Published in final edited form as:

Soc Forces. 2015 ; 93(3): 881–903. doi:10.1093/sf/sou086.

Gene by Social-Environment Interaction for Youth Delinquency and Violence: Thirty-Nine Aggression-related Genes

Hexuan Liu[1,3], **Yi Li**[1], and **Guang Guo**[1,2,3]

[1]Department of Sociology, the University of North Carolina at Chapel Hill

[2]Carolina Center for Genome Sciences, the University of North Carolina at Chapel Hill

[3]Carolina Population Center, the University of North Carolina at Chapel Hill

Abstract

Complex human traits are likely to be affected by many environmental and genetic factors, and the interactions among them. However, previous gene-environment interaction (G×E) studies have typically focused on one or only a few genetic variants at a time. To provide a broader view of G×E, this study examines the relationship between 403 genetic variants from 39 genes and youth delinquency and violence. We find evidence that low social control is associated with greater genetic risk for delinquency and violence and high/moderate social control with smaller genetic risk for delinquency and violence. Our findings are consistent with prior G×E studies based on a small number of genetic variants, and, more importantly, we show that these findings still hold when a large number of genetic variants are considered simultaneously. A key implication of these findings is that the expression of multiple genes related to delinquency depends on the social environment: gene expression is likely to be amplified in low-social-control environments but, tends to be suppressed in high/moderate-social-control environments. This study not only deepens our understanding of how the social environment shapes individual behavior, but also provides important conceptual and methodological insights for future G×E research on complex human traits.

INTRODUCTION

Previous studies have shown that gene-environment interplay contributes to a variety of behavioral and social outcomes (Boardman *et al.* 2012; Caspi *et al.* 2002; Fowler *et al.* 2011; Guo *et al.* 2008; Pescosolido *et al.* 2008; Shanahan *et al.* 2008; Simons *et al.* 2011). Yet these studies have typically focused on one or only a few genetic variants at a time. The aim of our research is to provide a more comprehensive view of the gene-environment interplay by incorporating dozens of genes identified in animal studies; particularly, to show how the social environment moderates genetic risk for youth delinquent and violent behaviors.

Corresponding author: Hexuan Liu, Department of Sociology, University of North Carolina, Hamilton Hall, CB#3210, Chapel Hill, North Carolina 27599-3210, hexuanl@live.unc.edu.

Traits determined by a single gene or allele are rare in human beings (Glazier *et al.* 2002). The vast majority of human diseases (e.g., cancer, heart disease, and diabetes) are complex traits affected by a large number of genes (Crabbe 2002; Plomin *et al.* 2001). Likewise, almost all human traits of interest to social scientists are complex, such as personality, cognition, motivation, and health behaviors. These traits are likely the consequence of many genetic and environmental factors, as well as interactions among them (Hirschhorn and Daly 2005; Lander and Botstein 1986; Lander and Schork 1994). Therefore it is important to incorporate multi-genetic and multi-environmental factors in gene-environment interaction (G×E) research on complex social outcomes.

In this study, we consider 403 genetic variants from 39 genes shown in animal studies to be related to aggression (Maxson 2009; Maxson and Canastar 2003; Miczek *et al.* 2001). We assess the collective contribution of these genetic variants to youth delinquency and violence using a recently developed mixed linear model approach in genomics studies that simultaneously accounts for a large number of genetic variables in a single regression analysis (Yang *et al.* 2011b). Moreover, we compare the collective genetic contribution to delinquency and violence between individuals exposed to environments with lower levels of social control and those who were exposed to environments with higher levels of social control (e.g., low parental attachment versus high/moderate parental attachment; loose school discipline versus strict/moderate school discipline; and disadvantaged neighborhoods versus non-disadvantaged neighborhoods). We find consistent evidence that genetic risk for adolescent delinquency and violence is largely context-dependent: genetic risk is amplified among individuals under low-social-control (LSC) conditions, but suppressed among those under high/moderate-social-control (HMSC) conditions.

