Race, Genetic Ancestry and Response to Antidepressant Treatment for Major Depression

Eleanor Murphy*[1], Liping Hou[1], Brion S Maher[2], Girma Woldehawariat[1], Layla Kassem[1], Nirmala Akula[1], Gonzalo Laje[1] and Francis J McMahon[1]

[1]Human Genetics Branch, Intramural Research Program, National Institute of Mental Health, National Institutes of Health, US Department of Health and Human Services, Bethesda, MD, USA; [2]Johns Hopkins School of Public Health, Baltimore, MD, USA

The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Study revealed poorer antidepressant treatment response among black compared with white participants. This racial disparity persisted even after socioeconomic and baseline clinical factors were taken into account. Some studies have suggested genetic contributions to this disparity, but none have attempted to disentangle race and genetic ancestry. Here we used genome-wide single-nucleotide polymorphism (SNP) data to examine independent contributions of race and genetic ancestry to citalopram response. Secondary data analyses included 1877 STAR*D participants who completed an average of 10 weeks of citalopram treatment and provided DNA samples. Participants reported their race as White ($n = 1464$), black ($n = 299$) or other/mixed ($n = 114$). Genetic ancestry was estimated by multidimensional scaling (MDS) analyses of about 500 000 SNPs. Ancestry proportions were estimated by STRUCTURE. Structural equation modeling was used to examine the direct and indirect effects of observed and latent predictors of response, defined as change in the Quick Inventory of Depressive Symptomatology (QIDS) score from baseline to exit. Socioeconomic and baseline clinical factors, race, and anxiety significantly predicted response, as previously reported. However, direct effects of race disappeared in all models that included genetic ancestry. Genetic African ancestry predicted lower treatment response in all models. Although socioeconomic and baseline clinical factors drive racial differences in antidepressant response, genetic ancestry, rather than self-reported race, explains a significant fraction of the residual differences. Larger samples would be needed to identify the specific genetic mechanisms that may be involved, but these findings underscore the importance of including more African-American patients in drug trials.

Neuropsychopharmacology (2013) **38**, 2598–2606; doi:10.1038/npp.2013.166; published online 24 July 2013

Keywords: Pharmacology; Genetics; Ethnicity; African Americans; SSRI

INTRODUCTION

Recent studies have highlighted a racial disparity in antidepressant treatment outcomes with generally poorer results for African Americans compared with Whites (Lesser *et al*, 2007). Although most studies have found that environmental factors such as lower socioeconomic status (SES), greater medical or psychiatric comorbidity, and higher dropout contribute to the poorer outcomes among African Americans (Warden *et al*, 2007; Warden *et al*, 2009; Lesser *et al*, 2010; Lesser *et al*, 2011), other studies have suggested that additional biological or genetic factors may also play an important role (Strickland *et al*, 1997; Luo *et al*, 2006; McMahon *et al*, 2006; Lesser *et al*, 2007). However, conclusions from both types of studies have been limited by

*Correspondence: Dr E Murphy, Human Genetics Branch, Intramural Research Program, National Institute of Mental Health, National Institutes of Health, US Department of Health and Human Services, 35 Convent Drive, Building 35, Porter Bldg, RM 1A-209, Bethesda, MD 20892, USA, Tel: +1 301 451 3813, Fax: +1 304 402 9081, E-mail: eleanor.murphy@nih.gov

methods that were unable to separate individual effects among many intercorrelated variables.

The idea that there may be biological—specifically genetic—reasons for racial differences in response to psychopharmacological treatment goes back at least as far as first-generation antidepressants and mood stabilizers. For example, pharmacokinetic research on the tricyclic antidepressants (TCAs) has identified several polymorphisms in genes encoding P-450 isoenzymes (*2D6* and *2C19*) that affect drug metabolism. Compared with people of European ancestry, African Americans were more likely to carry alleles that slow TCA metabolism, leading to a 'slow metabolizer' phenotype, higher blood TCA plasma level, and more rapid response—but greater adverse effects or toxicity (Henry *et al*, 1971; Raskin and Crook 1975; Strickland *et al*, 1991; Strickland *et al*, 1997). Further studies examined the effects of additional P-450 variants (eg, *2D6* and *2C19*) on SSRI antidepressants such as paroxetine, sertraline, and fluoxetine, but found no racial or ethnic group differences in drug response (Keers and Aitchison, 2011).

Candidate gene studies of the pharmacodynamic components of SSRI response, including the extensively studied

serotonin transporter gene, *SLC6A4*, and variants in the serotonin receptor genes, *HTR1A* and *HTR2A*, have shown differences in allele frequencies among racial and ethnic groups (Lotrich *et al*, 2003; McMahon *et al*, 2006; Ruhe *et al*, 2009). On the other hand, small sample sizes, uncertain genotype–phenotype associations, and other differences in study design have often led to ambiguous or contradictory conclusions. For example, different studies suggest that African Americans respond better or worse than European Americans to SSRIs, and a few studies have found no differences (Keers and Aitchison, 2011). Thus, it remains unclear what role these genetic variants have in treatment response differences among racial and ethnic groups.

In addition to genetic contributions, baseline clinical differences may also influence treatment outcomes. For example, comorbid anxiety in depressed patients is associated with lower remission and response rates (Brown *et al*, 1996; Trivedi *et al*, 2006; Baldwin and Lopes, 2009). In recent clinical trials including the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study, African Americans reported more comorbid anxiety disorders and more anxious symptoms than other racial groups (Lesser *et al*, 2007; Lesser *et al*, 2011). This is consistent with earlier studies of psychiatric patients suffering from unipolar depression, among whom African Americans reported more worry and exhibited more muscular tension, general anxiety, and autonomic symptoms than others (Simon *et al*, 1973; Uhlenhuth and Paykel, 1973; Fabrega *et al*, 1988). However, little is known about how anxiety might mediate the association between race and ethnicity in antidepressant treatment outcome.

