Applying novel methods for assessing individual- and neighborhood-level social and psychosocial environment interactions with genetic factors in the prediction of depressive symptoms in the Multi-Ethnic Study of Atherosclerosis

Erin B. Ware[1,2], **Jennifer A. Smith**[2], **Bhramar Mukherjee**[3], **Seunggeun Lee**[3], **Sharon L. R. Kardia**[2], and **Ana V. Diez-Roux**[4]

[1]Institute for Social Research, University of Michigan, Ann Arbor, MI

[2]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI

[3]Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI

[4]Department of Epidemiology and Biostatistics, School of Public Health, Drexel University, Philadelphia, PA

Abstract

Background—Complex illnesses, like depression, are thought to arise from the interplay between psychosocial stressors and genetic predispositions. Approaches that take into account both personal and neighborhood factors and that consider gene regions as well as individual SNPs may be necessary to capture these interactions across race and ethnic groups.

Methods—We used novel gene-region based analysis methods (Sequence Kernel Association Test (SKAT) and meta-analysis (MetaSKAT), Gene-Environment Set Association Test (GESAT)), as well as traditional linear models to identify gene region and SNP × psychosocial factor interactions at the individual- and neighborhood-level, across multiple race/ethnicities.

Results—Multiple regions identified in SKAT analyses showed evidence of a significant gene-region association with averaged depressive symptom scores across race/ethnicity (MetaSKAT p-values < 0.001). One region × neighborhood-environment interaction was significantly associated with averaged depressive symptom score across race/ethnicity after multiple testing correction (chr 18:21454070-21494070, Fisher's combined p-value = 0.001).

Corresponding author/Requests for reprints: Erin Ware, PhD #4614 SPH Tower 1415 Washington Heights Ann Arbor, MI 48109
Phone: 734.615.9725 Fax: 734.764.1357 ebakshis@umich.edu.

Online Resources:

Online Resource 1: Extended Methods

Online Resource 2: Region information for the 47 regions identified from the individual-SNP based GWAS of averaged depressive symptom score in the Multi-Ethnic Study of Atherosclerosis(Ware *et al.* 2015)

Online Resource 3: Sequence Kernel Association Testing and Meta-analysis results for the 47 regions identified from the individual-SNP based GWAS of averaged depressive symptom score in the Multi-Ethnic Study of Atherosclerosis(Ware *et al.* 2015)

Online Resource 4: Fisher's combined p-values and race/ethnicity-specific p-values for each SNP in region 46

Conclusions—The examination of gene regions jointly with environmental factors measured at multiple levels (individuals and their contexts) may shed light on the etiology of depressive illness across race/ethnicities.

Keywords

Gene × environment; depressive symptoms; GESAT; gene-set testing; gene environment set association testing

Introduction

Complex diseases, such as depressive illness, likely arise from a combination of genetic, individual, and environmental factors that interact to affect susceptibility or resilience. It is likely that both individual- and neighborhood-level characteristics play a role in modifying genetic associations with depressive symptoms. Further, it is plausible that the interactions between genetic factors and these individual- and neighborhood-level environmental factors may impact the development of depressive illness.

It has been postulated that neighborhood contextual characteristics may be related to mental health over and above the effects of individual characteristics. Neighborhood constructs investigated have included general measures of neighborhood socioeconomic position (SEP), as well as more specific measures of chronic stressors (including violence, disorder, and aesthetic quality) and measures of neighborhood social cohesion/social support. Although not all studies have been consistent, a large number of studies have documented associations of various features of neighborhood environments with depression or depressive symptoms (reviewed in (Mair *et al.* 2008a; Mair *et al.* 2008b)).

Despite evidence that neighborhood environments play a role in the etiology of depressive illness, few studies have investigated how these neighborhood factors may interact with genetic predictors of depression or depressive symptoms. For example, most work on gene-by-environment interactions in depression has focused on the serotonin transporter gene (*5-HTT*) promoter variant, and has been limited to either personal characteristics or broader county level factors (rather than neighborhoods which are likely more homogenous in environmental characteristics than counties). Studies of gene-by-environment (G × E) interactions involving *5-HTT* have found interactions with a personal history of stressful life events, and with the proportion of individuals receiving public assistance (county-level), county-level infant mortality rates, and county-level crime rates, suggesting that that both individual (e.g. stressful life events) and broader environmental (e.g. county-level crime rates) factors may convey different risks of depression for individuals with different genotypes (Caspi *et al.* 2003; Kendler *et al.* 2005; Middeldorp *et al.* 2008; Koenen *et al.* 2009; Uddin *et al.* 2010).

There is no strong consensus in the field as to whether these interactions have been sufficiently replicated. A meta-analysis of fourteen studies investigating *5-HTT* × stressful life events interactions found no evidence of interaction effect between genotype and stressful life events on depression (Risch *et al.* 2009). It has been suggested that limited environmental measurement may have at least in part contributed to lack of replication

(Koenen & Galea 2009; Lotrich & Lenze 2009; Rieckmann *et al.* 2009; Schwahn & Grabe 2009). Publication bias may also have contributed substantially to these results as G × E interaction studies are traditionally underpowered and often have an increased false discovery rate (Duncan & Keller 2011). In addition, replication studies often struggle to find similar studies for homogeneous comparison, which limits their generalizability. For example, the vast majority of these studies have been performed only in European subsamples which do not allow for comparison across genetically distinct ethnic groups limiting the applicability of study findings beyond European populations.

Another major limitation of prior work on G × E interactions in depressive illness has been the type and quality of the social environment measures available. The ability to replicate findings has been shown to differ when measures of environment are defined objectively versus subjectively. In the context of the serotonin transporter gene, a review found that studies involving objective measures assessing environment adversity replicated G × E interactions either fully or in part, whereas studies relying on self-reported measures often did not replicate (Uher & McGuffin 2010). Depression has been connected to several biological pathways including metabolic pathways (e.g. (Marazziti *et al.* 2013)), inflammatory pathways (e.g. (Felger & Lotrich 2013)), and neurobiological pathways (e.g. (Krishnan & Nestler 2008)). These pathways may be activated by external environments such as chronic burden or neighborhood stressors (e.g. (Pariante & Lightman 2008)). There is need to further investigate G × E interactions using larger and diverse samples, novel assessment of interactions using gene region approaches, and improved environmental measures at multiple levels.

Recent work on the genetics of depression has focused on alternative phenotypic measures (such as measures of depressive symptoms rather than assessment of depressive illness which increases power) and has broadened the genetic factors examined to genome-wide analyses (moving beyond the candidate gene approaches describe above). In a recent GWAS of depressive symptoms using the Center for Epidemiologic Studies depressions Scale (CES-D) in European Ancestry individuals, no loci reached genome-wide significance in the discovery sample (composed of 34,549 individuals), but one SNP reached genome-wide significance (p-value = 4.78×10^{-8}) in overall meta-analysis of the combined discovery and replication samples (n = 51,258) (Hek *et al.* 2013). Another GWAS performed with the goal of assessing longitudinal depressive phenotypes across four ethnicities (African -, European-, Chinese-, and Hispanic-Americans) found several novel variants at the genome-wide suggestive level ($5 \times 10^{-8} <$ p-value 5×10^{-6}) in each race/ethnicity for each approach to analyzing longitudinal depressive symptoms (Ware *et al.* 2015).