CONCEPTUAL FRAMEWORK AND RESEARCH HYPOTHESES

Gene-environment interaction for delinquency

Genetic factors affect but do not determine human behavior, and their effect largely depends on the environment in which individuals live (Rutter *et al.* 2006). As animal and human studies show, changes in environmental conditions can influence expression of genes related to various phenotypes (Barr *et al.* 2004; Bennett *et al.* 2002; Chen *et al.* 2009; Cole *et al.* 2010; Newman *et al.* 2005; Tung *et al.* 2012). With respect to delinquent and violent behaviors, the *environmental triggering/suppressing* perspective offers important contributions to our understanding of how the social environment moderates genetic influence.[1]

There are two components to the *environmental triggering/suppressing* perspective. First, adverse environments are likely to "*trigger*" the expression of risk alleles (Shanahan and Hofer 2005). This *"triggering"* mechanism is also referred to as the *diathesis stress* model

[1]There is also a growing *differential susceptibility* perspective. Accordingly, individuals who are sensitive to adverse environments also tend to be susceptible to favorable environments. This implies that those who benefit the most from advantaged social conditions may be the same as those who suffer most in adverse social environments. As demonstrated by Simons et al. (2011), when exposed to adverse social environments with low social control, children with both s-allele *5-HTTLRP* and l-allele *DRD 4*, relative to those with other genotypes, show higher levels of violence-related characteristics such as "aggression, anger, hostile view of relationships, and concern with toughness." Yet the same children tend to have fewer such characteristics than others when exposed to low adversity and high social control.

(Ellis *et al.* 2011). Central to this model is the coaction of the risk allele and the risk environment. For example, Caspi *et al.* (2002) identify an association between monoamine oxidase A (*MAOA*) genotypes and antisocial behaviors, but mainly among test subjects who experienced childhood maltreatment. Second, favorable environments may suppress the expression of risk alleles. Particularly, social norms and structural constraints can inhibit individuals' behavior and choices, thereby reducing genetic influence (Shanahan and Hofer 2005). As shown by Pescosolido *et al.* (2008), the association between gamma-aminobutyric acid receptor subunit alpha-2 (*GABRA2*) and alcoholism is reduced by family support. Similarly, the dopamine D2 receptor (*DRD2*) is found to contribute less to delinquency among male youths who had a closer relationship with their parents (Guo *et al.* 2008).

Most of these studies focus on a single or only a few genetic variants at a time (Beaver *et al.* 2008; Caspi *et al.* 2002; Foley *et al.* 2004; Guo *et al.* 2007; Kim-Cohen *et al.* 2006; Simons *et al.* 2011; Vanyukov *et al.* 2007). However, delinquent and violent behaviors are complex human traits that can be affected by a large number of genetic factors with small to moderate effects.[2] Therefore, it is crucial to investigate the collective contribution of multi-genetic factors to delinquency and violence.

How do we identify genes that potentially contribute to human delinquency and violence? Animal studies may shed some light on gene selection insofar as the molecular functions of a large number of genes are conserved to a great extent across species (Robinson *et al.* 2005). According to the Mouse Genome Sequencing Consortium, human and mouse genomes include similar numbers of genes. Approximately 99% of mouse genes have direct counterparts in humans (Gunter and Dhand 2002). Because of the high degree of homology between human and mouse genes, gene selection in human studies could be motivated by findings from rodent studies (Case *et al.* 2005; Murphy *et al.* 2001; Shih and Thompson 1999).

Heretofore, rodent studies have shown dozen of genes involved in mouse aggression. For instance, transgenic mice [3]overexpressing a mutant form of amyloid precursor protein (*APP*) or phenylethanolamine N-methyltransferase (*PNMT*) tend to display increased aggressive behavior (Moechars *et al.* 1998). Aggressive behavior is increased in β estrogen receptor knockout (ERKO) mice [4], and greatly reduced in both α ERKO and αβ ERKO mice (Ogawa *et al.* 2000; Ogawa *et al.* 1997; Scordalakes and Rissman 2003). Moreover, in a series of behavioral studies on aggression and mating behavior, male neuronal nitric oxide synthase (*nNOS*) knockout mice are shown to display a dramatic loss of behavioral inhibition characterized by persistent fighting and mounting behavior (Nelson *et al.* 1995). Besides, there is evidence that *nNOS* is also associated with female mice's maternal aggression (Gammie and Nelson 1999; Gammie *et al.* 2000). These findings could help us select genes for research in human delinquent and violent behaviors.

