Published in final edited form as:
Pharmacogenet Genomics. 2009 September ; 19(9): 666. doi:10.1097/FPC.0b013e32832e4bcd.

Genome-wide Association Study of Suicidal Ideation Emerging During Citalopram Treatment of Depressed Outpatients

Gonzalo Laje, MD, MHSc[1], **Andrew S. Allen, PhD**[3], **Nirmala Akula, MS**[1], **Husseini Manji, MD**[2], **A. John Rush, MD**[6], and **Francis J. McMahon, MD**[1]

[1]Genetic Basis of Mood and Anxiety Disorders Unit, Mood & Anxiety Program, National Institute of Mental Health, National Institutes of Health, Dept. of Health & Human Services, Bethesda, MD

[2]Laboratory of Molecular Pathophysiology, Mood & Anxiety Program, National Institute of Mental Health, National Institutes of Health, Dept. of Health & Human Services, Bethesda, MD

[3]Department of Biostatistics and Bioinformatics, Duke University, Durham, NC

[6]Depts. of Clinical Sciences and Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX

Abstract

Suicidal ideation is an uncommon but worrisome symptom than can emerge during antidepressant treatment. We have previously described association between treatment emergent suicidal ideation (TESI) and markers in genes encoding the glutamate receptors GRIK2 and GRIA3. The present genome-wide association study was conducted to identify additional genetic markers associated with TESI that may help identify individuals at high-risk who may benefit from closer monitoring, alternative treatments, and/or specialty care. A clinically-representative cohort of outpatients with non-psychotic major depressive disorder enrolled in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial were treated with citalopram under a standard protocol for up to 14 weeks. DNA samples from 90 white participants who developed TESI and a gender and race matched equal number of treated participants who denied any suicidal ideas were genotyped with 109,365 single nucleotide polymorphisms on the Illumina's Human-1 BeadChip. One marker was found to

Correspondence: Gonzalo Laje, MD, MHSc. Genetic Basis of Mood and Anxiety Disorders, National Institute of Mental Health, 35 Convent Drive, Rm 1A207, Bethesda, MD 20892-3719; telephone: 301-496-7730 / fax 301-402-9081 / gonzalo.laje@nih.gov.

Previous Presentation: Data from this manuscript have been presented at the Society of Biological Psychiatry Annual Meeting, held from May 17 to May 19th, 2007, in San Diego, California, the World Congress on Psychiatric Genetics, held from October 7, 2007 to October 11 2007 in New York, NY and the American Society of Human Genetics Annual Meeting, held from October 23 to October 27, 2007, in San Diego, California

Contributions: Gonzalo Laje contributed to the study design, data analysis, interpretation and the manuscript drafting. Andrew Allen performed the statistical analysis. Nirmala Akula performed the genotyping. A. John Rush designed and executed STAR*D and also provided manuscript editing. Husseini Manji contributed to the study design, interpretation, and manuscript editing. Francis McMahon oversaw the study design, interpretation, and manuscript writing.

Disclosures and Acknowledgements: Dr. Allen and Ms Akula report no competing interests. Drs. Laje, Manji, Rush and McMahon are listed as co-inventors on a patent application the NIH has filed based in part on the diagnostic technology described in this paper. NeuroMark of Boulder, CO, executed a non-exclusive license with the NIH to develop this technology commercially. The authors were not involved in the negotiation and execution of this license, but under federal law the NIH is required to pay them a portion of any royalties the NIH receives under the license. The authors do not endorse any commercial use of the patent. Dr. Rush has served as an advisor, consultant, or speaker for or received research support from Advanced Neuromodulation Systems, Inc.; AstraZeneca; Best Practice Project Management, Inc.; Bristol-Myers Squibb Company; Cyberonics, Inc.; Eli Lilly & Company; Forest Pharmaceuticals, Inc.; Gerson Lehman Group; GlaxoSmithKline; Healthcare Technology Systems, Inc.; Jazz Pharmaceuticals; Magellan Health Services; Merck & Co., Inc.; the National Institute of Mental Health; Neuronetics; Ono Pharmaceutical; Organon USA Inc.; Pamlab; Personality Disorder Research Corp.; Pfizer Inc.; the Robert Wood Johnson Foundation; the Stanley Medical Research Institute; the Urban Institute; and Wyeth-Ayerst Laboratories Inc. He has equity holdings in Pfizer Inc and receives royalty/patent income from Guilford Publications and Healthcare Technology Systems, Inc.

be associated with TESI in this sample at the experiment-wide adjusted p<0.05 level (marker rs11628713, allelic p= 6.2×10^{-7}, OR = 4.7, permutation p=0.01). A second marker was associated at the experiment-wide adjusted p=0.06 level (rs10903034, allelic p = 3.02×10^{-6}, OR = 2.7, permutation p=0.06). These markers reside within the genes PAPLN and IL28RA, respectively. PAPLN encodes papilin, a protoglycan-like sulfated glycoprotein. IL28RA encodes an interleukin receptor. Together with our previous report, these findings may shed light on the biological basis of TESI and may help identify patients at increased risk of this potentially serious adverse event.

Keywords

treatment-emergent suicidal ideation; TESI; pharmacogenetics; citalopram; antidepressant; major depression; SSRI; STAR*D; genome-wide

Introduction

Most suicides are attributable to major depressive episode. Effective treatment with antidepressant medication prevents suicide. There is, however, an ongoing controversy over whether antidepressants, especially the class known as selective serotonin reuptake inhibitors (SSRIs), can actually trigger suicidal thoughts or behavior (suicidality) in some people. [1,2] The FDA determined that there was sufficient evidence to issue a “black box warning” highlighting the risk of treatment-emergent suicidality in those under the age of 25 yr prescribed antidepressants [3]. A subsequent decrease in the use of SSRIs, especially in children, [4] and a spike in the suicide rate among young people [5,6] have only served to heighten the controversy.

