschel M, Rajewska-Rager A, Petrovic A, Zobel A, Henigsberg N, Mendlewicz J, Aitchison KJ, Farmer A, McGuffin P. Early and delayed onset of response to antidepressants in individual trajectories of change during treatment of major depression: a secondary analysis of data from the Genome-Based Therapeutic Drugs for Depression (GENDEP) study. Journal of Clinical Psychiatry. 2011; 72(11):1478–1484. [PubMed: 22127194]

Uher R, Muthén B, Souery D, Mors O, Jaracz J, Placentino A, Petrovic A, Zobel A, Henigsberg N, Rietschel M, Aitchison KJ, Farmer A, McGuffin P. Trajectories of change in depression severity during treatment with antidepressants. Psychological Medicine. 2010; 40(8):1367–1377. [PubMed: 19863842]

Uher R, Perroud N, Ng MY, Hauser J, Henigsberg N, Maier W, Mors O, Placentino A, Rietschel M, Souery D, Zagar T, Czerski PM, Jerman B, Larsen ER, Schulze TG, Zobel A, Cohen-Woods S, Pirlo K, Butler AW, Muglia P, Barnes MR, Lathrop M, Farmer A, Breen G, Aitchison KJ, Craig I, Lewis CM, McGuffin P. Genome-Wide Pharmacogenetics of Antidepressant Response in the GENDEP Project. American Journal of Psychiatry. 2010; 167:555–564. [PubMed: 20360315]

Uher R, Tansey KE, Rietschel M, Henigsberg N, Maier W, Mors O, Hauser J, Žagar T, Placentino A, Souery D, Farmer A, Aitchison KJ, Craig I, McGuffin P, Lewis CM, Ising M, Lucae S, Binder EB, Kloiber S, Holsboer F, Müller-Myhsok B, Ripke S, Hamilton SP, Laje G, McMahon FJ, Fava M, Rush AJ, Perlis RH. Common genetic variation and antidepressant efficacy in major depressive disorder: a meta-analysis of three genome-wide pharmacogenetic studies. American Journal of Psychiatry. 2012 in press.

Weiss LA, Arking DE, Gene Discovery Project of Johns Hopkins & the Autism Consortium. Daly MJ, Chakravarti A. A genome-wide linkage and association scan reveals novel loci for autism. Nature. 2009; 461(7265):802–808. [PubMed: 19812673]

Zhang K, Cui S, Chang S, Zhang L, Wang J. i-GSEA4GWAS: a web server for identification of pathways/gene sets associated with traits by applying an improved gene set enrichment analysis to genome-wide association study. Nucleic Acids Research. 2010; 38:W90–W95. [PubMed: 20435672]

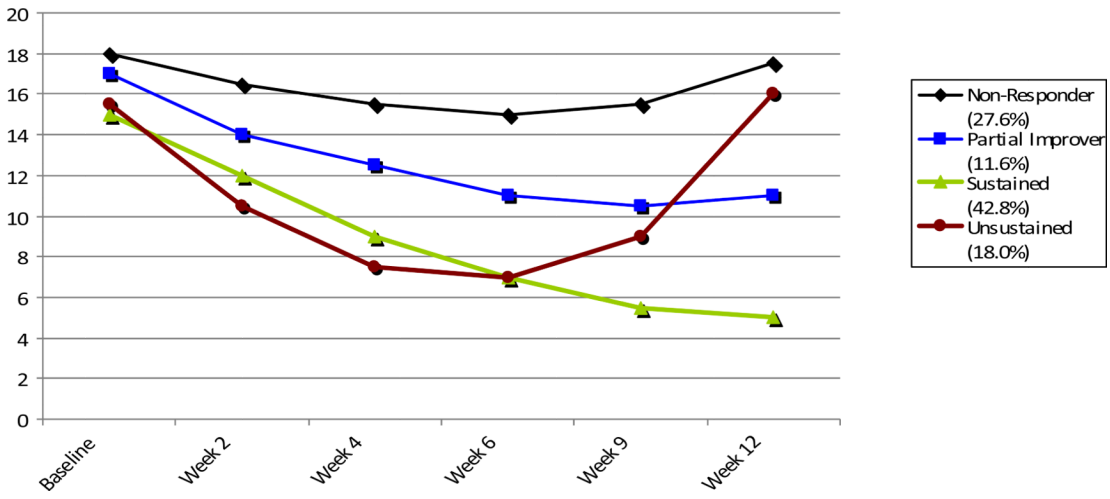

Figure 1.
Estimated mean QIDS-C scores (y-axis) across 12 weeks of citalopram treatment (x-axis) for four classes of subjects in STAR*D[2].