[2]Many other complex human traits (e.g., most common diseases) have been shown to be determined by multi-environmental and multi-genetic factors, where individual genetic variants generally have a small effect (Hirschhorn and Daly 2005).

[3]The transgenic technique is used to determine the function of a gene by forcing the expression of a gene and examining the consequences. A famous example is the use of transgenic mice to identify Sry (termed SRY for humans), the sex-determining region Y (Koopman et al. 1991). In the experiment, Sry gene sequences were microinjected into fertilized eggs. As a result, among the transgenic mice, two chromosomally female mice developed male phenotypes.

[4]Gene knockout is used to determine the function of a gene by removing a gene and examining the consequences.

Social moderators for delinquency

In this paper, we focus on interaction of delinquency-related genes and three important social institutions in childhood or adolescence: the family, the school, and the neighborhood. These social institutions not only contribute to inhibiting or reducing children's deviant acts, but also have a long-term impact on their development of characteristics relevant to future delinquency or crime (Gottfredson and Hirschi 1990; Hirschi 1969; Sampson and Laub 1993). Of particular interest to us are the roles of these institutions in shaping individual propensity or self-control that can have persistent influence over the life course.

Parenting factors, such as parental attachment and supervision, are the most important source of self-control. According to Gottfredson and Hirschi (1990), self-control is cultivated during early childhood through careful rearing and effective discipline, whereas low self-control is mainly attributed to ineffective parenting. That is, if the caregivers of a child neglect to monitor his/her behavior, fail to recognize his/her deviant behaviors or punish such behaviors, as a consequence, the child may lack the ability to delay gratification, be insensitive to others' needs and interests, as well as be unwilling to accept restrictions on his/her behavior, and become more likely to use forcible or violent means to achieve his/her ends. Cullen *et al.* (2008) summarize results from 13 empirical studies examining the relationship between self-control and various dimensions of parenting. Twelve of the 13 studies have provided evidence that less effective parenting is associated with weaker self-control.

School is another powerful social institution that helps adolescents develop self-control (Gottfredson and Hirschi 1990). Because the school has a particular interest in maintaining a good educational environment, it is expected to recognize and prevent antisocial behavior and it has the authority and means to implement effective discipline. As Denise Gottfredson (2001) suggests, "schools have the potential to teach self-control and to engage informal social controls to hold youthful behavior in check." Turner *et al.* (2005) show that the influence of school socialization on self-control is more effective for children of parents who failed in their task to teach self-control. Accordingly, school socialization may work to "pick up the slack" for inadequate parenting practices. This is consistent with the study of Meldrum (2008), in which self-control is found to be significantly predicted by school monitoring, even after controlling for familial factors.

In addition to family and school, neighborhood conditions are also critical for the development of self-control. Wikström and Sampson (2003) propose that individuals with weaker self-control are more likely to be found in disadvantaged neighborhoods with weak community capital and low collective efficacy (i.e., weak social cohesion among neighbors and their expectations to achieve common good), because these neighborhoods often lack resources and services, such as time, money, and knowledge, to support familial socialization practices. Empirical studies have offered mixed support for this position. Pratt *et al.* (2004) provide evidence that self-control is predicted by neighborhood conditions. In a more recent study, Gibson *et al.* (2010) also find support for associations between neighborhood structural characteristics and self-control, but these associations became nonsignificant after taking into account individual-level characteristics.

In summary, prior studies have demonstrated associations among the social environment, delinquency, and self-control. Although they do not directly address genetic factors, these studies are consistent with the G×E interaction view that the social environment may moderate individual propensities that have a long-term influence on delinquency. From the environmental triggering/suppressing perspective, we hypothesize that *genetic risk for delinquency and violence is greater among young adults who were weakly attached to parents and schools, loosely disciplined by parents or school authorities, or lived in disadvantaged neighborhoods than those who were closely/moderately attached to their parents and schools, strictly/moderately disciplined by parents or school authorities, or lived in non-disadvantaged neighborhoods.* Our study extends previous G×E research by incorporating a larger number of genetic variants selected from animal studies. Using 403 genetic variants from 39 genes shown by transgenic and knockout studies to be related to aggression in mice, we examine the genetic variants' collective contribution to youth delinquency and violent behaviors.