While examining single nucleotide polymorphisms (SNPs) through GWAS is an important first step in identifying genetic risk factors for depressive symptoms, SNP-set based analyses may help us better understand the association between genetic variants and complex phenotypes by identifying genetic regions that are associated with the phenotype across different ethnicities (Mukherjee *et al.* 2013). SNP-set testing jointly analyzes SNPs in a defined region, overcoming some of the power limitations of candidate gene research and individual-SNP GWAS by reducing the number of tests being performed. Additionally, because relevant variability in a given genetic region may be indexed by different SNPs in

different ethnicities, the failure to perform gene-region analyses may result in underestimates of the effects of genetic variability on phenotypic variability. For example, there have been a number of conditions (e.g., bipolar disorder, coronary artery disease, hypertension, Parkinson's, amyotrophic lateral sclerosis, Crohn's disease, rheumatoid arthritis, types I and II diabetes, and age-related eye disease) for which analyses of genetic regions identified important genetic predictors whereas traditional SNP analyses did not (Peng *et al.* 2010).

Using quantitative depressive symptom scores measured with the CES-D, we investigated interactions between variability in selected genetic regions and both individual-level and neighborhood-level measures of social environments to elucidate cross-race/ethnicity G × E interactions, paying particular attention to the contributions of each race/ethnic group to the cross-race/ethnicity analyses. The genetic regions for investigation were identified based on prior GWAS (Ware *et al.* 2015) and gene region analysis. Three environmental factors were investigated: neighborhood-level social factors (an index score combining neighborhood social cohesion, perceived neighborhood safety, and neighborhood aesthetic quality, which have previously been found to be significantly associated with depressive symptoms in the MESA cohort (Mair *et al.* 2009)), and individual-level measures of social support and chronic burden from stress.

Methods

MESA

This study uses data from the Multi-Ethnic Study of Atherosclerosis, which is described extensively elsewhere (Bild *et al.* 2002). Briefly, MESA is a longitudinal study supported by NHLBI consisting of individuals from six field centers. All MESA cohort members who provided DNA samples and were included in the SNP Health Association Resource (SHARe) project are included in this analysis. Informed consent was obtained from all individual participants included in this study. Analyses were performed in African Americans (AA), European Americans (EA), Chinese Americans (CA), and Hispanic Americans (HA). Participants were ascertained from six study sites (Baltimore, MD; Chicago, IL; Forsyth County, NC; Los Angeles, CA; New York, NY; and St. Paul, MN).

The outcome of interest, depressive symptom score, was assessed using the 20-item CES-D Scale (Radloff 1977), at MESA exams 1, 3 and 4 over a total of approximately ten years. CES-D scores were adjusted for anti-depressant use and averaged over all exams for which the measure was administered (Hek *et al.* 2013; Ware *et al.* 2015). Adult socioeconomic position (ASEP) is included as a covariate in the region-level analyses. Depressive symptom score, adjustment for anti-depressant use, and ASEP are further discussed in Online Resource 1.

Approximately one million SNPs were genotyped for MESA participants using the Affymetrix Genome-Wide Human SNP Array 6.0. The IMPUTE 2.1.0 program was used in conjunction with HapMap Phase I and II reference panels (CEU+YRI+CHB+JPT, release 22 - NCBI Build 36 for African-, Chinese- and Hispanic-American participants; CEU, release 24 - NCBI Build 36 for European Americans) to increase the number of available SNPs to

approximately 2.5 million markers. We accounted for population substructure by including the top four race/ethnicity-specific principal components (estimated from genome-wide data) as adjustment covariates in all analyses, as proposed previously (Setiawan *et al.* 2009; Sun *et al.* 2009). SNP dosages were used in the SKAT and GESAT models.

Two individual-level social environments are used in these analyses: chronic burden (CB) and emotional social support (SS). These measures represent different dimensions of individual stressors and were analyzed in separate models. CB was measured at two exams in MESA (exams 1 and 3) and is based off of the chronic burden scale developed for the Healthy Women Study (Bromberger & Matthews 1996). Higher values of CB indicate higher chronic burden (i.e. more burdens). Emotional social support was measured at exams one and three of MESA and is based on a scale from the Enhancing Recovery in Coronary Heart Disease study (Enrichd Investigators 2001). Higher scores indicate more social support. Information on questions included and index coding are available in Online Resource 1.

Neighborhood social environment is summarized into a neighborhood index score (NIS) composed of three dimensions: aesthetic quality (AQ), safety (SF), and social cohesion (SC) measured with a 1-mile radius as the definition of neighborhood. The separate neighborhood dimension scales (AQ, SF, SC) pooled information from MESA respondents and from a separate Community Survey (CS) in order to obtain a more reliable measure of the neighborhood context that is not based exclusively on the MESA participant's report (Mujahid *et al.* 2007). The NIS was created by averaging the responses for all MESA and CS participants living within 1 mile of the MESA participant. We averaged 1-mile means for AQ, SF, and SC across the three exams by neighborhood dimension and then averaged the three exam-specific averages. The final NIS score was standardized to the combined race/ethnicity mean and standard deviation. Higher NIS indicates "more positive" overall neighborhood environments, such as a high degree of SF, good AQ, and/or good SC.

Overview of statistical analysis

To investigate $G \times E$, we first conducted a GWAS of longitudinal depressive symptoms, averaged over time. GWAS was performed separately by race/ethnicity (Ware *et al.* 2015), then meta-analyzed across race/ethnicity. Meta-analysis p-values were used to select regions for $G \times E$ investigation. Next, we performed region-based analysis to investigate genetic main effects on averaged depressive symptoms, in each ethnicity separately, followed by region-based $G \times E$ (GESAT), separately for each race/ethnicity. Region-based $G \times E$ were meta-analyzed across ethnicity using Fisher's method (Fisher 1925). We then followed up significant $G \times E$ with linear regression models in order to estimate the direction and magnitude of effect of each individual SNPs.

Region identification

We previously conducted race/ethnicity-specific GWAS in MESA's four ethnic groups, using depressive symptom score averaged across all exams as the outcome measure (Ware *et al.* 2015). Linear modeling, adjusting for age at baseline, sex, site at baseline, and the top four race/ethnicity-specific principal components was used to assess the relationship

between SNPs and depressive symptoms. SNPs were analyzed as dosages using an additive genetic model.

Gene regions for further analysis were selected by first ranking the 5,000 SNPs with the lowest p-values from the averaged depressive symptom GWAS (filtered at race/ethnicity-specific minor allele frequency (MAF) > 0.05) within each race/ethnicity. Once those SNPs were identified, the union of the SNPs (n_{SNP} = 19,932) was obtained, and each SNP was analyzed in a fixed-effects meta-analysis across the four ethnicities using METAL, weighted by sample size (Willer *et al.* 2010). From the meta-analysis results, SNPs were retained if they had MAF > 0.05 in more than one race/ethnicity (n_{SNP} = 18,645). SNPs were ranked by lowest p-value (meta-analysis p-value: $P_{(1)}$ $P_{(2)}$ ... $P_{(nSNP)}$) and the top 100 meta-analysis SNPs ($P_{(1)} - P_{(100)}$) were identified.