Race is a social construct; however, it also serves as a surrogate for biological differences. This duality presents a challenge for researchers investigating health disparities associated with race or ethnicity. Race-associated differences in treatment outcomes may reflect ancestral genetic variation, social disparities that are correlated with race, or both. Race is a poor proxy for biological differences across populations, leading some to argue that it should be abandoned in medical research (Yaeger *et al*, 2008; Gravlee *et al*, 2009; Lee 2009). Indeed, racial differences tend to disappear or diminish after social factors are taken into account. However, a large body of data over many years has been classified by race, pointing to persistent disparities in a variety of health outcomes.

By use of genetic marker data, it is now possible to estimate proportions of genetic ancestries that comprise an individual's total genomic ancestry and assess to what degree these estimates correlate with race, social variables, and health outcomes. These methods also allow us to partition out the direct effects of self-reported race from those of genetic ancestry and test for mediation by social variables (Wassel *et al*, 2011; Sucheston *et al*, 2012).

The primary objective of the present study is to examine the extent to which self-reported race and genetic ancestry (here and later in the paper, genetic ancestry refers specifically to African ancestry unless otherwise noted) independently predict therapeutic response to first-level treatment of major depression with citalopram. A secondary objective is to investigate whether anxiety mediates any apparent relationships between race, genetic ancestry, and therapeutic response.

MATERIALS AND METHODS

Participants

Participants were part of the STAR*D Study—to date, the largest clinical trial for the pharmacological treatment of major depression. The original sample consisted of 4041 depressed adult outpatients who had undergone rigorous screening procedures before enrollment. These participants were followed from baseline through treatment Level 4, with a range of clinical and research assessments administered at regularly scheduled intervals. The study protocol including recruitment, screening, enrollment, methods, and measures has been described extensively in other publications (Rush *et al*, 2004; Trivedi *et al*, 2006; Lesser *et al*, 2007). Participants were enrolled without regard to race, but they described themselves as being Asian, Black/African American, Native American, Pacific Islander, White, multi-/biracial or 'other.' They were not asked to provide information about parental race or ancestry. In our analyses, we divided the sample into three racial categories, based on participants' self-reports: Black/African American, White, and 'other,' which collapsed the remaining groups into one solely for statistical purposes.

Genotyping and Quality-Control Procedures

A subset of 1953 participants provided DNA samples for genotyping and further analyses. The characteristics of this subsample and how they differed from the full sample have been detailed in earlier reports (McMahon *et al*, 2006; Laje *et al*, 2007; Perlis, 2007; Laje *et al*, 2009; Garriock *et al*, 2010). The genotyping and quality-control methods are extensively described elsewhere (Garriock *et al*, 2010; Cutler *et al*, 2012). Briefly, lymphocyte- or lymphoblastoid-derived DNA was genotyped on Affymetrix 500 K or 5.0 human SNP arrays ($\geqslant 99\%$ concordance was observed in the 12 samples genotyped on both platforms). During quality-control procedures, a total of 76 cases were omitted from our analyses because of missing genotypes, duplicates, or sex discrepancies. The omitted cases did not differ significantly from the remaining sample by self-reported race or other predictors used in our study. After pruning markers with minor allele frequency $<5\%$, Hardy–Weinberg deviation $p < 1 \times 10^{-05}$, and missing genotypes $>2\%$, our final sample of 1877 cases contained a total of 328 845 SNPs for analysis.

Measures

Predictors

Genetic ancestry. The study sample consisted of 299 self-reported Black/African Americans, 1464 Whites, and 114 'other.' MDS was used to extract genetic ancestral clusters using PLINK v. 1.07 (http://pngu.mgh.harvard.edu/purcell/plink/) (Purcell *et al*, 2007). Individual scores along the first four dimensions were plotted according to self-reported race. Using Eigenstrat (Price *et al*, 2006), principal component analysis (PCA) on 53 439 SNPs was performed by integrating the study sample data with the HapMap Phase III reference panels that included CEU (Utah residents of Northern/Western European ancestry, CEPH), TSI (Toscans in Italy), ASW (African ancestry in

Southwest USA), YRI (Yoruba in Ibadan, Nigeria), LWK (Luhya in Webuye, Kenya), and MKK (Masai in Kinyawa, Kenya) (Altshuler *et al*, 2010). For the PCA, all non-autosomal SNPs were removed.

To facilitate interpretation of the results and validate the MDS analysis, we ran STRUCTURE (version2.3.4) (Pritchard *et al*, 2000), which derives estimates of relative proportions of different population ancestries for each individual. After LD pruning (pair-wise $r^2 < 0.01$), 10 000 autosomal SNPS were randomly selected for analysis. Two sets of estimates were obtained for number of populations (k) equal to 3 and 4.

Non-genetic predictors. We included a subset of demographic, socioeconomic, baseline clinical, and psychosocial factors previously found to affect the treatment outcome and differ significantly across racial groups in the original study sample (Trivedi *et al*, 2006; Lesser *et al*, 2007) (see Supplementary Table S1). In addition to variables used for our analyses, we also examined adverse effects and drug tolerability in our study sample and found no significant racial group differences in adverse effects and tolerability, similar to earlier analyses in the full STAR*D sample.

Outcome. Our primary outcome was reduction in depressive symptoms based on longitudinal scores on the QIDS, clinician and self-rated versions (QIDS-C, QIDS-SR). We computed the reduction in depressive symptoms from baseline to exit, calculated as the difference between the baseline score and exit score (or last score recorded), expressed as a percentage of baseline. The Hamilton Rating Scale of Depression (17-item) ($HRSD_{17}$) was not included among outcome measures in all our models because over 15% of the sample had missing exit scores for this measure.