DATA AND MEASUREMENT

Data

Data for this study come from the National Longitudinal Study of Adolescent Health (Add Health). Add Health is a longitudinal survey of U.S. adolescents in grades 7 through 12 from 1994 to 1995 (In-School, N = 90,118; Wave I, N = 20,745). The Add Health cohort was followed up in 1996 (Wave II, N = 14,738) and again from 2001 to 2002 (Wave III, N = 15,197) (Harris *et al.* 2003). Based on responses to the in-school survey, twin, full, half, and step siblings were oversampled for in-home interviews, resulting in 5,740 individuals. At Wave III, twins and full siblings (N = 2,600) were asked to provide buccal cells for genotyping (Harris *et al.* 2013). Our genotyping was supported by a major National Science Foundation (NSF) grant. We targeted 1,536 single-nucleotide polymorphisms (i.e., genetic variants that occur when a single nucleotide [e.g., A, T, C, or G] in the genome is altered) in an Illumina 1536-SNP array; the 1,536 SNPs included 186 ancestral informative markers and genetic markers in 57 candidate genes associated with aggressive behavior in mice (Maxson 2009). In the standard quality control, we excluded individuals with 10% or more missing genotype data and SNPs with a call rate of less than 99% or a minor allele frequency smaller than 0.01. The quality control yielded 403 SNPs from 39 autosomal genes (see Table A1 for more details about the 39 genes, Table A2 for rs ids of the 403 SNPs, and Figure A1 for SNP correlations) for 2,262 individuals from 1,425 families. Because our analytic model requires genetically unrelated individuals to obtain unbiased results, we randomly selected one individual from each family, thereby reducing the effective sample size to the number of families.

Variable measurement

Outcome variables: serious delinquency and violence scores—Our outcome variables are based on 12 items from Add Health questionnaires at Wave III: (1) deliberately damaged others' property, (2) so badly hurt someone that medical treatment was needed, (3) used a weapon to get something from someone, (4) took part in group fights, (5) carried a weapon, (6) pulled a knife or gun on someone, (7) shot or stabbed someone, (8) took part in

fights in which self was injured, (9) stole something worth more than $50, (10) broke into a house or building to steal, (11) sold drugs, and (12) stole something worth less than $50 (Cronbach's alpha = .68). To be consistent with the delinquency literature (Hagan and Foster 2003; Hannon 2003), we divided the 12 questions into violent and nonviolent categories. The serious delinquency score is a summed index of all twelve items that ranges from 0 to 36, with higher scores indicating greater delinquency. The violence score is a summed index based upon the first 8 items.[5] We chose outcomes from a single wave because our analytic model does not allow repeated measures. Also, we used outcomes measured at Wave III and social-environmental measures from Wave I to minimize reverse causality.

Socio-environmental variables: Parenting factors—To simplify the G×E analysis, we constructed each social-environmental variable as a dichotomous variable. We assessed *parental attachment* using two Wave I questions asking how close a respondent felt to his or her mother and father and a question concerning the respondent's feeling about how his or her parents cared about him or her (alpha = .62). If the average of a respondent's answers to three questions was greater than or equal to the first sample tertile (i.e., 1/3 cut-off), for him or her, *Parental attachment* was coded as 1, indicating high/moderate parental attachment, and 0 otherwise (indicating low parental attachment). *Parental supervision* was constructed based on seven Wave I questions asking the respondent if his or her parents allowed him or her to make their decisions about the following: the time they must be home on weekend nights; the people they hang around with; what they wear; how much television they watch; which television programs they watch; what time they go to bed on week nights; and what they eat (alpha = .62). *Parental supervision* was coded as 1 if the average of a respondent's answers to seven questions was greater or equal to the first sample tertile (indicating strict/moderate parental supervision), and 0 otherwise (indicating loose parental supervision).