Antidepressants decrease suicidality in most patients, but may increase it in others. Thus, it would be useful to identify at-risk groups who may benefit from closer monitoring or alternative therapy while decreasing unnecessary concern among others that may lead to underuse of an otherwise safe and effective class of medications.

Actual suicide attempts during antidepressant treatment are uncommon [7], but suicidal ideation is always a concern. Two prior studies suggest that genetic markers may identify patients who experience treatment-emergent suicidal ideation (TESI) [8,9]. Both studies used data from the Sequenced Treatment Alternatives to Relieve Depression cohort (STAR*D) (www.star-d.org) (Rush, et al, Controlled Clinical Trials), but the two studies utilized somewhat different case definitions. Perlis et al. tested markers near CREB1 and detected evidence of association with onset or worsening of suicidal thoughts during citalopram treatment, but only in males. [9] In a study of 68 candidate genes, Laje et al. [8] detected an association between markers near GRIK2 and GRIA3 and the onset of thoughts of death or suicide during citalopram treatment. To our knowledge, no study has yet attempted to identify markers of TESI at the genome-wide level.

This report provides results of a genome-wide association study of TESI in the STAR*D cohort based on over 100,000 gene-based markers.

Methods

Sample

Ascertainment and evaluation for the STAR*D study are detailed elsewhere [10,11]. Briefly, investigators implemented a standard study protocol at 18 primary care and 23 psychiatric care settings across the United States. Participants provided separate written informed consents for the treatment study and the DNA collection. Participation in DNA collection was optional.

DNA collection was initiated approximately 12 months after initiation of the treatment study. Outpatients, 18-75 years of age, with a baseline 17-item Hamilton Depression Rating Scale [12,13] ($HRSD_{17}$) score of ≥14 who met DSM-IV [14] criteria for non-psychotic, major depressive disorder (MDD) were eligible. Patients with bipolar or psychotic disorders were excluded, as were those with a primary diagnosis of obsessive-compulsive or eating disorders, general medical conditions that contraindicated study medications, substance dependence requiring inpatient detoxification, clear non-response or intolerance to any protocol antidepressant during the current episode, or those who were pregnant or breast-feeding [10, 11].

At the first treatment step, all participants received the SSRI citalopram, typically starting at 20 mg/day, with dose increases following recommended procedures (up to 40mg/day by week 4 and 60mg/day by week 6) [15]. The protocol required an adequate dose of citalopram for a sufficient time to ensure that those who did not improve were most likely unresponsive to the medication [15]. No concomitant psychotropic medications were allowed, aside from benzodiazepines, hypnotics, or trazodone (up to 200 mg/d) for sleep, if needed. The sample characteristics were presented elsewhere [15,16].

DNA samples were collected from 1953 participants. A sample of 20 ml of whole blood was collected and shipped to the Rutgers Cell Repository, where lymphocytes were extracted and cryopreserved using standard methods. DNA was extracted using Gene-Pure chemistry (Qiagen). Samples were arrayed robotically, then gender-verified with a set of three X-linked and two Y-linked markers. A CONSORT (Consolidated Standards of Reporting Trials) diagram of the current study sample is shown elsewhere [8].

Those who consented to DNA collection were similar to those in the full sample, but differed on some variables (for details, see [16]). These differences cannot affect the genetic association results, which derive from comparisons among the genotyped subjects, but do limit generalizability. Those who provided DNA did not differ in the frequency of TESI or in those variables that were associated with TESI in this sample [8].

All of the 90 white (including Hispanics), participants who developed TESI and a gender, race and ethnicity matched set of 90 treated participants who denied any history of suicide attempts and suicidal ideas during their citalopram treatment were used for the genome-wide scan. Due to limited resources we could not genotype the whole STAR*D sample, so to improve power controls were further selected based on a minimum of 3 visits, a baseline QIDS-C total score (CRC) >=10, adherence to medication, recurrent depression, no history of suicide attempts and further, no report of suicidal thoughts throughout citalopram treatment on $QIDS\text{-}C_{16}$ or $QIDS\text{-}SR_{16}$. An expanded set of 94 controls, selected on the same criteria, were added post-hoc to test the robustness of the findings.

Phenotype Definition

Consistent with previous definitions of TESI [8,9], we used the “Thoughts of Death or Suicide” question (item # 12) on the 16-item Quick Inventory of Depressive Symptomatology – Self-report ($QIDS\text{-}SR_{16}$), a reliable and well-validated measure of symptom severity that has been shown to correlate well with the Hamilton Depression Rating Scale [17-21]. The $QIDS\text{-}SR_{16}$ was chosen here over the clinician-rated version (the $QIDS\text{-}C_{16}$) since TESI is a subjective phenomenon and we wished to avoid clinician bias and patient reluctance to verbally report the phenomenon to the clinician.

The $QIDS\text{-}SR_{16}$ was administered at baseline and at each of the protocol-recommended clinic visits at baseline, weeks 2, 4, 6, 9, and 12. Possible responses to item #12 include: “I do not think of suicide or death” (=0), “I feel that life is empty or wonder if it's worth living”(=1), “I

think of suicide or death several times a week for several minutes" (=2), and "I think of suicide or death several times a day in some detail, or I have made specific plans for suicide, or have actually tried to take my life (=3). Participants who scored "0" on this item before citalopram treatment and 1, 2 or 3 at least once during treatment were defined as TESI cases. Although a score is obtained from the scale's item this trait cannot be considered quantitative since the intervals between anchor points are not equal.