Starting with the SNP with the lowest meta-analysis p-value ($P_{(1)}$ – referred to as the index SNP), a region was defined including all SNPs (not only SNPs in the meta-analysis) within a 20 kilobase (kb) region up and downstream of the index SNP (eliminating any SNPs in the meta-analysis top 100 in this region from being an index SNP of a second region). The 40kb total region was selected to conservatively capture the average size of a linkage disequilibrium (LD) block. In EA and AA populations, average LD, calculated by r^2, declines to approximately 0.15 - 0.25 at a distance of 40kb (Shifman *et al.* 2003). We continued this process until all regions were identified from the top 100 SNPs from meta-analysis ($n_{regions}$ = 47). It is possible that regions overlap slightly if the index SNP of $region_i$ and the index SNP of $region_j$ (i   j) are more than 20kb but less than 40kb away from each other. This occurred five times (regions 3 and 4; 5 and 6; 7 and 8; 27, 28, and 29; and 41 and 42]).

Region analyses

Regions were selected for G × E analysis based on statistical tests for marginal genetic association performed through Sequence Kernel Association Testing (SKAT) and meta-analysis (MetaSKAT) of those results (Wu *et al.* 2011; Lee *et al.* 2013; Lee 2014; Lee *et al.* 2014). For the 47 identified regions, SKAT was performed for each region for each race/ethnicity separately (adjusted for age, sex, site, top four race/ethnicity-specific principal components (PC), and ASEP), and MetaSKAT was conducted across all four ethnicities. SKAT and MetaSKAT methods are described extensively in Online Resource 1, region information is presented in Online Resource 2, and results from the SKAT and MetaSKAT analysis can be found in Online Resource 3. Any region that had a MetaSKAT p-value less than 0.20 was included in the interaction analyses. The threshold of a MetaSKAT p-value of 0.20 was selected to allow for the possibility of qualitative interaction which could result in an insignificant marginal effect for the region (for example if the genetic variant is positively associated with CES-D in the presence of the environmental factor but inversely associated in the absence of the factor resulting in a marginally significant or null marginal effect).

Gene-Environment Set Association Test (GESAT)

The GESAT is a gene or region based test for G × E interaction and was performed to elucidate interaction effects in the context of depressive symptoms (Lin *et al.* 2013). Suppose *n* subjects are genotyped in a region with *p* SNPs. For GESAT, the interaction model is:

$$E\left(\bar{y}_{i\cdot}\right) = \alpha_0 + \boldsymbol{\alpha_1}' \boldsymbol{X}_i + \boldsymbol{\alpha_2}' \boldsymbol{G}_i + \alpha_3 E_i + \boldsymbol{\beta}' \boldsymbol{S_i},$$

where $\bar{y}_{i\cdot}$ is the log-transformed (depressive symptom score averaged across exams plus 1) for individual *i*, $\mathbf{X}_i$ is a vector of non-genetic covariates, E_i is the environmental factor, $\mathbf{G}_i$ is a vector of genetic markers, $\mathbf{S}_i$ is a vector of G × E interaction terms, and $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)$ is a vector of regression coefficients for the interaction terms. It assumes that each of the β_j, *j* = 1, ..., *p*, independently follows an arbitrary distribution with mean zero and common variance τ^2. Testing H_0: $\tau^2 = 0$ is equivalent to testing H_0: $\boldsymbol{\beta} = 0$, which tests whether at least one of the interaction terms is non zero. Covariates include age, sex, site, top four race/ethnicity-specific PCs, and adult socioeconomic position (ASEP). The variance-component score statistic for τ is

$$Q_{GESAT} = \left(\bar{\boldsymbol{y}} - \hat{\boldsymbol{\mu}}_{main}\right)' \boldsymbol{S} \boldsymbol{S}' \left(\bar{\boldsymbol{y}} - \hat{\boldsymbol{\mu}}_{main}\right),$$

where $\boldsymbol{S} = \left(S_1', \ldots, S_n'\right)'$ is an $n \times p$ matrix of G × E interactions in the region, $\hat{\boldsymbol{\mu}}_{main}$ is the estimated mean of $\bar{\boldsymbol{y}}$ under the main effect only model of no interactions, i.e. H_0: $\tau^2 = 0$.

P-values for each region from G × E models were combined across ethnicities using Fisher's method (Fisher 1925):

$$\chi_{2K}^2 \sim -2 \sum_{k=1}^{K} ln\left(p_k\right),$$

where p_k is the p-value for the G × E interaction for each race/ethnicity *k*, and *K* is the number of ethnicities. This statistic follows a chi-square distribution with 2*K* degrees of freedom. Only cases where two or more ethnicities contribute to the statistic were included.

Individual SNP-level × Environment Linear Regression Models

Gene-level interaction analyses were followed by individual SNP × E analyses using linear regression to estimate both the magnitude and direction of each SNP × E interaction term for regions that showed significant evidence of interactions in the analysis using GESAT. The following model was used:

$$E\left(\bar{y}_{i\cdot}\right) = \alpha_0 + \boldsymbol{\alpha_1}' \boldsymbol{X}_i + \beta_{SNP} SNP_i + \beta_E E_i + \gamma SNP_i * E_i,$$

where $\bar{y}_{i\cdot}$ is the depressive symptom measure averaged across exams and log transformed for participant i, $\mathbf{X}_i$ is a vector of non-genetic covariates, E_i is the environmental factor, SNP_i is a SNP genotype, $SNP_i * E_i$ is a G × E interaction term, and γ is a regression coefficient of the interaction term. This is a cross-sectional analysis using the subject-level average as outcome. To better control for Type I error rate for G × E analyses, we used the generalized estimating equation (GEE) with a robust sandwich estimator of variance, instead of the model-based estimator of variance, as recommended in Voorman, et al. 2011 (Voorman *et al.* 2011). Fisher's method (Fisher 1925) was again used to combine results across ethnicity for the SNP × E interaction p-values. This method gives an estimate of the overall effect of the SNP × E interaction across ethnicities.

RESULTS

Demographics from the MESA sample, by race/ethnicity, are presented in Table I. There were 2,514 EA, 1,603 AA, 1,443 HA, and 775 CA participants included in these analyses. Averaged CES-D score ranged from 6.2 (standard deviation (SD) 10.4) in the CA sample to 10.2 (SD 8.5) in the AA sample. Average age was 62 years, with a majority of female participants in each race/ethnicity. Chronic burden, social support, and neighborhood index score were all significantly associated (p-value < 0.05) with averaged depressive symptom score in bivariate models for AA, EA and CA, and adjusted (for age, sex, site, and adult socioeconomic position) models for AA and EA. After GWAS, region selection, SKAT analyses, and MetaSKAT analyses, a total of 37 regions (of 47) had MetaSKAT p-values <0.20 and were included in GESAT interaction analyses. Of the regions included in the GESAT analyses, four gene regions (on chromosome (chr) 6 and chr 8) were significantly associated with depressive symptoms across ethnicities after multiple testing correction (p-value range: 7.5×10^{-4} to 9.0×10^{-5}, $\alpha_{Bonferroni} = 0.001$) (Online Resource 3).

GESAT tests for gene region × environment

Chronic Burden and Social Support—Results from all regions for the region × CB interactions are shown in Table II. Of the 37 regions investigated, one region showed marginally significant (p-value < 0.05) evidence of interaction (region 13, chr 2: 192565881-192605881, Fisher's combined p-value=0.03) with CB across race/ethnicity, though this region did not pass the Bonferonni correction ($\alpha_{Bonferroni}=0.001$). Two of the 37 regions had Fisher's combined p-values for interactions with SS less than 0.10 across race/ethnicity, though due to multiple testing, these regions may be considered marginal at best.