School factors—We used two Wave I measures to assess school factors: *school attachment* and *school discipline*. To measure *school attachment*, we averaged responses to three questions (alpha = .77) asking whether a respondent (rated on a scale of 1 to 5) felt close to people at school, felt like being part of the school, or felt happy at school, and to measure *school discipline*, we averaged school administrators' responses to eleven questions (alpha = .73) asking in their schools what happens to a student who is caught the second time fighting with another student, injuring another student, possessing alcohol, possessing an illegal drug, possessing a weapon, drinking alcohol at school, using an illegal drug at school, smoking at school, verbally abusing a teacher, physically injuring a teacher, and stealing school property (1 = no policy; 2 = verbal warning; 3 = minor action; 4 = in-school suspension; 5 = out-of-school suspension;6 = expulsion). Like *parental attachment* and *parental supervision*, *school attachment* and *school discipline* were dichotomized on the basis of the first sample tertile (coded as 1 if the average of the items was equal to or greater than the first sample tertile, indicating high/moderate school attachment and strict/moderate

[5]Both outcome variables are right-skewed. We conducted sensitivity analysis to examine whether the right-skewness affects the results. For example, we compared results based on transformed outcomes (e.g., log-transformed outcomes) and those based on original outcomes. Those results are consistent, indicating that our findings are robust to tests of distributional assumptions. Results are available from the authors upon request.

school discipline, and 0 otherwise, indicating low school attachment or loose school discipline).

Neighborhood—We assessed neighborhood environment using four Wave I block level variables from the Add Health Public Contextual Database: *proportion of aged 25+ individuals with college degree or more, proportion of households with income less than $15,000, unemployment rate* and *proportion of own children under 18 years in families and subfamilies not living with both parents*. Block is a geographic area defined by the U.S. Bureau of the Census, which in 1990, averaged 452 housing units or 1,100 people (U.S. Bureau of the Census 1993). It is the lowest level of geography in sample data published by the census bureau, and therefore captures the most localized available contextual characteristics of the areas in which individuals live (Billy *et al.* 1998). We recoded each of the four variables into a 0–1 indicator. For example, the *unemployment* variable was coded as 1 if the *unemployment rate* of the block where the respondent lived was lower than or equal to the second sample tertile (indicating non-disadvantaged neighborhoods).[6]

Control variables—We controlled for bio-ancestry scores, gender, age, and age squared in all analyses of the collective genetic contribution to serious delinquency and violence. Bio-ancestry scores of Africans, Europeans and East Asians were calculated based on 186 ancestral informative markers (AIMs) using the Structure procedure (Pritchard *et al.* 2000). For each individual, the three scores sum to 1. These AIMs was developed to detect and correct population stratification for genetic association studies (Enoch *et al.* 2006). Moreover, associations between school or neighborhood factors and the outcomes might be confounded by family-level factors. For example, both living in a disadvantaged neighborhood and having higher levels of delinquency are possibly consequences of low family SES. Therefore, in G×E analyses involving school or neighborhood factors, we also controlled for family socioeconomic status and family structure.[7]

ANALYTICAL STRATEGY

At the first stage of our analysis, we employed a mixed linear model to estimate the collective genetic contribution of the 403 SNPs. The model was fit using the Genome-wide Complex Trait Analysis (GCTA) software package, a tool based on the work of Yang *et al.* (2011b) to estimate the overall genetic variance for complex human traits.

The mixed linear model offers the substantial advantage of simultaneously accounting for a large number of genetic variants. It was developed to address the “missing heritability” issue in genome-wide association studies (GWAS) (Yang *et al.* 2010). For example, whereas 80% of variance in human height is believed to be heritable, SNPs discovered by GWAS together can explain less than 10% of observed height variation (Visscher *et al.* 2012). In contrast to single-variant association analysis where each SNP is tested against an adjusted p-value

[6]We conducted sensitivity analysis using dichotomized variables based on other cut-offs such as the first quartile and the median. The main findings remain, suggesting that our findings are robust to different grouping strategies.

[7]To test the robustness of the results, we fit the models in various ways, such as controlling for family socioeconomic status, family structure, and census region in all models and controlling for Wave I delinquency or violence in addition to other covariates. The major findings were very similar in all models.

(e.g., 5×10^{-8} or smaller), the mixed linear model approach treats all SNP effects as random effects. Using this approach, Yang *et al.* (2011a) show common SNPs collectively explain 41.9%, 15.9%, 25.4%, and 16.8% of the total phenotypic variances in human height, body mass index (BMI), von Willebrand factor (vWF), and OT interval (QTi), whereas highly significant and well replicated SNPs identified by GWAS merely account for 10%, 1.5%, 13%, and 7%, respectively. This method has also been employed for common diseases (Lee *et al.* 2011), schizophrenia (Lee *et al.* 2012), intelligence (Chabris *et al.* 2012; Davies *et al.* 2011), personality traits (Vinkhuyzen *et al.* 2012), subjective well-being (Rietveld *et al.* 2013), economic and political phenotypes (Benjamin *et al.* 2012), but not yet for delinquency and violence.