Population structure

We evaluated the possibility of population structure leading to inflated association results by using as covariates in the primary allelic test the posterior probabilities calculated by STRUCTURE [22,23] to evaluate the worst case scenario of 2 cryptic populations within the white subset (k=2). A set of 344 unlinked SNPs selected from the total dataset was run on STRUCTURE for 20,000 burn-in steps followed by 20,000 replications (admixture model) to assign to each individual a posterior probability of ancestry for each of the inferred populations. Additionally, to detect any residual bias we plotted the observed −log10 p-values against the expected −log10 p-values under the null hypothesis (Q-Q plot) (figure 1).

Selection of single nucleotide polymorphism (SNP) markers

We used the Illumina Human-1 BeadChip. This chip samples 109,365 single nucleotide polymorphisms, 48% fall within 10K of exons, 23% of the snps are in the transcripts, 15% fall in highly conserved regions, and 14% are uniformly spaced in the genome.

Genotyping methods

A total of 90 white TESI cases and 90 gender, race and ethnicity matched controls were genotyped the Illumina Beadstation 500 according to manufacturer's recommended protocol. The markers of interest were further genotyped in the total sample using Taqman Assay (Applied Biosystems, Foster City, CA) followed by scanning the samples in Analyst HT (Molecular Devices, Sunnyvale, CA). In the 274 samples that were genotyped with both the Taqman and Illumina assays, there was 100% agreement at the genotype level. The genotypes are scored by a clustering algorithm (kindly provided by Sam Chen, Virginia Commonwealth University). The call rates for this sample were 99.5%. For quality control purposes 8501 SNPs were excluded if missingness exceeded 10%, if minor allele frequencies were lower than 5% and if Hardy –Weinberg deviation was lower than 1×10^{-3}. No individuals were excluded due to gender mismatches.

Statistical methods

Stratified score tests were developed from a retrospective logistic regression model adjusted for confounders via stratum specific intercepts (see appendix for details). In this analysis, autosomal markers were tested using a standard allelic dosing model in which each extra copy of the minor allele has a multiplicative effect on risk. In females, the genetic model used for X-linked markers also corresponded to an allelic dosing model. In males, however, the genetic model was parameterized so that the penetrance of a male with the minor allele would correspond to the penetrance of a female with two minor alleles. This parameterization was chosen because it corresponds to the allelic dosing model under X-inactivation. The score tests developed here are asymptotically equivalent to the "allelic" likelihood ratio test statistics given by the software program UNPHASED [24]. However, unlike the likelihood ratio tests, the score tests can be given in closed form (details in appendix) making for extremely rapid computation. All analyses were stratified on each individual's most likely subpopulation membership as output from the structure 2 subpopulation model detailed above.

We controlled the experiment-wide error rate by using a standard permutation scheme. This scheme aims at breaking any correlation between the SNP markers and the phenotype, reflecting a global null hypothesis, while maintaining the correlation structure among the markers. To implement, we randomly permuted the phenotypes between individuals to form permutation datasets. Since adjusted analyses were conducted, permutations were performed separately within each strata. On each resulting permutation dataset an entire genome scan was performed and the minimum p-value noted. By repeating the permutation/analysis steps a large number of times (50,000), we were able to estimate the permutation distribution of the minimum p-value under the global null hypothesis. By comparing our original raw p-values with the permutation distribution allowed us to assess their experiment-wide significance and compute permutation adjusted p-values.

Additional models, such as genotypic or haplotype tests, could have been performed but given the potential for correction of a large number of tests, no other models beyond the allelic were tested in our primary analyses. We ran a genotypic model on a post-hoc basis.

Clinical parameters (risk) calculation

The effect of the two SNPs presented here (rs11628713 and rs10903034) as well as two previously published SNPs [8] found in GRIA3 (rs4825476) and GRIK2 (rs2518224) on an individual's risk of developing TESI was estimated, after supplemental genotyping of rs11628713 and rs10903034, from the total STAR*D sample (n=1915). Since this sample represents a prospective study of TESI development, risk estimation was conducted directly on this sample using a standard multivariable logistic regression model. The adequacy of model fit was assessed via the Hosmer and Lemeshow test [25].

Results

Descriptive variables

Demographic and clinical characteristics of the sample are presented elsewhere [8]. There were no significant demographic differences between cases (n=90) and controls (n=90), and no differences in several clinical variables that are known predictors of suicide [26]. Consistent with our previous report [8] in this subsample, TESI cases also received a significantly higher maximum citalopram dose, were significantly less likely to go into remission (full recovery, as defined previously [15,16]), and were more likely to move to the secondary treatment step following citalopram [15]. As these variables could act as potential confounders, we performed a logistic regression that included these variables and the markers of interest. The results suggest these variables are not significant confounders (Supplemental table 1).

Association analysis

Figure 2 presents results of the genetic association test. Two markers, rs11628713 and rs10903034 stand out. Both markers were in Hardy-Weinberg equilibrium in this sample.

One marker, rs11628713 (allelic $p= 6.2 \times 10^{-7}$, OR = 4.9, permutation p=0.01) was associated with TESI at the experiment-wise $p<0.05$ level. This is an intronic marker at about 50bp from exon 13 in the PAPLN gene in chromosome 14q24.2. This SNP is a C/T polymorphism in a non-conserved region where the ancestral C-allele is over-represented in TESI cases (allele frequency in cases 0.93 and controls 0.74). The association findings remained significant in the expanded reference set (total controls n=184), but the effect size was slightly lower (OR 4, 95%CI 2.08, 7.46).