Statistical Analyses

Analyses were done primarily in SPSS/AMOS version 19, and structural equation models were validated via analyses done in SAS version 9.2 and MPLUS version 6. Baseline group differences were assessed with one-way ANOVA and χ^2 tests. To facilitate model specification, we obtained the bivariate Pearson's correlations among all predictor and outcome variables. We used maximum likelihood structural equation models to investigate the effects of African genetic ancestry (C1 from the MDS analysis), self-reported race, and other predictors on the reduction of depressive symptoms. Advantages of this statistical method include the ability to model predictors that are highly intercorrelated, specify the underlying data structure with latent variables, correct for error and bias inherent in individual variables, and compare various models on a selection of fit indices (Tomarken and Waller, 2005).

We derived latent constructs where possible from observed predictor and outcome variables through factor analysis, employing two widely used methods—principal axis factoring (PAF) and PCA. Although similar, the two methods use somewhat different approaches in extracting communalities among variables. PCA derives components by using the sums of squares across standardized variables summing to 1 (diagonal on correlation matrix R). PAF computes communality by estimating the proportion of variance in one variable that is shared among other variables. Thus, PAF can identify underlying constructs, whereas PCA serves mainly as a data reduction technique (Costello and Osborne, 2005). We compared PCA varimax rotation (with Kaiser normalization) with PAF oblique rotation (with Kaiser normalization). We obtained similar results for both approaches in terms of the number of components extracted and the variables that loaded best on those components (See Supplementary Tables S5 and S6).

Finally, we used the estimates of genetic ancestry proportions (which corresponded with the African (YRI) ancestral component) obtained in STRUCTURE and correlated them with the outcome measures separately (QIDS-C, QIDS-SR, and $HRSD_{17}$) using bivariate and multiple linear regression. The standardized coefficients were compared with those from the structural models that used genetic ancestry from MDS analysis, and latent treatment outcome measures. The unstandardized coefficients were used to facilitate interpretation for clinical applicability of findings by estimating the degree to which proportions of genetic ancestry were associated with changes in treatment outcome.

Structural Equation Models

Model 1. This model examined the effects of non-genetic predictors. This model is similar to the traditional approach in the literature (eg, Lesser *et al*, 2007). *Treatment Response*, the outcome, was modeled as a latent construct based on change scores from the QIDS-C and QIDS-SR, two highly correlated measures of symptomatic improvement. A latent predictor, *SES*, was based on education, occupational status, and insurance (private or public). A second latent predictor, *psychosocial functioning*, was based on scores on the quality of life enjoyment and satisfaction questionnaire, the work and social adjustment scale, and perceived mental functioning as measured by the Short Form 12-item health survey.

Model 2. This model differs from model 1 by adding a term for genetic ancestry based on the C1 score in the MDS analysis discussed above. Model 2 examined the effects of all the latent and observed predictors on treatment outcome and is thus the most comprehensive model.

Model 3. This model omitted variables found to be nonsignificant predictors of treatment in model 1. The 'dropout' variable was also omitted because it did not appreciably influence the effect of race and genetic ancestry on treatment response in models 1 and 2.

Model 4. This model only included the variables of interest, ie, genetic ancestry, self-reported race, and anxiety, and allowed us to determine whether anxiety mediated the relationship between ancestry or race and the outcome. Thus, this is the most parsimonious model. We further specified an alternate version of this model, where the HRSD17 was included with the QIDS-C/SR as part of the latent outcome, *treatment response.*

Goodness-of-fit. Goodness-of-fit indices are used to evaluate how well the data fit underlying theoretical relation-

ships specified among the variables. Typically, the range of indices evaluates different aspects of the model and also provides a cohesive guideline for optimal specification. For more extensive description and evaluation of fit indices used for structural equation models, see Tomarken and Waller (2005) and Hooper *et al* (2008).

In this paper, we used two absolute fit indices. The model χ^2 tests the hypothesis that the specified model is not significantly different from a 'perfectly fitting' model. A smaller χ^2 is more likely to be nonsignificant, indicating better fit. Sample sizes >200 make it more likely that the null is rejected, causing the χ^2 to be highly sensitive to sample size. The root mean square error of approximation (RMSEA) evaluates how well the covariance structure of the specified model matches that of the population of interest. An ideal RMSEA value ranges from 0 to 0.07, with values >0.08 considered 'poorer' fits. The RMSEA favors parsimony and gives better fit to models with fewer parameters.

In addition to the absolute fit indices, we used three relative fit indices. The comparative fit index (CFI) evaluates how well the specified model compares with the null model in which all variables are uncorrelated with each other. Values on the CFI range from 0 to 1, with values closer to 1 being better. The general cutoff for a good fitting model is 0.90, with an ideal fit $\geqslant 0.95$. The Tucker–Lewis index (TLI) is similar to the CFI and less sensitive to sample size. Values usually range from 0 to 1 but can exceed 1 in rare cases. Like the CFI, the TLI cutoff for a good fitting model is closer to 1; ideal fit $\geqslant 0.95$. The Akaike information criterion is a parsimony fit index that rewards models with fewer parameters. This minimizes the temptation to specify

as many relationships as possible in order to obtain saturation and 'better' fit on some of the more traditional fit criteria. Smaller Akaike information criterion values are indicative of better fit and more parsimony.

RESULTS

Table 1 reports the differences in observed predictor and outcome variables, broken down by self-reported race. Compared with the White group, the Black group had higher socioeconomic disadvantage and greater medical and depressive symptom severity. However, the Black group reported better perceived mental health at entry than the White group.

Relationship between Self-Reported Race and Genetic Ancestry

Figure 1a shows the corresponding genetic ancestry clusters C1 and C2 from the MDS analysis, differentiating individuals of African ancestry (C1) from those of European ancestry (C2) in the sample. Using the scores on each cluster, cases were classified into predicted genetic ancestral groups. The genetic ancestry cluster coincident with sub-Saharan Africans contained 97% of self-reported Blacks, 0.11% of self-reported Whites, and 15% of 'other' (see Supplementary Table S2 for genetic ancestry classification among individuals who reported being Black, White, or 'other' race).