Neighborhood Index Score—Two regions were found to have significant ($\alpha_{Bonferroni} =0.001$) or marginally significant (p-value < 0.05) interactions with NIS (Table III). The interaction between region 46 × NIS was significantly associated with depressive symptoms across race/ethnicity (chr 18:21454070-21494070, Fisher's combined p-value = 0.001). Within race/ethnicity analyses showed an NIS interaction with region 46 as marginally significant in three ethnicities (AA p-value = 0.04, EA p-value = 0.03, HA p-value = 3.94×10^{-3}). Region 46 does not have any established genes within ±100kb of the index SNP. Region 38 had a marginal interaction with NIS across ethnicity (chr 13:33802221-33842221, Fisher's combined p-value=0.03), though this result was primarily

driven by the Hispanic ethnicity (HA p-value=5.91 ×10^{-3}). Region 38 does not have any established genes within ±100kb of the index SNP.

Individual SNP-level × Environment Linear Regression Models—Instances where there was a significant cross-race/ethnicity region-level interaction ($\alpha_{\text{Bonferroni}}$ =0.001) for any genomic region × E were considered to provide the strongest evidence of cross-race/ethnicity region-level interaction effects. Only one region × E (region 46 × NIS) reached significance and individual SNP × E interactions were examined to determine which SNPs were driving the region-level interaction.

Neighborhood Index Score—Region 46 had 24 SNP × E interactions that reached significance. The lowest p-values in this region were for rs4800653 (p-value 1.47×10^{-6}) and rs1840444 (p-value 7.65×10^{-6}) (Figure 1). The race/ethnicity-specific sample sizes, SNP × E p-values, and Fisher's chi-square, degrees of freedom, and p-value for cross-race/ethnicity comparison for each region can be found in Online Resource 4.

Discussion

These analyses used novel methods (GESAT (Lin *et al.* 2013)) to elucidate gene region × social environment interactions associated with averaged depressive symptoms, across race/ethnicity, using regions defined from a previous study (Ware *et al.* 2015). Incorporating three different environments, two at the individual-level (chronic burden and social support) and one at the neighborhood-level (neighborhood index score), we found two regions with marginally significant ($p < 0.05$) G × E interactions (one from chronic burden (region 13, chr 2, Fisher's combined p-value = 0.03), and one from the NIS analyses (region 38, chr 13, Fisher's combined p-value = 0.03)), and one gene region with a significant interaction with NIS (region 46, chr 18, Fisher's combined p-value = 0.001), across race/ethnicity. Investigating the significant gene region × NIS association at the SNP level and combining across race/ethnicity provided suggestive evidence of multiple SNP × E interactions with the neighborhood index score environmental measure.

Unlike a previously published GWAS/region analysis (Mukherjee *et al.* 2013), we did not select our top SNPs from race/ethnicity-specific GWAS. Rather, we selected the top SNPs from a meta-analysis across the four ethnicities, since our goal was to find regions associated with depressive symptom phenotype across multiple ethnicities. This is the reason we may not have seen our strongest signal in a region as that region's index SNP. It is also apparent that there may be effects in only a subset of the ethnicities (e.g., in EA and HA only, in CA, HA, and AA only, etc.) as opposed to across all four of the examined ethnicities. Future research should consider all combinations of ethnicities in gene region analysis to elucidate regions that are associated with phenotypes under study.

GESAT in particular – an extension of SKAT, which has been shown through simulation and real data applications to be a more powerful method over others (e.g. weighted sum statistics (Madsen & Browning 2009), cohort allelic sum tests (Morgenthaler & Thilly 2007), or C-alpha test (Neale *et al.* 2011)) – is computationally efficient, robust, and has several advantages over traditional SNP × environment analysis. GESAT allows for

covariate adjustment and can test common variants through the use of an unweighted linear kernel. Since our analyses filtered out any rare variants (MAF < 0.05), this option is particularly important. GESAT also does not assume that all variants will produce effects of similar direction and magnitude by allowing the variance of an individual variant to differ from a mean of zero. Finally, GESAT allows for a test of biologically meaningful regions rather than individual SNPs that may vary in distribution across ethnicities due to evolutionary patterns and may not be functional genetic variants (Lander & Schork 1994; Goddard *et al.* 2000; Bryc *et al.* 2010). Unfortunately, GESAT does not yet allow for testing of phenotypes over time in repeated measures models, accounting for correlation between measures on the same individual. The extension of GESAT to allow for repeated measures modeling would greatly enhance the ability to detect genetic effects for phenotypes that are characterized over time.

Region 46, the region with a significant interaction with NIS, showed evidence for regulatory function (several large elevations in enhancer- and promoter-associated histone marks (H3K4Me1) in the positions approximately +30kb and to a lesser extent −45kb from the index SNP, as well as many dense DNaseI hypersentivity clusters and transcription factor binding sites, and many SNPs that are in SINEs or LTRs). Typically, regulatory areas modulate gene expression in response to developmental, tissue specific or environmental signals. Influences on gene expression from developmental signals may lay down a basis for methylation across the life course and consequently lead to higher (or lower) depressive symptoms later in life. The regulation of tissue-specific signals could possibly set up the brain's ability to successfully (or unsuccessfully) adapt to chemical stimuli, while these regulatory regions may also influence how the body responds, at a molecular level, to neighborhood stimuli – leading to biological plausibility of this gene region × neighborhood interaction's influence on depressive symptoms. This bioinformatic evidence, taken from the ENCODE database, is indicative of a potential regulatory effect of the genetic region involved in G × NIS interactions related to depressive symptoms – suggesting that further functional characterization is warranted (Kent *et al.* 2002; Consortium *et al.* 2012).

Though regions were chosen using a fairly conservative genetic distance of ±40kb, regions of true association could be larger or smaller than our selected size. We detected regions that overlapped, which may imply larger regions ought to be created from these abutting regions. Future research should determine biologically relevant regions while still incorporating the SNP-based GWAS information, or possibly use other approaches (moving windows, LD block refinement, gene-regions, etc.) to elucidate genetic regions. We assumed an additive effect for each SNP; that is, for every additional copy of the coded allele, the mean response (averaged depressive symptom score) increases (or decreases) linearly. However, it is likely that the additive model may not be the best-fitting model for every variant within a region. Additional testing with different genetic effect assumptions is warranted to better estimate the true genetic effects of these variants on depressive symptoms.