The other marker, rs10903034, was associated with TESI at an experiment-wise p=0.06 level (allelic $p =3.02 \times 10^{-6}$, OR = 2.7, permutation = 0.06). This marker resides in chromosome

1p36.11 within the 3′UTR of the IL28RA gene. The ancestral T-allele is over-represented in TESI cases (allele frequency in cases 0.68 and controls 0.44).

One additional marker in PAPLN and four additional markers in IL28RA were associated with TESI in this sample (Supplemental table 2). A post hoc genotypic test was conducted, but no additional markers were found to have experiment-wise significance (data not shown).

Finally, rs11628713 (PAPLN) and rs10903034 (IL28RA) were genotyped in the whole STAR*D sample (N=1915, cases n=120, controls n=1742, missing n=53). The PAPLN marker had an allele frequency in cases of 0.91 and 0.83 in controls (OR: 2.1, p=0.001). The IL28RA marker had an allele frequency in cases of 0.61 and 0.54 in controls (OR: 1.33, p=0.037).

Association with previously described markers in GRIK2 and GRIA3

The markers previously associated with TESI are not included on the Illumina Infinium I chip used in the present study. As seen in supplemental figure 1a, there were no markers included on the Illumina Infinium I array that sampled rs2518224 through linkage disequilibrium in this population. However, marker rs5911588 does sample rs4825476 at an r^2 value of 0.46 (supplemental figure 1b), representing moderate correlation between these neighboring markers. As expected, rs5911588 is associated with TESI (p = 0.0004), consistent with the findings previously described in this sample.

Additive effect of risk alleles

Figure 3 presents risk estimates derived from a multiple logistic regression model which includes additive effects of four genotypes (rs11628713, rs10903034, rs4825476 and rs2518224) as well as the effect of sex in the whole STAR*D sample (n=1915). Race and ethnicity were not found to have significant main effects or to significantly modify genotypic effect parameters. Hence, race and ethnicity were excluded from the final model. The final model was found to provide a reasonable fit to the observed data (Hosmer and Lemeshow p=0.7; small p-values would indicate a poor model fit).

The highest risk of developing TESI in the context of citalopram treatment would occur with one or two G alleles in GRIA3, the CC genotype in GRIK2, the CC genotype in PAPLN and the AA genotype in IL28RA, a very uncommon combination. All other possible allelic combinations and their predicted risks are shown in (figure 3).

Discussion

To our knowledge this is the first genome-wide study to detect genetic markers related to TESI. We found two novel markers that were associated with TESI in this sample. Although the marker on IL28RA did not achieve experiment-wise significance at the 0.05 level, we chose to make it a part of the predictive model because of the strength of the association, the consistency observed in the signal of neighboring markers, and biological plausibility. However, until functional alleles are demonstrated or replication is shown in an independent sample, these findings must be viewed as preliminary.

Due to the relatively low frequency of TESI, replication will require a very large prospectively ascertained treatment sample or a retrospective case-control design. It remains unclear whether our findings extend to other antidepressants. However, due to the extensive use in psychiatry and primary care of citalopram, and its low risk for pharmacokinetic interactions at the cytochorme P-450-2D6 level, these findings have relevance even if confined to citalopram and related compounds. Moreover, the effect size estimates reported herein are likely to be inflated (winner's curse) and only after replication in independent samples we would be able to estimate

the “true” effect size. Until such experiments have been carried out it is impossible to make statements about any clinical utility of the reported markers.

This study has several limitations. Although consistent with prior reports [8,9], TESI was defined on the basis of a single item on a depression rating scale. Neither the instrument, the item, nor the STAR*D study itself, were designed primarily to address TESI. Due to the lack of a placebo group it is impossible to determine what fraction of TESI in this study is directly attributable to antidepressant treatment. These limitations are characteristic of all existing large antidepressant treatment samples, suggesting the need for future large, placebo-controlled studies.

Although the present study took a genome-wide approach, covering over 100,000 markers, current genotyping technology now allows for the study of at least 5 times more potentially relevant polymorphisms. Moreover, as described above, the design of the Illumina Infinium I chip is exon-centric. Therefore, the markers used did not sample potential regulatory regions between genes. Of course, factors other than sequence polymorphisms may be of relevance in TESI, such as copy number variation, methylation and acetylation patterns, and environmental influences.

As in all case-control genetic association studies, hidden population stratification is always a risk. When the trait of interest and marker allele frequencies differ among ancestral groups, spurious associations (and inflated test statistics) may arise. We have previously shown that there is no relationship between race and TESI in this sample. [8] We further adjusted the results for any differences in allele frequencies, using covariates. The final distribution of p-values was consistent with expectations under the null hypothesis (Figure 1), and the genomic inflation factor ($lambda_{GC}$), based on the median of all the observed chi-square values, was 1.00. Thus, our findings are unlikely to be the result of population stratification.

PAPLN is a 37kb, 26 exon gene on chromosome 14 that encodes a proteoglycan-like sulfated glycoprotein called papilin. There is a paucity of information about papilin in humans.

The other gene implicated in this study, IL28RA, is a 33kb, 6 exon gene on chromosome 1 that encodes for a protein in the class II cytokine receptor family. Viral infection can induce the expression of interleukin 29 (IL29) (or interferon lambda 1), interleukin 28A (IL28A) (or interferon lambda 2) and interleukin 28B (IL28B) (or interferon lambda 3) [27]. While little is known about the role of this receptor or interleukin 28 (interferon lambda 2), or their relevance to mood disorders and suicidality, it is notable that the therapeutic use of interferon alpha is associated with a significant incidence of depression and suicidality [28].