Our model is described by the following equation:

$$Y=X\beta+W\mu+\varepsilon, \quad \text{(Equation 1)}$$

where Y is the outcome variable; β is a vector of fixed effects such as age, sex and other controls; μ is a vector of SNP effects with $\mu_i \sim N(0, \boldsymbol{\sigma^2_\mu})$ where i = 1,…, I, with I being the number of SNPs; ε is a vector of residual effects with $\varepsilon_j \sim N(0, \boldsymbol{\sigma^2_\varepsilon})$ where j = 1,…, J, with J being the number of individuals; W is a standardized genotype matrix with the ij[th] element $w_{ij}=(s_{ij}-2p_i)/\sqrt{[2p_i(1-p_i)]}$ where s_{ij} is the number of copies of the reference allele for the i[th] SNP of the j[th] individual[8] and p_i is the frequency of the reference allele.

Yang et al. (2010) innovatively applied a previous result that has been known in animal genetics (Goddard et al. 2009). The result defines g = Wμ, $\mathbf{A} = WW^T/I$ and $\boldsymbol{\sigma^2_g} = I\boldsymbol{\sigma^2_\mu}$. Then Equation 2 is mathematically equivalent to Equation 1:

$$Y=X\beta+g+\varepsilon, \text{ with } V=\mathbf{A}\boldsymbol{\sigma^2_g}+I_\varepsilon\sigma^2_\varepsilon, \quad \text{(Equation 2)}$$

where g is an n*1 vector of the total genetic effects of the individuals with $g \sim N(0, \mathbf{A}\boldsymbol{\sigma^2_g})$, **A** is the genetic relationship matrix (GRM) between individuals and $\boldsymbol{\sigma^2_g}$ is the total genetic variance explained by the SNPs. Hence $\boldsymbol{\sigma^2_g}$ can be estimated by the restricted maximum likelihood (REML) approach, depending on the GRM estimated from all SNPs. In this study, the collective genetic contribution is assessed using the proportion of total variance in the outcome explained by all SNPs, which can be expressed as $\boldsymbol{\sigma^2_g/(\sigma^2_g+\sigma^2_\varepsilon)}$.

As noted earlier, the mixed linear model requires genetically unrelated individuals. Due to common environmental effects, including individuals from the same families could have resulted in a biased estimate of the genetic variance (Yang *et al.* 2011b). Because of this, we randomly selected one individual from each family to form a subsample. Using the subsample, we applied the mixed linear model to estimate the collective genetic contribution after controlling for potential confounding factors such as age, sex, bio-ancestry scores and

[8]Common SNPs typically have only two alleles. There are three possible combinations of two alleles in a population (e.g., CG, CC and GG). Either of the two alleles can be chosen as the reference allele. For example, for a SNP that includes alleles "C" and "G," suppose we choose "G" as the reference allele. If the ith SNP of the jth individual is "CC", then s_{ij}, the number of copies of the reference allele, equals 0 as there is no "G" in the combination "CC." Similarly, in cases of "CG" or "GC", $s_{ij} = 1$ as there is one copy of "G" in either of the two combinations, and if "GG", $s_{ij} = 2$ as there are two copies of "G."

etc. However, either member of siblings in a family was equally likely to be included in the subsample. To avoid arbitrariness, we repeated the steps 500 times (estimated the collective genetic contribution using each of the 500 subsamples) and averaged the results.

Next, we performed two types of hypothesis testing to test whether genes interact with social environments influencing youth delinquent and violent behavior.[9] In the first type of hypothesis testing, we compared the collective genetic contribution to delinquency and violence between individuals under LSC conditions and those under HMSC conditions. We split the whole sample into two strata on the basis of each constructed dichotomous socio-environmental variable (e.g., one stratum only includes individuals under LSC conditions and the other includes those under HMSC conditions).[10] Within each stratum we selected 500 subsamples, for each of which we applied the mixed linear model to estimate the collective genetic contribution. For each stratum, results of 500 replications provided an empirical distribution of the collective genetic contribution. We then compared the empirical distributions between two strata. Secondly, we assessed individual SNP effects using the best linear unbiased predictors (BLUPs) estimated by the mixed linear model,[11] and employed the F test to compare the distribution of individual SNP effects under LSC and HMSC conditions.