Despite these limitations, this novel work in examining the impact of G × E interactions on depressive symptoms, across multiple gene regions, environment definitions, and race/ethnicities was possible through innovative gene-environment set association test techniques, and through detailed assessments of individual-level psychosocial environment

and objective neighborhood dimensions. These methods permit an examination of genes/regions across ethnicities, where individual SNPs may not replicate across ethnicities due to race/ethnicity specific patterns of linkage disequilibrium or differences in allele frequencies across ethnic groups (Lander & Schork 1994; Goddard *et al.* 2000; Bryc *et al.* 2010). Using these novel approaches, we found suggestive evidence that neighborhood context may interact with genetic factors in shaping depressive symptoms. Replication in other samples is necessary before firm conclusions can be drawn.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

REFERENCES

Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR Jr. Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M, Tracy RP. Multi-ethnic study of atherosclerosis: objectives and design. Am J Epidemiol. 2002; 156:871–81. [PubMed: 12397006]

Bromberger JT, Matthews KA. A longitudinal study of the effects of pessimism, trait anxiety, and life stress on depressive symptoms in middle-aged women. Psychol Aging. 1996; 11:207–13. [PubMed: 8795049]

Bryc K, Velez C, Karafet T, Moreno-Estrada A, Reynolds A, Auton A, Hammer M, Bustamante CD, Ostrer H. Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations. Proc Natl Acad Sci U S A. 2010; 107(Suppl 2):8954–61. [PubMed: 20445096]

Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington HL, McClay J, Mill J, Martin J, Braithwaite A. Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene. American Association for the Advancement of Science. 2003

Consortium EP, Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012; 489:57–74. [PubMed: 22955616]

Duncan LE, Keller MC. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. Am J Psychiatry. 2011; 168:1041–9. [PubMed: 21890791]

Enrichd Investigators. Enhancing Recovery in Coronary Heart Disease (ENRICHD) study intervention: rationale and design. Psychosom Med. 2001; 63:747–55. [PubMed: 11573023]

Felger JC, Lotrich FE. Inflammatory cytokines in depression: neurobiological mechanisms and therapeutic implications. Neuroscience. 2013; 246:199–229. [PubMed: 23644052]

Fisher, RA. Statistical Methods for Research Workers. Edinburgh; Oliver and Boyd: 1925.

Goddard KA, Hopkins PJ, Hall JM, Witte JS. Linkage disequilibrium and allele-frequency distributions for 114 single-nucleotide polymorphisms in five populations. Am J Hum Genet. 2000; 66:216–34. [PubMed: 10631153]

Hek K, Demirkan A, Lahti J, Terracciano A, Teumer A, Cornelis MC, Amin N, Bakshis E, Baumert J, Ding J, Liu Y, Marciante K, Meirelles O, Nalls MA, Sun YV, Vogelzangs N, Yu L, Bandinelli S,

Benjamin EJ, Bennett DA, Boomsma D, Cannas A, Coker LH, de Geus E, De Jager PL, Diez-Roux AV, Purcell , Hu FB, Rimm EB, Hunter DJ, Jensen MK, Curhan G, Rice K, Penman AD, Rotter JI, Sotoodehnia N, Emeny R, Eriksson JG, Evans DA, Ferrucci L, Fornage M, Gudnason V, Hofman A, Illig T, Kardia S, Kelly-Hayes M, Koenen K, Kraft P, Kuningas M, Massaro JM, Melzer D, Mulas A, Mulder CL, Murray A, Oostra BA, Palotie A, Penninx B, Petersmann A, Pilling LC, Psaty B, Rawal R, Reiman EM, Schulz A, Shulman JM, Singleton AB, Smith AV, Sutin AR, Uitterlinden AG, Volzke H, Widen E, Yaffe K, Zonderman AB, Cucca F, Harris T, Ladwig KH, Llewellyn DJ, Raikkonen K, Tanaka T, van Duijn CM, Grabe HJ, Launer LJ, Lunetta KL, Mosley TH Jr. Newman AB, Tiemeier H, Murabito J. A Genome-Wide Association Study of Depressive Symptoms. Biol Psychiatry. 2013

Kendler KS, Kuhn JW, Vittum J, Prescott CA, Riley B. The interaction of stressful life events and a serotonin transporter polymorphism in the prediction of episodes of major depression: a replication. Arch Gen Psychiatry. 2005; 62:529–35. [PubMed: 15867106]

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. Genome Res. 2002; 12:996–1006. [PubMed: 12045153]

Koenen KC, Aiello AE, Bakshis E, Amstadter AB, Ruggiero KJ, Acierno R, Kilpatrick DG, Gelernter J, Galea S. Modification of the association between serotonin transporter genotype and risk of posttraumatic stress disorder in adults by county-level social environment. Am J Epidemiol. 2009; 169:704–11. [PubMed: 19228812]

Koenen KC, Galea S. Gene-environment interactions and depression. JAMA. 2009; 302:1859. author reply 61-2. [PubMed: 19887662]

Krishnan V, Nestler EJ. The molecular neurobiology of depression. Nature. 2008; 455:894–902. [PubMed: 18923511]

Lander ES, Schork NJ. Genetic dissection of complex traits. Science. 1994; 265:2037–48. [PubMed: 8091226]

Lee S. MetaSKAT: Meta analysis for SNP-set (Sequence) Kernel Association Test. R package version 0.40. 2014

Lee S, Teslovich TM, Boehnke M, Lin X. General framework for meta-analysis of rare variants in sequencing association studies. Am J Hum Genet. 2013; 93:42–53. [PubMed: 23768515]

Lee S, Miropolsky L, Wu M. SKAT: SNP-set (Sequence) Kernel Association Test. R package version 1.0.1. 2014

Lin X, Lee S, Christiani DC, Lin X. Test for interactions between a genetic marker set and environment in generalized linear models. Biostatistics. 2013; 14:667–81. [PubMed: 23462021]

Lotrich FE, Lenze E. Gene-environment interactions and depression. JAMA. 2009; 302:1859–60. author reply 61-2. [PubMed: 19887661]

Madsen BE, Browning SR. A groupwise association test for rare mutations using a weighted sum statistic. PLoS Genet. 2009; 5:e1000384. [PubMed: 19214210]

Mair C, Diez Roux AV, Shen M, Shea S, Seeman T, Echeverria S, O'Meara ES. Cross-sectional and longitudinal associations of neighborhood cohesion and stressors with depressive symptoms in the multiethnic study of atherosclerosis. Ann Epidemiol. 2009; 19:49–57. [PubMed: 19064189]

Mair C, Hawes SE, Agne HD, Sow PS, N'Doye I, Manhart LE, Fu PL, Gottlieb GS, Kiviat NB. Factors associated with CD4 lymphocyte counts in HIV-negative Senegalese individuals. Clin Exp Immunol. 2008a; 151:432–40. [PubMed: 18190600]

Mair C, Roux AV, Galea S. Are neighbourhood characteristics associated with depressive symptoms? A review of evidence. British medical journal. 2008b; 62:940–6.

Marazziti D, Rutigliano G, Baroni S, Landi P, Dell'osso L. Metabolic syndrome and major depression. CNS Spectr. 2013:1–12.

Middeldorp CM, Cath DC, Beem AL, Willemsen G, Boomsma DI. Life events, anxious depression and personality: a prospective and genetic study. Psychological medicine. 2008; 38:1557–65. [PubMed: 18294422]

Morgenthaler S, Thilly WG. A strategy to discover genes that carry multi-allelic or mono allelic risk for common diseases: a cohort allelic sums test (CAST). Mutat Res. 2007; 615:28–56. [PubMed: 17101154]

Mujahid MS, Diez Roux AV, Morenoff JD, Raghunathan T. Assessing the measurement properties of neighborhood scales: from psychometrics to ecometrics. Am J Epidemiol. 2007; 165:858–67. [PubMed: 17329713]

Mukherjee S, Kim S, Ramanan VK, Gibbons LE, Nho K, Glymour MM, Ertekin-Taner N, Montine TJ, Saykin AJ, Crane PK, the Alzheimer's Disease Neuroimaging I. Gene-based GWAS and biological pathway analysis of the resilience of executive functioning. Brain Imaging Behav. 2013

Neale BM, Rivas MA, Voight BF, Altshuler D, Devlin B, Orho-Melander M, Kathiresan S, Purcell SM, Roeder K, Daly MJ. Testing for an unusual distribution of rare variants. PLoS Genet. 2011; 7:e1001322. [PubMed: 21408211]

Pariante CM, Lightman SL. The HPA axis in major depression: classical theories and new developments. Trends in Neurosciences. 2008; 31:464–8. [PubMed: 18675469]

Peng G, Luo L, Siu H, Zhu Y, Hu P, Hong S, Zhao J, Zhou X, Reveille JD, Jin L, Amos CI, Xiong M. Gene and pathway-based second-wave analysis of genome-wide association studies. Eur J Hum Genet. 2010; 18:111–7. [PubMed: 19584899]

Radloff L. The CES-D scale: A self-report depression scale for research in the general population. Applied psychological measurement. 1977; 1:385–401.