In summary, we have identified markers in two novel genes that together with previous reports may shed light on the biological basis of suicidal ideation that emerges during antidepressant treatment. If replicated, these findings may take us a step closer toward developing markers with clinically meaningful predictive value for this potentially serious adverse event. Further work is needed to replicate these findings and uncover the functional variation that underlies the association signals we observe.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was funded by the Intramural Research Program of the National Institute of Mental Health, NIH, by a NARSAD Independent Investigator Award to FJM, Duke Clinical Research Institute and NIH #1K25-077663 to ASA.

The authors thank the STAR*D research team for acquisition of clinical data and DNA samples and Forest Laboratories for providing citalopram at no cost for the STAR*D study. Data and sample collection were funded with federal funds from the NIMH, NIH, under contract N01MH90003 to University of Texas Southwestern Medical Center at Dallas (Principal Investigator, A. John Rush).

We thank Jo Steele for technical assistance, the Rutgers University Cell and DNA Repository for extracting DNA and providing samples to our laboratories, and most importantly, we thank the participants of the STAR*D trial, without whom this study would not be possible. The content of this publication does not necessarily reflect the views or policies of the DHHS, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Appendix

Retrospective score tests for autosomal SNP/trait association

Denote genotype by G and assume that it is coded as the number of minor alleles present so that $G = 1$ indicates a heterozygote while $G = 2$ indicates a homozygote for the minor allele. Let cases be indicated by $D = 1$ and controls by $D = 0$. Assume there are t strata indicated by the random variable Z. Let d_{jz} and c_{jz} be the number of cases and controls, respectively, having genotype j in stratum z. Let $n_{jz=}c_{jz} + d_{jz}$ be the total number of individuals in the study having genotype j in stratum z. Let $d_z = d_{0z} + d_{1z} + d_{2z}$ and $c_z = c_{0z+}c_{1z} + c_{2z}$ We also define p_z, the frequency of the minor allele in stratum z, by $p_z = \exp(\gamma_z)/\{1+\exp(\gamma_z)\}$. Assume Hardy-Weinberg equilibrium among the controls in stratum z and that the odds of being a case are given by

$$\frac{\Pr\left(D=1\middle|G=j,Z=z\right)}{\Pr\left(D=0\middle|G=j,Z=z\right)}=\exp\left(\alpha_z+X_j^T\beta\right),$$

where X_j is a design vector that codes the genetic model. Then the likelihood, conditional on both disease status and strata, can be written as

$$\prod_{z=1}^{t}\prod_{j=0}^{2}\left(\frac{\exp\{j\gamma_z+I(j=1)\ln(2)\}}{\sum_{j^*=0}^{2}\exp\{j^*\gamma_z+I(j^*=1)\ln(2)\}}\right)^{c_{jz}} \times\left(\frac{\exp\left\{X_j^T\beta+j\gamma_z+I(j=1)\ln(2)\right\}}{\sum_{j^*=0}^{2}\exp\left\{X_{j^*}^T\beta j^*\gamma_z+I(j^*=1)\ln(2)\right\}}\right)^{d_{jz}},$$

where $I(\ )$ is an indicator function. Note that the intercept parameter α_z cancel out of the likelihood. From this likelihood and assuming an allelic dosing model in which $X_j = j$, it is easy to derive the score statistic U under the null hypothesis $\beta = 0$, i.e.,

$$U=\sum_{z=1}^{t}\sum_{j=0}^{2}d_j\ (j_z-2\widehat{p_z}),$$

where $\hat{p}_z = (2n_{2z} + n_{1z})/ 2n_z$. The variance of U_β is given by

$$V=\sum_{t=1}^{z}\frac{d_z c_z}{n_z}2\widehat{p_z}\ (1-\widehat{p_z})\,.$$

A score test is formed by $T = U^2/V$. Asymptotically, T will have a χ^2 distribution on 1 degree of freedom. Note that the only d_{jz} in U varies from permutation to permutation. The minor allele frequencies within strata do not vary and can be stored for computation of U The variance V is completely invariant to permuation and therefore does not need to be recomputed each iteration.

Retrospective score tests for SNP/trait association on the X chromosome

We modify our notation by allowing the above quantities to be indexed by the sex of the patient. For example, we let d_{jsz} and c_{jsz} be the number of cases and controls, respectively, having genotype j and sex s in stratum z. We assume Hardy-Weinberg equilibrium among the controls in stratum z (note that Hardy-Weinberg here implies Hardy-Weinberg frequencies in the females and similar allele frequencies between males and females). Let the odds of being a case are given by

$$\frac{\Pr\left(D=1\middle|G=j,Z=z,S=s\right)}{\Pr\left(D=0\middle|G=j,Z=z,S=s\right)}=\exp\left(\alpha_{zs}+X_{js}^{T}\beta\right),$$

Then the likelihood can be written as

$$\prod_{s=m}^{f}\prod_{z=1}^{t}\prod_{j=0}^{2^{I(s=f)}}\left(\frac{\exp\{j\gamma_z+I(j=1,s=f)\ \mathrm{In}\ (2)\}}{\sum_{j^*=0}^{2}\exp\{j^*\gamma_z+I(j^*=1,s=f)\ \mathrm{In}\ (2)\}}\right)^{c_{jz}}$$
$$\times\left(\frac{\exp\left\{X_{js}^{T}\beta+j\gamma_z+I(j=1,s=f)\ \mathrm{In}\ (2)\right\}}{\sum_{j^*=0}^{2}\exp\left\{X_{js^*}^{T}\beta j^*\gamma_z+I(j^*=1,s=f)\ \mathrm{In}\ (2)\right\}}\right)^{d_{jz}}.$$