Figure 1b shows the sample integrated with the HapMap reference panels. As expected, the self-reported

Table 1 Racial Group Differences on Relevant Predictor and Outcome Variables

| Measures	Racial group category		
	Black[a] (n = 299)	White (n = 1464)	Other (n = 114)
Age at enrollment, Mean (SD)	44.08 (12.59)	42.89 (13.55)	37.27 (12.10)***
Female (%)	64	61	62
Married (%)	31	45***	36
Years of schooling, mean (SD)	12.89 (2.55)	13.70 (3.36)***	14.06 (3.29)**
Employed (%)	46	58***	53
Private insurance (%)	37	57***	43**
Cumulative Illness Severity index, mean (SD)	1.41 (0.51)	1.26 (0.58)***	1.16 (0.53)***
Anxiety score, mean (SD)	6.70 (2.55)	6.10 (2.53)***	5.79 (2.47)**
Psychiatric comorbidity (%)	64	52***	61
Quality of life enjoyment and satisfaction questionnaire (QLESQ), mean (SD)	40.66 (14.49)	42.36 (15.22)	43.62 (13.22)
Work and social adjustment scale (WSAS)[b], mean (SD)	24.65 (9.79)	23.08 (9.06)*	22.75 (8.73)
Perceived mental functioning (SF12), mean (SD)	28.63 (9.48)	26.69 (8.55)***	26.97 (8.94)
Length of time (days) in study, mean (SD)	73.06 (30.76)	75.46 (26.16)	76.39 (26.50)
QIDS-C change[c], median	38	56***	53*
QIDS-SR change, median	33	50***	50
HRSD17 change, mean (SD)	35 (0.42)	43 (0.43)*	42 (0.44)

Ethnic group differences tested with one-way ANOVA, or χ^2-test, two-tailed significance; ***$p \leqslant 0.001$, **$p \leqslant 0.01$, and *$p \leqslant 0.05$.
Bonferroni correction for multiple pair-wise testing.
[a]'Black' is reference group.
[b]Higher scores indicate lower perceived adjustment.
[c]Average percent (%) improvement from baseline to exit on QIDS-C (median), QIDS-SR (median), and HRSD17 (mean).

a

b
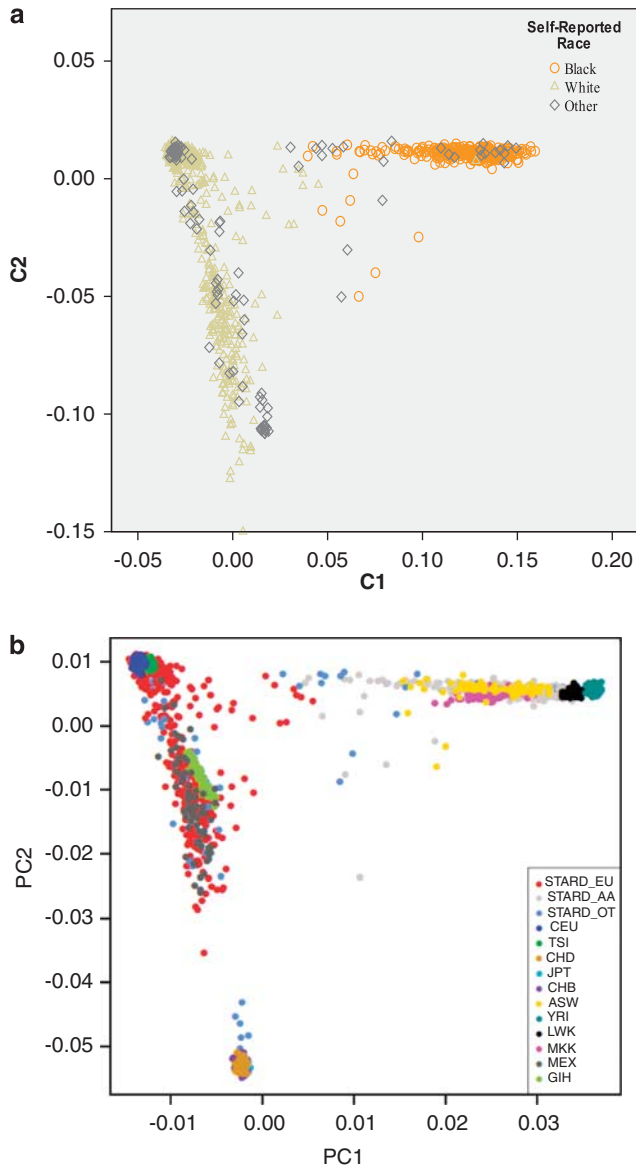
Figure 1 (a) Genetic ancestry cluster dimensions among the self-reported racial groups. X-axis shows scores along C1 dimension depicting African ancestry, whereas Y-axis shows scores along C2 dimension depicting European ancestry. The other clusters (C3 and C4) obtained from MDS analyses failed to differentiate along genetic ancestry dimensions and therefore are not depicted in the figure. (b) STAR*D genetic ancestry clusters with HapMap phase III reference panels. ASW, African ancestry in Southwest USA; CEU, Utah residents with northern and western European ancestry (CEPH); CHD, Chinese in metropolitan Denver, Colorado; GIH, Gujarati Indians in Houston, Texas; JPT, Japanese in Tokyo; LWK, Luhya in Webuye, Kenya; MEX, Mexican ancestry in Los Angeles, California; MKK, Masai in Kinyawa, Kenya; STARD_EU/STARD_AA/STARD_OT, self-reported White/Black or African American/other race in STAR*D sample; TSI, Tuscans in Italy: CHB, Han Chinese in Beijing: and YRI, Yoruba in Ibadan, Nigeria.

Black/African Americans in our sample clustered largely past the midpoint of PCA1 with the four African HapMap samples (ASW, YRI, LWK, and MKK), whereas the Whites clustered largely past the midpoint of PCA2 with the two European HapMap samples (CEU and TSI). The 'other' group showed varying degrees of African, European, Asian, and Native American ancestry, clustering along PCA1 and PCA2.