Rieckmann N, Rapp MA, Muller-Nordhorn J. Gene-environment interactions and depression. JAMA. 2009; 302:1861. author reply -2. [PubMed: 19887659]

Risch N, Herrell R, Lehner T, Liang KY, Eaves L, Hoh J, Griem A, Kovacs M, Ott J, Merikangas KR. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. JAMA : the journal of the American Medical Association. 2009; 301:2462–71. [PubMed: 19531786]

Schwahn C, Grabe HJ. Gene-environment interactions and depression. JAMA. 2009; 302:1860–1. author reply 1-2. [PubMed: 19887660]

Setiawan VW, Doherty JA, Shu XO, Akbari MR, Chen C, De Vivo I, Demichele A, Garcia-Closas M, Goodman MT, Haiman CA, Hankinson SE, Henderson BE, Horn-Ross PL, Lacey JV Jr. Le Marchand L, Levine DA, Liang X, Lissowska J, Lurie G, McGrath M, Narod SA, Rebbeck TR, Ursin G, Weiss NS, Xiang YB, Yang HP, Zheng W, Olson SH. Two estrogen-related variants in CYP19A1 and endometrial cancer risk: a pooled analysis in the Epidemiology of Endometrial Cancer Consortium. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2009; 18:242–7.

Shifman S, Kuypers J, Kokoris M, Yakir B, Darvasi A. Linkage disequilibrium patterns of the human genome across populations. Hum Mol Genet. 2003; 12:771–6. [PubMed: 12651872]

Sun YV, Peyser PA, Kardia SL. A common copy number variation on chromosome 6 association with the gene expression level of endothelin 1 in transformed B lymphocytes from three racial groups. Circulation. Cardiovascular genetics. 2009; 2:483–8. [PubMed: 20031624]

Uddin M, Koenen KC, de Los Santos R, Bakshis E, Aiello AE, Galea S. Gender differences in the genetic and environmental determinants of adolescent depression. Depression and anxiety. 2010

Uher R, McGuffin P. The moderation by the serotonin transporter gene of environmental adversity in the etiology of depression: 2009 update. Mol Psychiatry. 2010; 15:18–22. [PubMed: 20029411]

Voorman A, Lumley T, McKnight B, Rice K. Behavior of QQ-plots and genomic control in studies of gene-environment interaction. PLoS One. 2011; 6:e19416. [PubMed: 21589913]

Ware EB, Mukherjee B, Sun YV, Diez Roux AV, Kardia SL, Smith JA. Comparative genome-wide association studies of a depressive symptom phenotype in a repeated measures setting by race/ethnicity in the Multi-Ethnic Study of Atherosclerosis. BMC Genetics. 2015 Under Review.

Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26:2190–1. [PubMed: 20616382]

Wu MC, Lee S, Cai T, Li Y, Boehnke M, Lin X. Rare-variant association testing for sequencing data with the sequence kernel association test. Am J Hum Genet. 2011; 89:82–93. [PubMed: 21737059]
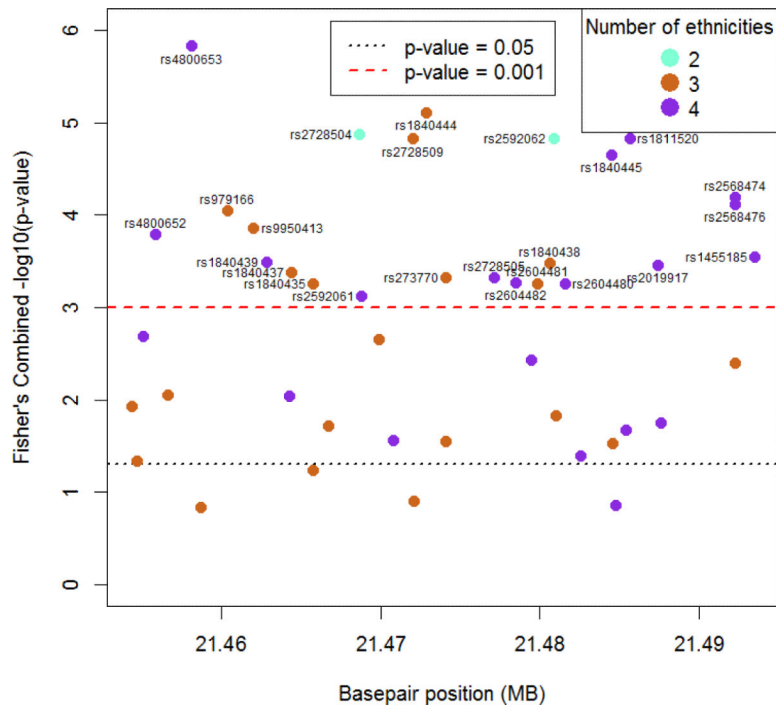
Figure 1.

Fisher's combined $-\log_{10}$(p-value) for each SNP $\times$ environment (neighborhood index score) interaction in region 46 plotted against genomic position

SNPs with Fisher's combined p-value < 0.001 are identified by rs number. Colors indicate the number of ethnicities that were used in calculating the Fisher's combined p-value. Only SNPs with race/ethnicity-specific minor allele frequency > 5% were included in the Fisher's combined analysis.

Table I

Descriptive statistics, Multi-Ethnic Study of Atherosclerosis

	MESA n = 6,335							
	European American n=2,514		African American n=1,603		Hispanic American n=1,443		Chinese American n=775	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Depressive symptom score[a]								
Baseline CES-D	8	(7.8)	7.6	(7.6)	9.9	(9.2)	6.3	(6.6)
Averaged CES-D	8.7	(7.4)	7.8	(6.7)	10.2	(8.5)	6.2	(5.6)
Age	62.6	(10.2)	62.2	(10.1)	61.4	(10.3)	62.4	(10.4)
Sex (N,% male)	1207	48.0	744	46.4	711	49.3	385	49.7
Site (N,%) Baltimore, MD	505	20.1	482	30.1	0	0.0	0	0.0
Chicago, IL	526	20.9	258	16.1	0	0.0	275	35.4
Forsyth County, NC	548	21.8	425	26.5	3	0.2	0	0.0
Los Angeles, CA	133	5.3	143	8.9	554	38.4	498	64.3
New York, NY	209	8.3	295	18.4	431	29.9	2	0.3
St. Paul, MN	593	23.6	0	0.0	455	31.5	0	0.0
Anti-depressant use (N,%)	307	12.2	61	3.8	84	5.8	19	2.5
Adult socioeconomic position[b]	6.5	2.3	5.0	2.8	3.0	2.5	4.2	3.0
Chronic burden[c]	0.9	1.0	0.2	1.1	0.1	1.0	−0.4	0.8
Social support[d]	5.0	4.9	5.2	4.8	5.3	5.1	4.8	4.3
Neighborhood index score[e]	0.4	0.9	−0.3	1.0	−0.6	1.0	0.2	0.7
	β	P	β	P	β	P	β	P
Chronic burden[c,f]	0.33	<0.001	0.31	<0.001	0.32	<0.001	0.40	<0.001
Chronic burden[c,g]	0.30	<0.001	0.27	<0.001	0.31	<0.001	0.34	<0.001
Social support[d,f]	−0.07	<0.001	−0.06	<0.001	−0.07	<0.001	−0.09	<0.001
Social support[d,g]	−0.06	<0.001	−0.06	<0.001	−0.06	<0.001	−0.08	<0.001
Neighborhood index score[e,f]	−0.12	<0.001	−0.14	<0.001	−0.03	0.30	0.10	0.03
Neighborhood index score[e,g]	−0.06	0.01	−0.06	0.04	0.04	0.21	−0.01	0.87