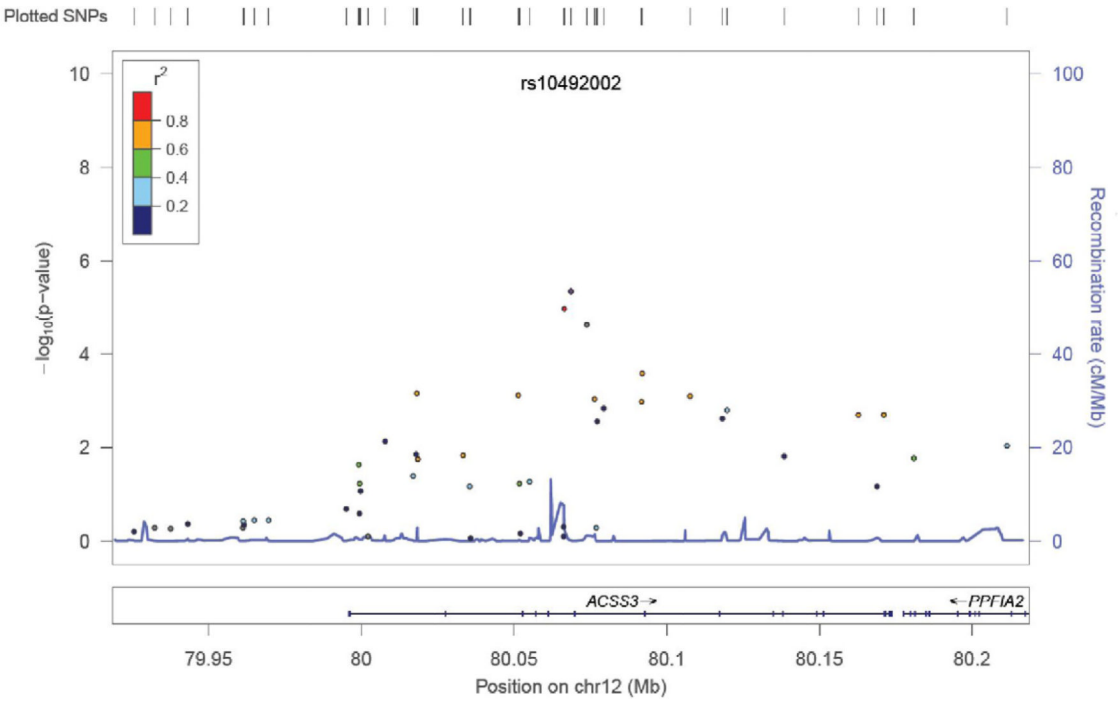

Figure 2.
SNPs in the in the region of the acyl-CoA synthetase short-chain family member 3 (*ACSS3*).

Table 1

p-values, odds ratios (ORs) and 95% confidence intervals (CIs) for the top 25 SNPs in the GWAS analysis.

SNP	CHR	BP	Allele	P	OR (95% CI)	Location	Nearby Gene	Distance (bp)
rs10492002	12	80068667	T	4.52E-06	0.61(0.50,0.75)	Intronic	ACSS3	-
rs7816924	8	19583918	A	7.72E-06	2.14(1.53,2.99)	Intronic	CSGALNACT1	-
rs1435964	12	80066346	C	1.07E-05	0.62(0.50,0.77)	Intronic	ACSS3	-
rs448038	5	9466495	C	2.23E-05	2.82(1.75,4.57)	Intronic	SEMA5A	-
rs7132119	12	80073806	A	2.39E-05	0.62(0.50,0.77)	Intronic	ACSS3	-
rs9497111	6	145246442	T	3.11E-05	0.26(0.14,0.49)	Intergenic	UTRN	30579
rs7158881	14	80661507	G	3.19E-05	0.64(0.52,0.79)	Intronic	TSHR	-
rs11931209	4	159551486	A	3.73E-05	0.60(0.47,0.76)	Intergenic	RXFP1	111011
rs6446618	4	7770149	A	3.74E-05	0.21(0.10,0.44)	Intronic	SORCS2	-
rs4439856	18	67021993	C	4.04E-05	1.71(1.32,2.21)	Intergenic	none	-
rs12080794	1	9073912	A	4.54E-05	0.57(0.44,0.75)	Intergenic	GPR157	13151
rs37747	7	110707416	G	4.86E-05	0.49(0.35,0.69)	Intronic	IMMP2L	-
rs11100172	4	159567770	G	5.27E-05	0.60(0.50,0.77)	Intergenic	RXFP1	94727
rs2476230	13	101380444	C	5.32E-05	1.69(1.31,2.17)	Intronic	FGF14	-
rs2346793	4	159585134	T	5.41E-05	0.60(0.47,0.77)	Intergenic	RXFP1	77363
rs6476077	9	28499551	A	6.81E-05	0.28(0.15,0.52)	Intronic	LINGO2	-
rs10032252	4	159596924	G	7.22E-05	0.61(0.48,0.78)	Intergenic	RXFP1	65573
rs2183110	13	108686568	T	7.22E-05	0.20(0.09,0.44)	Intergenic	MYO16	28212
rs17040318	9	137605136	C	7.47E-05	0.37(0.23,0.61)	Intergenic	PAEP	6693
rs7283500	21	26265159	G	7.78E-05	1.79(1.34,2.39)	Intronic	APP	-
rs17303101	9	118221615	A	8.45E-05	0.63(0.50,0.79)	Intergenic	ASTN2	5713
rs16892284	6	161538841	T	9.16E-05	0.55(0.40,0.74)	Intronic	AGPAT4	-
rs924693	11	19200953	T	9.17E-05	1.62(1.27,2.05)	Intergenic	E2F8	1233
rs4584622	12	70959090	G	9.22E-05	1.64(1.28,2.10)	Intronic	TRHDE	-
rs7138083	12	117939030	C	9.73E-05	0.51(0.36,0.72)	Intronic	SRRM4	-