The diagrams corresponding to models 1 and 3 are shown in the Supplementary Figures S1 and S2. The diagrams corresponding to models 2 and 4 are depicted below in Figures 2 and 3, respectively. The direct and indirect effects of self-reported race and genetic ancestry for all models are summarized in Supplementary Table S7.

Model 1–what are the effects of self-reported race on therapeutic response when clinical and social factors are taken into account? Model 1 (see Supplementary Figure S1) shows the effects of self-reported race and other variables on treatment response. Race had a small but statistically significant direct effect on treatment response ($\beta = 0.050$, $p < 0.05$) that was independent of clinical and social factors, or early study withdrawal, consistent with previous reports (Lesser *et al*, 2007). The indirect effect of race, mediated by other clinical and social factors in the model, was also small but significant ($\beta = 0.043$, $p \leqslant 0.05$). Thus, the total effect of self-reported race on treatment response was 0.093.

Model 2—what happens to the effects of self-reported race and other factors when genetic ancestry is added to the model? Figure 2 shows the independent and direct effects of both genetic ancestry and self-reported race on treatment response. The direct effects of the other clinical and social factors on response are also indicated. The strongest direct predictors of response were length of time spent in the study (V17) SES, and psychosocial functioning, which were all significant at $p < 0.001$. Anxiety was also significantly predictive of treatment response. Gender (V3), marital status (V4), psychiatric comorbidity (V10), suicide attempt history (V11), and general medical conditions (V8) were not significant predictors of therapeutic response.

Genetic ancestry had a significant direct effect on treatment response ($\beta = -0.07$, $p < 0.05$ in this model) independent of other factors. In contrast to model 1, in model 2, self-reported race had no direct effect on treatment response. For a summary of the direct and indirect effects of genetic ancestry and self-reported race, see Supplementary Table S7.

Model 3. The relative effects of genetic ancestry, self-reported race, and anxiety were similar in model 2 and in the reduced model 3. This model also did not substantially improve fit to the data compared with models 1 and 2.

Model 4—impact of genetic ancestry, self-reported race, and anxiety on therapeutic outcome. This parsimonious model also allowed us to estimate whether anxiety was a significant mediator of the observed effects of genetic ancestry and self-reported race on treatment outcome.

As shown in Figure 3, genetic ancestry had a significant direct effect on treatment outcome under model 4, whereas self-reported race did not. Interestingly, although genetic ancestry had a significant direct effect on anxiety, which in turn had a significant direct effect on treatment response, the indirect effect of genetic ancestry mediated through anxiety was not significant.
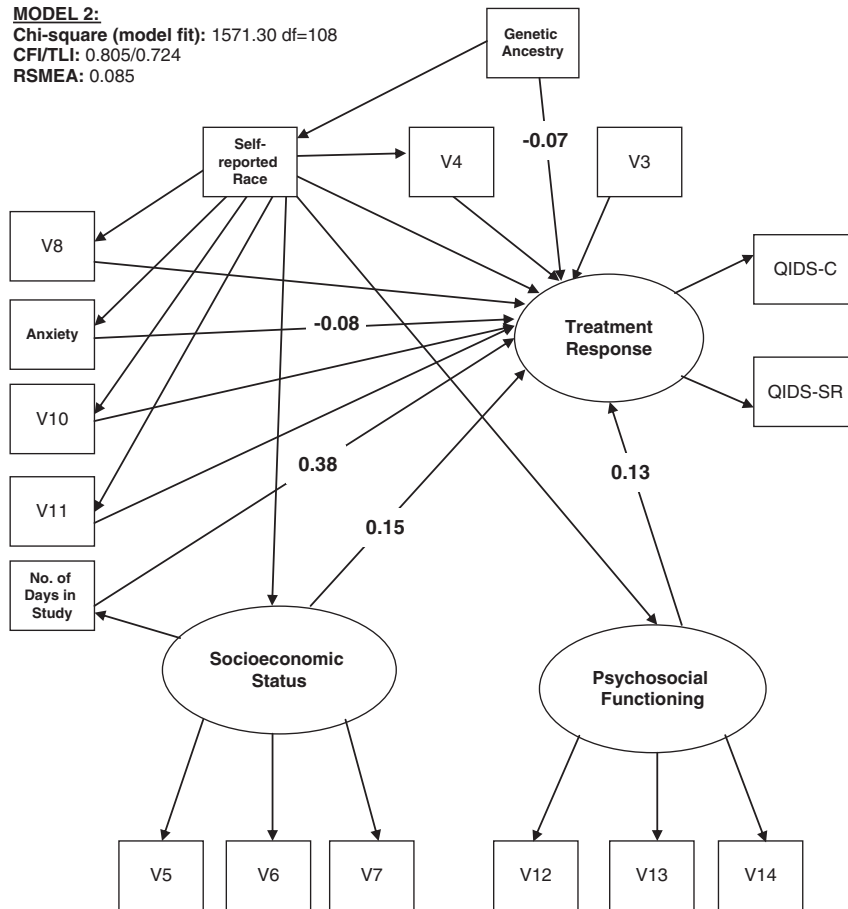
Figure 2 Effect of genetic ancestry, self-reported race, and clinical and social factors on treatment response. The diagram refers to model 2 as described in the text. Only coefficients for significant direct paths (at the $p < 0.05$ level (two-tailed)) to treatment response are depicted. For all models, regression coefficients use standardized units, ie, the effect of one SD change in predictors on SD change in treatment response, and are comparable across predictors.

A variation of model 4 (not shown but available upon request) showed that anxiety did not mediate the relationship between race and therapeutic response. When genetic ancestry and race were examined in the same model for their effects on anxiety (not shown in Figure), genetic ancestry continued to have a significant effect on anxiety, whereas race did not. Another variation of model 4 included $HRSD_{17}$ in the *treatment response* latent outcome; the results were similar to those in main model 4 (see Figure 3 for estimates). We also found that gender did not significantly moderate any of these relationships (data not shown but available upon request).