From this likelihood and assuming a genetic model in which $X_{js} = j2^{I(s=m)}$, we can derive the score statistic U under the null hypothesis $\beta = 0$, i.e.,

$$U=\sum_{s=m}^{f}\sum_{z=1}^{t}\sum_{j=0}^{2^{I(s=f)}}d_{jsz}\left(j2^{I(s=m)}-2\widehat{p_z}\right),$$

where $p\hat{}_z\ (2n_{2fz} + n_{fz} + n_{1mz})/(2_{fz} + n_{mz})$. The variance of U is given by

$$V=\sum_{t=1}^{z}\frac{2d_{mz}+d_{fz}-2d_z^2}{n_{mz}+2n_{fz}}2\widehat{p_z}\,(1-\widehat{p_z})\,.$$

Again, a score test is formed by $T = U^2/V$ which will asymptotically be distributed as χ^2 on 1 degree of freedom. Again, only d_{jsz} in U varies from permutation to permutation.

References

1. Teicher MH, Glod C, Cole JO. Emergence of intense suicidal preoccupation during fluoxetine treatment. Am J Psychiatry 1990;147:207–210. [PubMed: 2301661]
2. Khan A, Khan S, Kolts R, Brown WA. Suicide rates in clinical trials of SSRIs, other antidepressants, and placebo: analysis of FDA reports. Am J Psychiatry 2003;160:790–792. [PubMed: 12668373]
3. Food and Drug Administration. Antidepressant Use in Children, Adolescents and Adults. 2005 Online. 2-5-2006.
4. Murray ML, Thompson M, Santosh PJ, Wong IC. Effects of the committee on safety of medicines advice on antidepressant prescribing to children and adolescents in the UK. Drug Saf 2005;28:1151–1157. [PubMed: 16329717]
5. Centers for Disease Control and Prevention. Suicide Trends Among Youths and Young Adults Aged 10--24 Years --- United States, 1990--2004. Online. 9-5-2007. 12-15-2007.
6. Gibbons RD, Brown CH, Hur K, Marcus SM, Bhaumik DK, Erkens JA, et al. Early evidence on the effects of regulators' suicidality warnings on SSRI prescriptions and suicide in children and adolescents. Am J Psychiatry 2007;164:1356–1363. [PubMed: 17728420]
7. Food and Drug Administration. Suicidality in Children and Adolescents Being Treated With Antidepressant Medications. Online. 10-15-2004. 2-5-2006.
8. Laje G, Paddock S, Manji H, Rush AJ, Wilson AF, Charney D, et al. Genetic markers of suicidal ideation emerging during citalopram treatment of major depression. Am J Psychiatry 2007;164:1530–1538. [PubMed: 17898344]
9. Perlis RH, Purcell S, Fava M, Fagerness J, Rush AJ, Trivedi MH, et al. Association between treatment-emergent suicidal ideation with citalopram and polymorphisms near cyclic adenosine monophosphate response element binding protein in the STAR*D study. Arch Gen Psychiatry 2007;64:689–697. [PubMed: 17548750]
10. 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. x. [PubMed: 12778843]
11. 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. Control Clin Trials 2004;25:119–142. [PubMed: 15061154]
12. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62. [PubMed: 14399272]
13. Hamilton M. Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 1967;6:278–296. [PubMed: 6080235]
14. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV). Washington, DC: American Psychiatric Association Press Inc.; 1994.
15. 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]
16. 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]
17. Rush AJ, Carmody TJ, Reimitz PE. The Inventory of Depressive Symptomatology (IDS): clinician (IDS-C) and self-report (IDS-SR) ratings of depressive symptoms. Int J Methods Psychiatr Res 2000;9:45–59.
18. 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]

19. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol Med 2004;34:73–82. [PubMed: 14971628]

20. Rush AJ, Trivedi MH, Carmody TJ, Ibrahim HM, Markowitz JC, Keitner GI, et al. Self-reported depressive symptom measures: sensitivity to detecting change in a randomized, controlled trial of chronically depressed, nonpsychotic outpatients. Neuropsychopharmacology 2005;30:405–416. [PubMed: 15578008]

21. Rush AJ, Bernstein IH, Trivedi MH, Carmody TJ, Wisniewski S, Mundt JC, et al. An evaluation of the quick inventory of depressive symptomatology and the hamilton rating scale for depression: a sequenced treatment alternatives to relieve depression trial report. Biol Psychiatry 2006;59:493–501. [PubMed: 16199008]

22. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in structured populations. Am J Hum Genet 2000;67:170–181. [PubMed: 10827107]

23. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000;155:945–959. [PubMed: 10835412]

24. Dudbridge F, Koeleman BP. Efficient computation of significance levels for multiple associations in large studies of correlated data, including genomewide association studies. Am J Hum Genet 2004;75:424–435. [PubMed: 15266393]

25. Hosmer, D.; Lemeshow, S. Applied Logistic Regression. Second. New York: Wiley; 2000.

26. Hirschfeld RM, Russell JM. Assessment and treatment of suicidal patients. N Engl J Med 1997;337:910–915. [PubMed: 9302306]

27. Vilcek J. Novel interferons. Nat Immunol 2003;4:8–9. [PubMed: 12496969]

28. Janssen HL, Brouwer JT, van der Mast RC, Schalm SW. Suicide associated with alfa-interferon therapy for chronic viral hepatitis. J Hepatol 1994;21:241–243. [PubMed: 7989716]