CHR: chromosome; BP: base position on chromosome; Allele, risk allele of the SNP was used as the reference allele in the regression; P, p-value; OR: odds ratio with 95% confidence interval; Location; annotated functional position; Nearby gene, nearest gene; Distance, base pairs (bp) to nearest genes, with "-" for intronic SNPs.

Table 2

Replication results in GENDEP for top SNPs from STAR*D.

SNP	STAR*D(95% CI)	STAR*D P	GENDEP OR (95% CI)	GENDEP P	Meta-P	Direction
rs7816924	2.14 (1.53, 2.99)	7.72E-06	1.58 (1.12, 2.21)	0.008	2.11E-07	++
rs10492002	0.61 (0.50, 0.75)	4.52E-06	0.86 (0.92, 1.47)	0.21	3.51E-06	--
rs1435964	0.62 (0.50, 0.77)	1.07E-05	0.86 (0.68, 1.08)	0.19	6.77E-06	--
rs12080794	0.57 (0.44, 0.75)	4.54E-05	0.73 (0.97, 1.92)	0.07	8.52E-06	--
rs924693	1.62 (1.27, 2.05)	9.17E-05	1.29 (0.61, 0.99)	0.04	9.97E-06	++
rs7132119	0.62 (0.50, 0.77)	2.39E-05	0.84 (0.67, 1.06)	0.15	1.05E-05	--
rs7283500	1.79 (1.34, 2.39)	7.78E-05	1.24 (0.91, 1.68)	0.17	3.53-E05	++
rs448038	2.82 (1.75, 4.57)	2.23E-05	1.14 (0.57, 1.34)	0.54	5.51E-05	++
rs4439856	1.71 (1.32, 2.21)	4.04E-05	1.06 (0.73, 1.22)	0.66	0.0001	++
rs11931209	0.60 (0.47, 0.76)	3.73E-05	0.96 (0.74, 1.46)	0.83	0.0002	++
rs2476230	1.69 (1.31, 2.17)	5.32E-05	1.06 (0.76, 1.48)	0.72	0.0002	++
rs2346793	0.60 (0.47, 0.77)	5.41E-05	0.94 (0.75, 1.49)	0.74	0.0002	--
rs4584622	1.64 (1.28, 2.10)	9.22E-05	1.09 (0.84, 1.69)	0.52	0.0002	++
rs7138083	0.51 (0.36, 0.72)	9.73E-05	0.89 (0.61, 1.30)	0.56	0.0002	--
rs10032252	0.61 (0.48, 0.78)	7.22E-05	0.95 (0.75, 1.49)	0.75	0.0003	--
rs37747	0.49 (0.35, 0.69)	4.86E-05	1.06 (0.62, 1.45)	0.80	0.0005	-+
rs7158881	0.64 (0.52, 0.79)	3.19E-05	1.20 (0.94, 1.53)	0.15	0.003	-+
rs17303101	0.63 (0.50, 0.79)	8.45E-05	1.33 (1.02, 1.74)	0.03	0.01	-+

OR: odds ratio for STAR*D or GENDEP with 95% confidence interval; P, p-value for STAR*D or GENDEP; Meta-P, combined p-value using Stouffer's z trend test; Direction, concordance of OR's, with "-" conferring risk for the unsustained response phenotype.

Table 3

Polygenic scoring in GENDEP from STAR*D SNP profiling.

SNP threshold	Coef	P	R squared explained
0.001	0.035	0.41	0.001
0.01	-0.004	0.79	0.0001
0.1	-0.002	0.74	0.0002
0.2	0.007	0.23	0.003
0.3	0.004	0.42	0.001
0.4	0.006	0.31	0.002
0.5	0.004	0.46	0.001