How well do the models fit the data? For a summary of the fit index values for all the models tested, see Supplementary Table S8. In models 1 and 2, the CFI and TLI model fit indices ranged from 0.670 to 0.805 and RSMEA exceeded 0.080, indicating a relatively poor fit to the data. Reduced model 3 showed a slight but less than optimal improvement over the first two models, whereas model 4 (and variations) showed substantial improvement in fit on all indices. Model 4, with the fewest parameters of the models tested, reduced the likelihood of misspecification among variables and did better on the parsimony fit
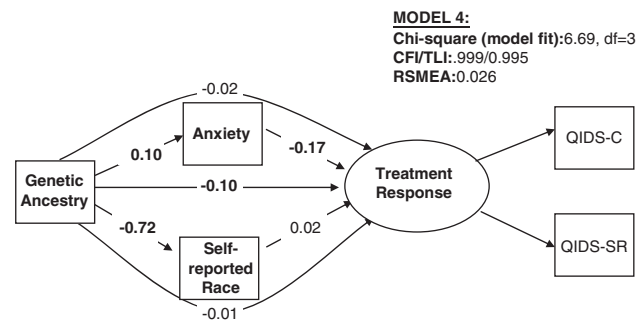
Figure 3 Effects of genetic ancestry, self-reported race, and anxiety on treatment response. The diagram refers to model 4 as described in the text. Direct effects are represented by straight arrows and indirect (mediator) effects are represented by curved arrows. For all models, coefficients (rounded to the nearest hundredth) represent standardized units, ie, the effect of one SD change in predictors on SD change in treatment response, and are comparable within models across predictors. All coefficients in bold are significant at $p < 0.05$ level (two-tailed). Genetic ancestry: total effects $= -0.129$***, direct effects $= -0.100$**, and indirect effects $= -0.029$); self-reported race: total effects $= 0.017$, direct effects $= 0.017$, and indirect effects $=$ N/A; anxiety: total effects $= -0.174$***, direct effects $= -0.174$, and indirect effects $=$ N/A. When HRSD17 is included as part of treatment response, the impact of genetic ancestry on TR is -0.108 ($p < 0.01$); the impact of self-reported race is 0.014 ($p > 0.500$); and the impact of anxiety on TR is -0.156 ($p < 0.001$).

criteria. Despite the varying degrees of fit across the four models, the relationships among the primary predictors and outcome of interest (genetic ancestry, self-reported race, and treatment response) did not change substantially.

How well do analyses from STRUCTURE estimates validate analyses from MDS genetic ancestry estimates? The standardized coefficients (β) from the MDS genetic ancestry estimates (Supplementary Table S7) and the STRUCTURE estimates (Supplementary Table S9) are approximately in agreement. Supplementary Table S9b shows the relative genetic ancestry proportions according to self-reported race, when the number of populations is set to 3 or 4. The adjusted regression coefficient for the MDS estimates in the best fitting structural model is -0.10, whereas those for the STRUCTURE estimates range from -0.08 to -0.13. As further depicted in Supplementary Table S9, the coefficients (β) for the STRUCTURE estimates show an 8–13% decrease in QIDS change scores for each percentage of African genetic ancestry. Accordingly, the 'average' self-reported African American with a median $\sim 88.6\%$ African (YRI) genetic ancestry would be expected to show a 7.1–11.5 point decrease in QIDS change scores after all social and baseline clinical factors are taken into account.

DISCUSSION

This is the first attempt to disentangle the effects of race and genetic ancestry on antidepressant treatment outcome. Although socioeconomic and baseline clinical factors were the most important reasons for racial differences in antidepressant response in Level 1 treatment, our results show an additional contribution of genetic ancestry to the residual disparity. This effect of genetic ancestry, while modest, could be of clinical significance when considered together with the other factors.

We found that self-reported race had virtually no direct contribution to the therapeutic response when genetic ancestry was taken into account. Three plausible mechanisms, not directly assessed in this study, could explain this: direct or indirect biological effects (eg, genes, epigenetics, endophenotypes, and other biomarkers); race-related differences in adherence to medication due to side effects, tolerability, convenience, etc; and race-related differences in therapeutic alliance (Murphy et al, 2013). We would expect the first mechanism to be explained by genetic ancestry and not race because race is a social construct. The second mechanism may be explained by both race and ancestry or by race alone, whereas the third mechanism is likely to be explained only by race, and not genetic ancestry. Because we also included race in all our models that had genetic ancestry as a predictor, the variance in treatment response that relates to race-associated adherence and therapeutic alliance with clinicians would be better accounted for by self-reported race, as well as dropout (time in study) and other psychosocial variables, but not genetic ancestry. Thus, it is unlikely that treatment adherence and therapeutic alliance are exerting their effects via genetic ancestry rather than by self-reported race or other psychosocial and environmental variables. Genetic ancestry eliminated the direct effects of race on outcome, but had no impact on the indirect effects of race—that is, the effects of race that were mediated by sociodemographic, clinical, or psychosocial factors.

Without the genetic ancestry data, our results are very consistent with published findings from the larger STAR*D study concerning racial differences in response to citalopram (Lesser et al, 2007). Two smaller, subsequently published studies showed no racial group differences in treatment outcome (Lesser et al, 2010; Lesser et al, 2011), although one of the studies reported a high dropout rate for African Americans (Lesser et al, 2011) relative to the other racial and ethnic groups. However, our ability to draw broad conclusions from these recent studies is limited, because both studies differ from STAR*D in terms of sample sizes, methodology, and outcome measures.