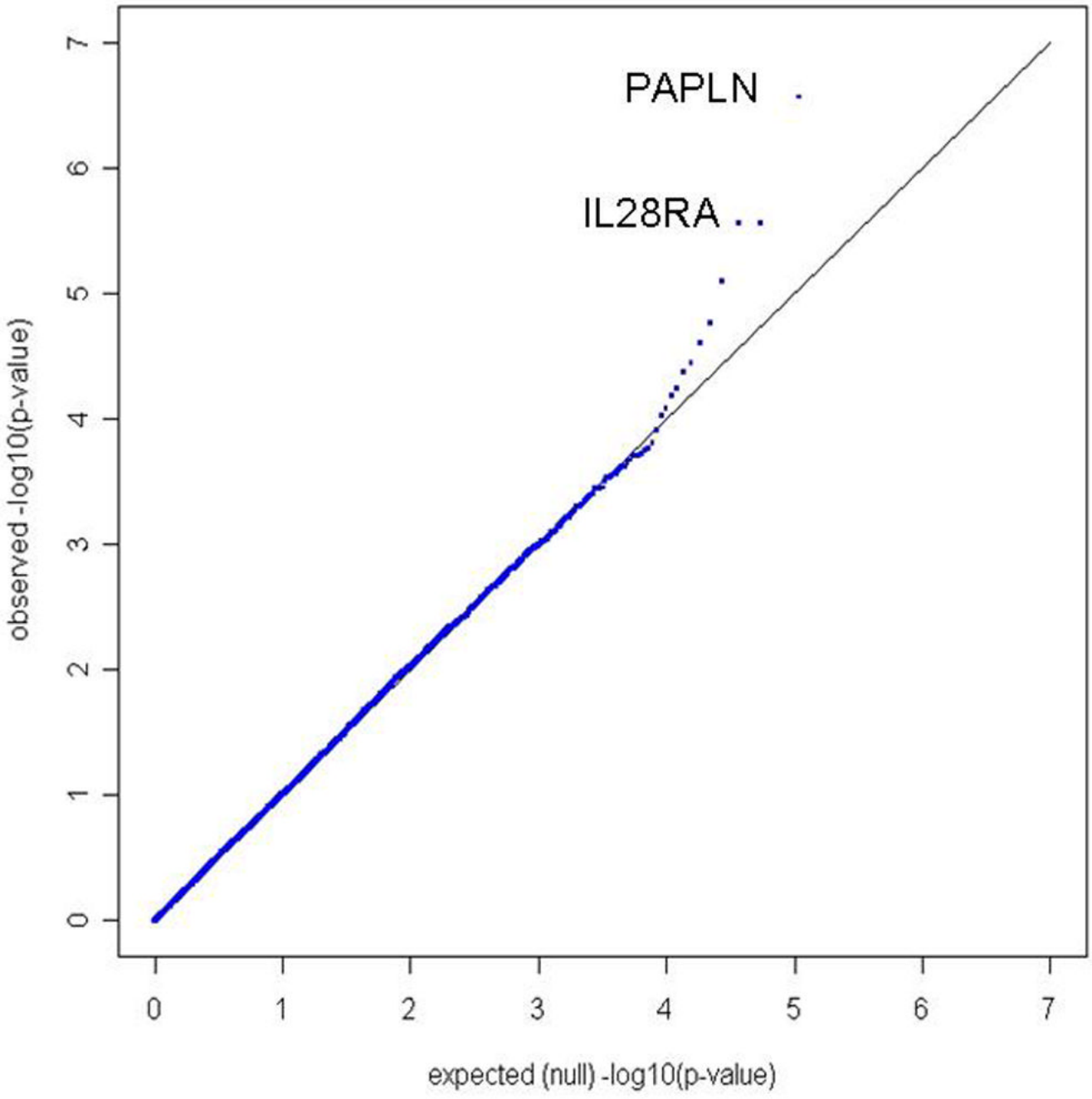


Figure 1.
Q-Q plot used for residual bias detection shows the observed –log10 p-values against the–log10 p-values expected under the null hypothesis.