CES-D: Center for Epidemiologic Studies – Depression; MD: Maryland, IL: Illinois, NC: North Carolina, CA: California, NY: New York, MN: Minnesota

[a] CES-D measured as 20-item sum ranging from 0 to 60;

[b] Adult Socioeconomic Position: EA n=2,376, AA n=1,402, HA n=1,379, CA n=747;

[c] Chronic burden: EA n=2,374, AA n=1,397, HA n=1,375, CA n=747;

[d] Social support: EA n=2,376, AA n=1,402, HA n=1,378, CA n=747;

[e] Neighborhood index score was standardized to the combined race/ethnicity mean and standard deviation: EA n=1,930, AA n=1,100, HA n=976, CA n=538;

[f] Linear regression model, unadjusted;

[g]Linear regression model, adjusted for age, sex, study site, and adult socioeconomic position

Table II

Gene-Environment Set Associations Tests (GESAT) interaction results for chronic burden × region and social support × region

Region	Chronic Burden						Social Support					
	AA	EA	CA	HA	Fisher's		AA	EA	CA	HA	Fisher's	
	p	p	p	p	χ^2 (df)	χ^2 p	p	p	p	p	χ^2 (df)	χ^2 p
1	9.91E-01	6.93E-02	3.53E-02	6.96E-01	12.8 (8)	0.12	3.25E-01	9.65E-01	3.69E-01	7.14E-01	5.0 (8)	0.76
2	7.61E-01	3.18E-01	3.94E-01	7.82E-01	5.2 (8)	0.74	9.22E-01	8.95E-01	6.97E-01	6.86E-01	1.9 (8)	0.99
3	7.64E-01	4.10E-01	8.35E-01	6.98E-01	3.4 (8)	0.91	7.73E-01	5.50E-01	5.00E-01	8.70E-01	3.4 (8)	0.91
4	3.20E-01	6.93E-01	9.20E-01	9.51E-01	3.3 (8)	0.92	3.89E-01	7.61E-01	4.19E-01	9.88E-01	4.2 (8)	0.84
5	1.73E-01	5.58E-01	9.03E-01	9.66E-01	4.9 (8)	0.76	6.87E-01	7.63E-01	4.28E-01	9.67E-01	3.1 (8)	0.93
6	1.95E-01	6.10E-01	8.54E-01	9.50E-01	4.7 (8)	0.79	5.46E-01	7.82E-01	7.40E-01	9.60E-01	2.4 (8)	0.97
7	1.27E-01	3.58E-01	5.32E-01	6.80E-01	8.2 (8)	0.41	3.67E-01	3.73E-01	3.42E-01	5.34E-02	12.0 (8)	0.15
8	3.41E-01	6.46E-01	6.10E-01	5.72E-01	5.1 (8)	0.74	1.91E-01	5.73E-01	3.55E-01	2.83E-02	13.6 (8)	0.09
9	1.57E-01	1.33E-01	1.43E-01	5.71E-01	12.7 (8)	0.12	2.09E-01	2.51E-01	1.82E-01	6.78E-01	10.1 (8)	0.26
10	4.62E-01	6.41E-01	4.93E-01	5.95E-01	4.9 (8)	0.77	4.75E-01	9.23E-01	7.62E-01	1.74E-01	5.7 (8)	0.68
11	4.69E-01	3.28E-01	3.70E-01	2.40E-01	8.6 (8)	0.38	8.98E-01	8.61E-01	9.95E-01	9.86E-01	0.6 (8)	1.00
13	1.85E-01	3.21E-01	2.63E-01	1.27E-02	17.1 (8)	**0.03**	6.40E-02	4.06E-01	2.15E-01	9.78E-01	10.4 (8)	0.24
14	5.52E-01	3.45E-01	9.77E-02	8.59E-01	8.3 (8)	0.41	6.63E-01	6.88E-02	5.79E-01	9.08E-01	7.5 (8)	0.49
16	2.91E-01	4.50E-01	6.68E-01	3.88E-01	6.8 (8)	0.56	4.27E-01	3.69E-01	--	7.12E-01	4.4 (6)	0.63
17	3.62E-01	--	1.81E-01	5.45E-01	6.7 (6)	0.35	6.37E-01	4.53E-01	2.90E-01	7.30E-01	5.6 (8)	0.69
19	9.54E-01	7.82E-01	5.03E-01	2.86E-01	4.5 (8)	0.81	1.55E-01	4.13E-02	8.99E-01	9.91E-01	10.3 (8)	0.24
20	1.75E-01	3.10E-01	6.17E-01	6.15E-01	7.8 (8)	0.46	2.41E-01	2.89E-01	3.12E-01	7.64E-01	8.2 (8)	0.41
21	5.33E-01	8.92E-01	1.24E-01	8.89E-01	5.9 (8)	0.66	9.43E-01	8.37E-01	7.78E-01	8.00E-01	1.4 (8)	0.99
22	5.57E-01	6.42E-01	7.20E-01	1.87E-01	6.1 (8)	0.64	2.69E-01	4.34E-01	8.66E-01	2.73E-01	7.2 (8)	0.52
23	6.51E-01	9.66E-01	2.91E-01	4.12E-01	5.2 (8)	0.74	5.67E-01	4.28E-01	6.56E-01	9.19E-01	3.8 (8)	0.87
24	6.87E-01	4.45E-01	2.29E-01	6.67E-01	6.1 (8)	0.63	8.18E-01	7.81E-01	7.50E-01	5.25E-01	2.8 (8)	0.95
26	7.65E-01	7.75E-01	7.49E-01	4.42E-01	3.3 (8)	0.92	6.06E-01	5.26E-02	9.44E-01	2.49E-01	9.8 (8)	0.28
27	6.32E-01	8.78E-01	8.40E-01	4.30E-01	3.2 (8)	0.92	4.82E-01	8.63E-02	9.44E-01	2.50E-01	9.2 (8)	0.32
28	5.62E-01	7.60E-01	2.28E-01	4.98E-01	6.1 (8)	0.64	3.37E-01	1.72E-01	7.52E-01	4.59E-01	7.8 (8)	0.45
29	3.81E-01	7.51E-01	5.90E-01	4.25E-01	5.3 (8)	0.73	4.19E-02	4.20E-01	3.99E-01	9.08E-01	14.7 (8)	0.06