The significant effect of genetic ancestry that we detected may point to underlying differences in allele frequencies of common variants that are known to differentiate continental groups. Some of these alleles may affect the absorption, distribution, metabolism, or excretion of antidepressant agents, leading to differences in treatment outcomes. Other alleles may have specific effects on antidepressant targets. An earlier study in the STAR*D sample identified a variant in the gene HTR2A that was associated with better treatment response and was more than six times less prevalent in African Americans than in European Americans (McMahon et al, 2006). However, the functional effect of this variant in African Americans still remains to be clarified. Subsequent association studies of HTR2A have been mixed (Kang et al, 2007; Illi et al, 2009; Kishi et al, 2009; Lucae et al, 2010), but a PET study (Laje et al, 2010) showed that the same HTR2A allele that was associated with better treatment response was also associated with lower brain expression of the serotonin transporter—the proximal target of SSRIs—suggesting a mechanism whereby genetic variation in HRT2A could influence SSRI response. A genome-wide association study of treatment outcome in African Americans would be the next logical step, but much larger samples would be required to detect small effects on the genome-wide level.

We found a significant association between anxiety and self-reported race that diminished when genetic ancestry was included in the model. In other words, genetic ancestry accounted for some of the racial differences in anxiety. Earlier studies do suggest that African Americans with depression report more anxiety relative to their White counterparts (Simon et al, 1973; Uhlenhuth and Paykel 1973; Fabrega et al, 1988; Lesser et al, 2007; Lesser et al, 2011), but these studies do not shed light on the mechanism driving the racial difference in clinical presentation. More research is needed to explain this key clinical difference that has importance for antidepressant treatment outcomes.

While suggesting directions for future research, we should also point out some important limitations of the current study. Our findings are based on secondary data analyses, and hence we were limited by the existing study measures. As noted above, these did not include biological phenotypes like drug plasma levels that are generally more accurate than responses to questionnaires in assessing adherence. Second, this study was unable to address possible biological traits that may be closely linked with genetic ancestry, but which have sociological implications that could potentially diminish or eliminate the effect of genetic ancestry. This idea was illustrated in a study of genetic ancestry, skin color, social

class, and hypertension in a sample of mixed-ancestry Puerto Ricans (Gravlee et al, 2009).

A related limitation is the fact that this sample comprised relatively little genetic variation within racial groups. The sample mainly included Whites, with African Americans comprising the next largest group. Together, these two groups comprised over 90% of the total sample. As noted in the results, self-reported race had the lowest sensitivity for correct ancestral classification among the participants who did not report to be of African-American/Black or White race. Thus, an ideal sample would be more racially mixed with higher variance in African ancestry, thereby decreasing the likelihood of confounding race with ancestry and yielding considerably more information than self-reported race. This is an important sampling consideration for future studies.

CONCLUSION

This study demonstrates a significant effect of genetic ancestry on antidepressant treatment outcome. Although socioeconomic and baseline clinical factors are clearly the most important reasons for racial differences in antidepressant response, our results suggest that a significant portion of the residual disparity is explained by genetic ancestry, not race. Rarer forms of genetic variation that are highly prevalent among persons of African ancestry may also have an important role; if specific genes are involved, larger samples would be needed to identify them. These findings underscore the importance of including more African American patients in drug development and efficacy trials.

Disorders (www.nimhgenetics.org). The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Study is registered at www.clinicaltrials.gov with identifier number NCT00021528.

DISCLAIMER

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 US Government.

REFERENCES

Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F et al (2010). Integrating common and rare genetic variation in diverse human populations. Nature 467: 52–58.

Baldwin DS, Lopes ATV (2009). The influence of comorbid anxiety disorders on outcome in major depressive disorder. Medicographia 31: 126–130.

Brown C, Schulberg HC, Madonia MJ, Shear MK, Houck PR (1996). Treatment outcomes for primary care patients with major depression and lifetime anxiety disorders. Am J Psychiatry 153: 1293–1300.

Costello AB, Osborne JW (2005). Best practices in exploratory factor analysis: recommendations for getting the most from your analysis. Practical Assessment Res Eval 10: 1–9.

Cultler JA, Rush AJ, McMahon FJ, Laje G (2012). Common genetic variation in the indoleamine-2, 3-dioxigenase genes and antidpressant treatment outcome in major depressive disorder. J Psychopharmacol 26: 360–367.

Fabrega H, Mezzich J, Ulrich RF (1988). Black-white differences in psychopathology in an urban psychiatric population. Compr Psychiatry 29: 285–297.

Garriock H, Kraft J, Shyn SI, Peters EJ, Yokoyama JS, Jenkins GD et al (2010). A genomewide association study of citalopram response in major depressive disorder. Biol Psychiatry 67: 133–138.

Gravlee CC, Non AL, Mulligan CJ (2009). Genetic ancestry, social classification, and racial inequalities in blood pressure in Southeastern Puerto Rico. PloS ONE 4: 1–9.

Henry BW, Overall JE, Markette JR (1971). Comparison of major drug therapies for alleviation of anxiety and depression. Dis Nerv Syst 32: 655–667.

Hooper D, Coughlan J, Mullen MR (2008). Structural equation modelling: guidelines for determining model fit. Electronic J Business Res Methods 6: 53–60.

Illi A, Setala-Soikkeli E, Viikki M, Poutanen O, Huhtala H, Mononen N et al (2009). 5-HTR1A, 5-HTR2A, 5-HTR6A, TPH1 and TPH2 polymorphisms and major depression. Neuroreport 20: 1125–1128.

Kang RH, Choi MJ, Paik JW, Hahn SW, Lee MS (2007). Effect of serotonin receptor 2A gene polymorphism on mirtazapine response in major depression. Int J Psychiatry Med 37: 315–329.

Keers R, Aitchison KJ (2011). Pharmacogenetics of antidepressant response. Expert Rev Neurother 11: 101–125.

Kishi T, Yoshimura R, Kitajima T, Okochi T, Okumura T, Tsunoka T et al (2009). HT2RA is associated with SSRI response in major depressive disorder in a Japanese cohort. Neuromolecular Med 12: 237–242.

Laje G, Cannon DM, Allen AS, Klaver JM, Peck SA, Liu X et al (2010). Genetic variation in HTR2A influences serotonin transporter binding potential as measured using PET and [11C]DASB. Int J Neuropsychopharmacol 13: 715–724.