RESULTS

Genetic contribution

Table 2 displays the estimates of the collective genetic contribution to serious delinquency and violence. As can be seen, estimates of the total variance in serious delinquency attributable to the 403 SNPs are non-significant at the .05 level. In the face of G×E, we might expect greater genetic risk for individuals exposed to LSC environments, and weaker risk for those who were exposed to HMSC environments in the sample. Next, we tested whether the collective genetic contribution to serious delinquency and violence differs under LSC and HMSC conditions.

Genetic contribution under differential conditions

Table 3 shows the results of comparing the collective genetic contribution of the 403 SNPs to serious delinquency and violence under differential conditions. Columns 1 and 3 contain estimates of the collective genetic contribution to serious delinquency under HMSC and LSC conditions, and columns 5 and 7 contain estimates for violence. Each number in the four columns is an average of 500 results. In Table 3, most estimates of the collective genetic contribution under LSC conditions are greater than those under HMSC conditions

[9] Yang et al. (2011a) already implemented a G×E interaction mixed linear model for GWAS data. The model is specified as: $Y=X\beta + g + ge + \varepsilon$, with $V = \mathbf{A_g}\,\sigma_g^2 + \mathbf{A_{ge}}\,\sigma_{ge}^2 + \boldsymbol{I_\varepsilon}\sigma_\varepsilon^2$, where ge is an n*1 vector of the G×E effects for all of the individuals with $\mathbf{A_{ge}} = \mathbf{A_g}$ for the pairs of individuals in the same environment and with $\mathbf{A_{ge}} = \mathbf{0}$ for the pairs of individuals in different environments. In addition to the genetic variance, this model estimate the variance explained by G×E. When statistically significant, this variance suggests that the SNPs of those in the same environment explains a higher portion of variance than those in different environments. However, this model cannot easily be used to test the hypothesis whether the proportion of the phenotypic variance explained by all SNPs and individual SNP effects differ between environmental conditions. We expand Yang et al.'s main effect mixed linear model to test such hypotheses.

[10] Observations with missing values in the socio-environmental variables were excluded in G×E analyses.

[11] As equations 1 and 2 (i.e. $Y=X\beta + W\mu + \varepsilon$ and $Y=X\beta + g + \varepsilon$) are mathematically equivalent, the BLUP of μ can be transformed from the BLUP of g by $\mu = W^T A^{-1} \hat{\boldsymbol{g}}$/I.

(with exceptions of neighborhood education and single-parent households for violence). For example, the proportion of total variance in the serious delinquency score explained by the 403 SNPs is estimated to be 3.1% for adolescents poorly attached to school, but the proportion drops to 0% for those who were closely/moderately attached to school.

Individual SNP effects under differential conditions

As mentioned earlier, the mixed linear model also provides estimates of individual SNP effects. Figure 1 plots the distributions of individual SNP effects on serious delinquency under differential conditions. As it shows, the spread of the SNP effects under most LSC conditions appears to be greater than HMSC conditions, suggesting a greater proportion of SNPs with a relatively large effect under LSC conditions than HMSC conditions. For example, for individuals poorly attached to school at Wave I, approximately 7% of the 403 SNPs have an effect size greater than 0.01 on serious delinquency,[12] while for those who were highly/moderately attached to school, none of the SNPs fall into that range. We used the F test to compare distributions of the individual SNP effects under LSC and HMSC conditions. As shown by Table 4, results are significant at the .05 level for most socioenvironmental variables (exceptions are neighborhood education and single-parent households).

To summarize, there is evidence that genetic risk for delinquency and violence is greater for adolescents who were weakly attached to parents and school, loosely disciplined by parents or school authorities, or lived in neighborhoods with lower income levels and higher unemployment rates as opposed to those who were closely attached to their parents and school, strictly disciplined by parents or school authorities, or lived in neighborhoods with higher income levels and lower unemployment rates.