Region	Chronic Burden						Social Support					
	AA	EA	CA	HA	Fisher's		AA	EA	CA	HA	Fisher's	
	p	p	p	p	χ² (df)	χ² p	p	p	p	p	χ² (df)	χ² p
30	6.67E-01	5.84E-01	4.62E-01	2.40E-01	6.3 (8)	0.62	4.98E-01	9.20E-01	9.05E-01	9.62E-01	1.8 (8)	0.99
32	7.93E-02	6.63E-01	6.07E-01	9.51E-01	7.0 (8)	0.54	1.68E-01	8.85E-01	5.84E-01	9.72E-01	4.9 (8)	0.76
33	5.12E-01	6.52E-01	4.92E-01	1.74E-01	7.1 (8)	0.52	1.53E-01	7.13E-01	3.16E-01	2.56E-01	9.5 (8)	0.30
34	6.98E-01	1.26E-01	5.31E-01	2.90E-01	8.6 (8)	0.38	9.02E-01	2.18E-01	1.02E-01	5.54E-01	9.0 (8)	0.34
35	1.34E-01	9.77E-01	4.25E-01	7.62E-01	6.3 (8)	0.61	8.64E-01	9.12E-01	3.43E-01	7.91E-01	3.1 (8)	0.93
38	8.42E-01	7.33E-01	9.18E-02	5.40E-01	7.0 (8)	0.54	8.92E-01	9.65E-01	4.38E-01	2.57E-02	9.3 (8)	0.32
40	8.45E-01	4.14E-01	2.11E-01	2.10E-01	8.3 (8)	0.40	1.04E-01	9.94E-01	6.35E-01	6.47E-01	6.3 (8)	0.61
41	9.65E-01	5.94E-01	5.25E-01	2.28E-01	5.4 (8)	0.72	5.29E-01	9.08E-01	4.44E-01	7.26E-01	3.7 (8)	0.88
42	2.11E-01	4.60E-01	1.89E-01	3.83E-01	9.9 (8)	0.27	2.67E-01	6.96E-01	3.62E-01	7.62E-01	5.9 (8)	0.65
43	6.09E-01	5.35E-01	7.45E-01	1.24E-01	7.0 (8)	0.54	1.10E-01	5.01E-01	8.86E-01	8.45E-01	6.4 (8)	0.61
44	5.95E-01	6.46E-02	9.17E-02	6.42E-01	12.2 (8)	0.14	5.78E-01	8.73E-01	6.79E-01	9.88E-02	6.8 (8)	0.56
46	7.61E-01	7.12E-01	6.51E-01	9.15E-01	2.3 (8)	0.97	7.23E-01	4.94E-01	7.55E-01	7.84E-01	3.1 (8)	0.93

AA: African Americans, EA: European Americans, CA: Chinese Americans, HA: Hispanic Americans, p: p-value

Only regions with MetaSKAT p-value <0.20 were investigated in the interaction analysis.

-- Indicates a model that did not converge, Fisher's combined p-values 0.05 are bolded

Table III

Gene-Environment Set Associations Tests (GESAT) interaction results for Neighborhood Index Score and region, by race/ethnicity

	AA	EA	CA	HA	Fisher's	
Region	p	p	p	p	χ^2 (df)	χ^2 p
1	8.61E-01	4.91E-01	6.07E-01	1.06E-01	7.2 (8)	0.51
2	8.84E-01	5.71E-01	6.81E-01	8.88E-01	2.4 (8)	0.97
3	2.12E-01	8.40E-01	7.56E-01	1.08E-01	8.5 (8)	0.39
4	4.68E-01	6.74E-01	4.66E-01	4.69E-01	5.3 (8)	0.72
5	7.31E-01	3.84E-01	5.36E-01	5.82E-01	4.9 (8)	0.77
6	5.76E-01	3.61E-01	6.07E-01	5.51E-01	5.3 (8)	0.72
7	3.98E-01	6.08E-01	9.34E-01	9.61E-01	3.1 (8)	0.93
8	2.28E-01	4.58E-01	8.49E-01	9.18E-01	5.0 (8)	0.76
9	1.91E-01	1.49E-01	3.62E-01	2.57E-01	11.9 (8)	0.16
10	3.69E-01	2.58E-01	5.55E-01	7.98E-02	10.9 (8)	0.21
11	6.08E-02	9.63E-01	6.24E-01	5.37E-01	7.9 (8)	0.45
13	4.33E-01	7.36E-01	1.21E-01	4.52E-01	8.1 (8)	0.42
14	8.03E-01	9.52E-01	5.51E-01	3.41E-01	3.9 (8)	0.87
16	9.42E-01	2.40E-01	--	1.77E-01	6.4 (6)	0.38
17	3.99E-01	3.82E-01	1.46E-01	3.82E-01	9.5 (8)	0.30
19	6.44E-01	7.97E-01	7.83E-02	4.56E-01	8.0 (8)	0.43
20	6.18E-02	3.46E-01	1.02E-01	4.64E-01	13.8 (8)	0.09
21	6.77E-01	5.33E-01	3.22E-01	6.56E-01	5.1 (8)	0.74
22	1.62E-01	9.15E-01	2.06E-01	7.78E-01	7.5 (8)	0.49
23	9.59E-01	4.53E-01	1.98E-02	3.97E-01	11.4 (8)	0.18
24	3.71E-01	4.25E-01	9.58E-01	3.21E-01	6.1 (8)	0.64
26	9.18E-01	6.75E-01	3.96E-01	2.93E-01	5.3 (8)	0.73
27	8.36E-01	7.32E-01	5.28E-01	6.11E-02	7.9 (8)	0.45
28	6.80E-01	9.01E-01	2.29E-01	1.06E-01	8.4 (8)	0.39
29	2.80E-01	2.07E-01	6.14E-01	6.42E-02	12.2 (8)	0.14
30	6.50E-01	6.48E-01	8.94E-01	8.76E-01	2.2 (8)	0.97
32	3.55E-01	9.52E-02	4.57E-01	4.98E-01	9.7 (8)	0.28
33	9.02E-01	8.81E-01	3.43E-01	1.28E-01	6.7 (8)	0.57
34	1.63E-01	8.34E-01	6.96E-01	1.41E-02	13.2 (8)	0.10
35	3.62E-01	7.10E-01	2.39E-01	4.70E-01	7.1 (8)	0.53
38	3.22E-01	2.75E-01	4.65E-01	5.91E-03	16.6 (8)	**0.03**
40	8.77E-01	8.38E-02	5.98E-01	5.49E-01	7.4 (8)	0.49
41	5.16E-01	2.84E-01	3.37E-01	3.59E-01	8.1 (8)	0.43
42	3.15E-01	6.07E-01	4.55E-01	2.72E-01	7.5 (8)	0.49
43	7.45E-01	4.43E-01	5.79E-01	8.34E-01	3.7 (8)	0.89

	AA	EA	CA	HA	Fisher's	
Region	p	p	p	p	χ^2 (df)	χ^2 p
44	6.25E-01	2.05E-01	2.28E-01	8.49E-01	7.4 (8)	0.49
46	4.23E-02	2.57E-02	4.97E-01	3.94E-03	26.1 (8)	**0.001**

AA: African Americans, EA: European Americans, CA: Chinese Americans, HA: Hispanic Americans, p: p-value

Only regions with MetaSKAT p-value <0.20 were investigated in the interaction analysis.

-- Indicates a model that did not converge, Fisher's combined p-values ≤ 0.05 are bolded