Laje G, Paddock S, Manji H, Rush AJ, Wilson AF, Charney D et al (2007). Genetic markers of suicidal ideation emerging during

citalopram treatment of major depression. *Am J Psychiatry* **164**: 1530–1538.

Laje G, Perlis RH, Rush AJ, McMahon FJ (2009). Pharmacogenetics studies in STAR*D: strengths, limitations, and results. *Psychiatr Serv* **60**: 1446–1457.

Lee SJ (2009). Defining statistical race and phenotypic race and their implications for health disparities. *Current Pharmacogenomics Personalized Med* **7**: 238–242.

Lesser IM, Castro DB, Gaynes BN, Gonzalez J, Rush AJ, Alpert JE et al (2007). Ethnicity/Race and outcome in the treatment of depression. Results From STAR*D. *Med Care* **45**: 1043–1051.

Lesser IM, Myers HF, Lin KM, Bingham Mira C, Joseph NT, Olmos NT et al (2010). Ethnic differences in antidepressant response: a prospective multi-site clinical trial. *Depress Anxiety* **27**: 56–62.

Lesser IM, Zisook S, Gaynes BN, Wisniewski SR, Luther JF, Fava M et al (2011). Effects of race and ethnicity on depression treatment outcomes: the COMED trial. *Psychiatr Serv* **62**: 1167–1179.

Lotrich FE, Pollock BG, Ferrell RE (2003). Serotonin transporter promoter polymorphism in African Americans: allele frequencies and implications for treatment. *Am J Pharmacogenomics* **3**: 145–147.

Lucae S, Ising M, Horstmann S, Baune BT, Arolt V, Müller-Myhsok B et al (2010). HT2RA gene variation is involved in antidepressant response. *Eur Neuropsychopharmacol* **20**: 65–68.

Luo HR, Poland RE, Lin KM, Wan YJ (2006). Genetic polymorphism of cytochrome P450 2C19 in Mexican Americans: a cross-ethnic comparative study. *Clin Pharmacol Ther* **80**: 33–40.

McMahon FJ, Buervenich S, Charney D, Lipsky R, Rush AJ, Wilson AF et al (2006). Variation in the gene encoding the serotonin 2A receptor as associated with outcome of antidepressant treatment. *Am J Hum Genet* **78**: 804–814.

Murphy EJ, Kassem L, Chemerinski A, Rush AJ, Laje G, McMahon FJ (2013). Retention and attrition among African Americans in the STAR*D Study: what causes research volunteers to stay or stray? *Depression and Anxiety* (in press).

Perlis RH (2007). Pharmacogenetic studies of antidepressant response: how far from the clinic? *Psychiatr Clin N Am* **30**: 125–138.

Price ALP, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**: 904–909.

Pritchard JK, Stephens M, Donelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D et al (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet* **81**: 559–575.

Raskin A, Crook TH (1975). Antidepressants in Black and White inpatients. Differential response to a controlled trial of chlorpomazine and imipramine. *Arch Gen Psychiatry* **32**: 643–649.

Ruhe HG, Ooteman W, Booij J, Michel MC, Moeton M, Baas F et al (2009). Serotonin transporter gene promoter polymorphisms modify the association between paroxetine serotonin transporter occupancy and clinical response in major depressive disorder. *Pharmacogenet Genomics* **19**: 67–76.

Rush AJ, Fava M, Wisniewski SR, Lavori PW, Trivedi MH, Sackeim HA et al (2004). Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Control Clin Trials* **25**: 119–142.

Simon J, Fleiss JE, Gurland BJ, Stiller PR, Sharpe L (1973). Depression and schizophrenia in hospitalized black and white mental patients. *Arch Gen Psychiatry* **28**: 509–512.

Strickland TL, Ranganath V, Lin KM, Poland RE, Mendoza R, Smith MW (1991). Pharmacologic considerations in the treatment of black American populations. *Psychopharmacol Bull* **27**: 441–448.

Strickland TL, Stein R, Lin KM, Risby E, Fong R (1997). The pharmacologic treatment of depression and anxiety in African Americans. Considerations for the general practioner. *Arch Fam Med* **6**: 371–375.

Sucheston LE, Bensen JT, Xu Z, Singh PK, Preus L, Mohler JL et al (2012). Genetic ancestry, self-reported race and ethnicity in African Americans and European Americans in the PCaP Cohort. *PloS ONE* **7**: e30950.

Tomarken AJ, Waller NG (2005). Structural equation modeling: Strengths, limitations and misconceptions. *Annu Rev Clin Psychol* **1**: 31–65.

Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L et al (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. *Am J Psychiatry* **163**: 28–40.

Uhlenhuth E, Paykel ES (1973). Symptom configuration and life events. *Arch Gen Psychiatry* **28**: 747–748.

Warden D, Rush AJ, Wisniewski SR, Lesser IM, Thase ME, Balasubramani GK et al (2009). Income and attrition in the treatment of depression: a STAR*D report. *Depression Anxiety* **26**: 622–633.

Warden D, Trivedi MH, Wisniewski SR, Davis L, Nierenberg AA, Gaynes BN et al (2007). Predictors of attrition during initial treatment for depression: A Star*D report. *Am J Psychiatry* **164**: 1189–1197.

Wassel CL, Jacobs DR, Duprez DA, Bluemke DA, Sibley CT, Criqui MH et al (2011). Association of self-reported race/ethnicity and genetic ancestry with arterial elasticity: The Multi-Ethnic Study of Atherosclerosis (MESA). *J Am Soc Hypertens* **5**: 463–472.

Yaeger R, Avila-Bront A, Abdul K, Nolan PC, Grann VR, Birchette MG et al (2008). Comparing genetic ancestry and self-described race in African Americans born in the United States and in Africa. *Cancer Epidemiol Biomarkers Prev* **17**: 1329–1338.

Supplementary Information accompanies the paper on the Neuropsychopharmacology website (http://www.nature.com/npp)