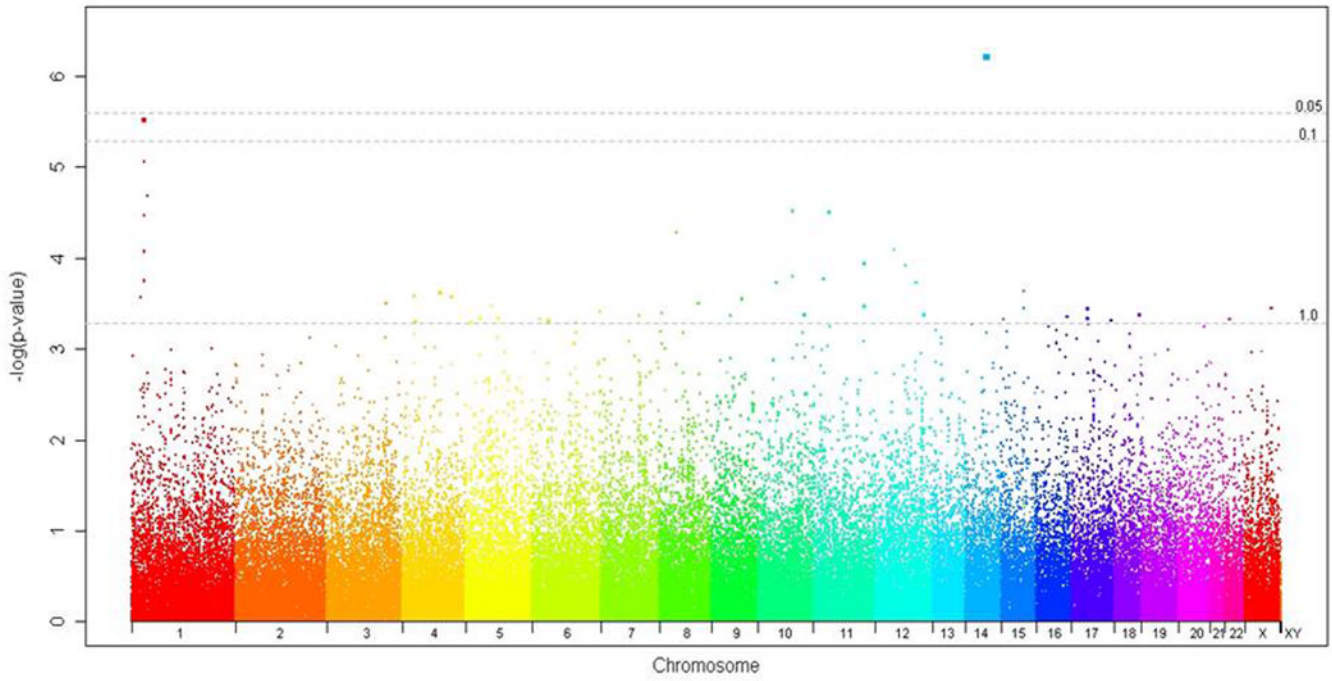
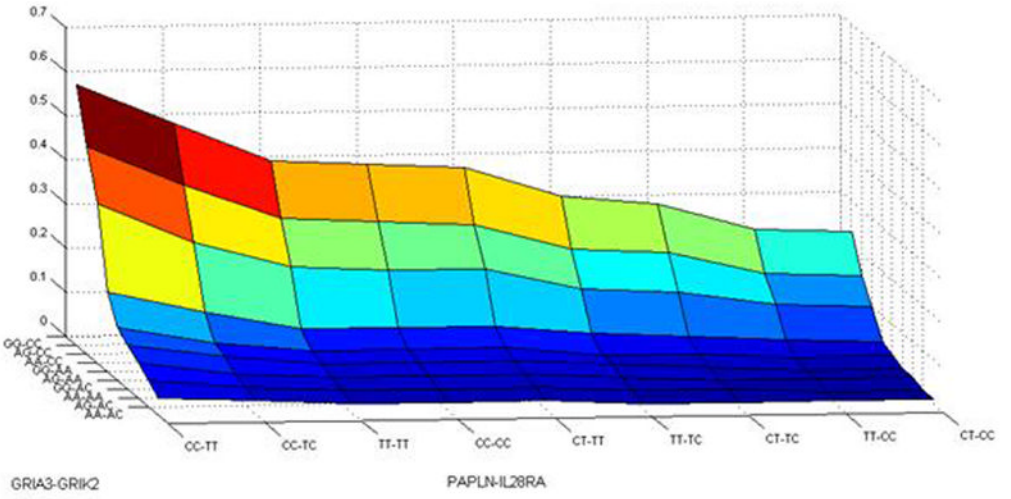
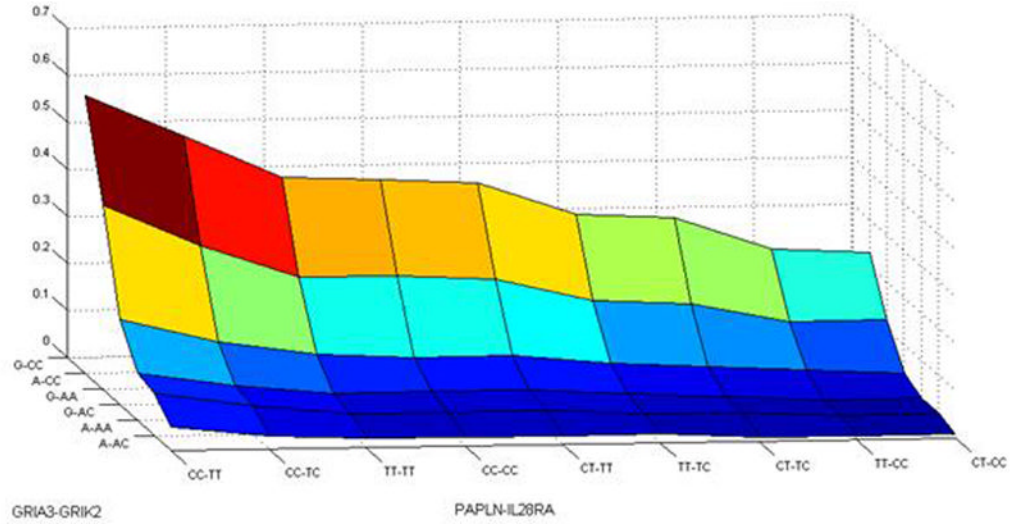


Figure 2.
Allelic comparisons for each of the 109,365 markers, ordered by physical position. The dashed lines represent permutation-corrected values.



A



B

Figure 3.

Risk estimates derived from a multiple logistic regression model which includes additive effects of four genotypes (rs11628713, rs10903034, rs4825476 and rs2518224) by gender (males above). Figures are separated by gender (A: females, B: males) since one of the markers resides on a chromosome X and thus males are hemizygous for this variant. The x-axis represents PAPLN and IL28RA genotypes, the y-axis represents the risk estimate, and the z-axis represents GRIA3 and GRIK2 genotypes