Assessing effects of population stratification and gene-environment correlation

While our analysis shows significant interactions of aggression-related genetic variants and socioenvironmental variables, the story is, in fact, more complicated. The results could be driven by population stratification or gene-environment correlation (rGE). In mixed linear models, GRM values are usually higher for pairs from similar racial groups than for those from different racial groups. Because of that, genetic contribution estimates might be confounded by population stratification. We compared model results with and without controlling for bio-ancestry scores. The effect size of the genetic contribution shrinks around 20% after including the bio-ancestry scores in the model. This suggests that the bio-ancestry scores are effective in adjusting for population stratification. Moreover, we fit the models to individuals from the same racial groups in the sample. The major findings remain in spite of reduced statistical power.

rGE occurs when one's exposure to an environment depends upon his or her genotype. The existence of rGE may confound the G×E effects (Caspi and Moffitt 2006; Jaffee and Price 2007; Wagner *et al.* 2013). To detect rGE, we applied the mixed linear model to examine the

[12]The effect could be in either positive or negative direction. An effect size of 0.01 means that an increase of 1 risk allele is associated with 0.01 unit increases in the serious delinquency score.

association between the 403 SNPs and each of the eight socio-environmental responses. Table 5 shows all the socio-environmental variables cannot be significantly predicted by the 403 SNPs, indicating an absent or weak correlation between the socioenvironmental variables and SNPs included in this study.

DISCUSSION AND CONCLUSIONS

In this paper we hypothesize that high social control suppresses genetic risk for youth delinquency and violence, and low social control exacerbates genetic risk. We examine the influences of crucial social institutions in childhood or adolescence, such as the family, the school, and the neighborhood, on the collective genetic contribution of more than 400 SNPs. Consistent with the *environmental triggering/suppressing* perspective, we find that favorable social conditions are associated with smaller collective genetic contribution, whereas adverse social conditions are associated with greater collective genetic contribution to adolescent delinquency and violence.

This study makes several important contributions to the G×E literature. First, we consider 403 SNPs from 39 aggression-related genes identified in animal transgenic and knockout studies. This is a crucial improvement over previous research, which normally studies one genetic factor or only a few at a time. Delinquent and violent behaviors are complex human traits, meaning they could be affected by a large number of genetic and environmental factors. It is likely that the effects of many genetic variants are too small to be detected by testing each one individually for an association with the phenotype. However, these variants, collectively, could make a substantial contribution.

Second, we find that genetic risk of the 403 SNPs is smaller under favorable conditions than adverse conditions. These findings are consistent with results in previous G×E research based on a one or a few genetic variants (Caspi *et al.* 2002; Guo *et al.* 2008; Pescosolido *et al.* 2008). What is more, our findings highlight the influence of social control on genetic risk of many variants at the same time. These findings illuminate one mechanism through which social control affects delinquency and violence: it is possible that the presence of social control simultaneously prevents the expression of a large number of genetic variants associated with aggression and violence. In an environment under high social control, such as high family attachment, there may be adolescents varying in their genetic propensities for delinquent behaviors; some may possess risk alleles related to delinquency. Yet, the expression of risk alleles is prevented due to strong social control. When the control is weakened, for example, parents pay less attention, the adolescent with high genetic propensities for delinquency, relative to one with low genetic propensities, may be more apt to show gene expression.

Our third contribution is methodological. We test G×E involving a large number of genetic variants. Our method is an extension of the recent mixed linear model approach (Yang *et al.* 2011b). Compared to conventional linear regression models, the key advantage of the mixed linear model is its ability to simultaneously account for a large number of genetic variants. To illustrate, in conventional linear models, one socioenvironmental factor and the 403 SNPs would generate 403 two-way interaction terms in total. Analyses dependent on such

models typically do not have sufficient statistical power to produce significant results. However, in the mixed linear model, being treated as random effects, the 403 SNPs could be considered simultaneously. That allows us to estimate and compare the collective genetic contribution of the 403 SNPs under differential social conditions.

Although this study provides important insights in understanding how the social environment moderates genetic influence on delinquency and violence, some limitations should be noted. Our 403 SNPs are selected based on mouse models. In animal studies, experimental techniques such as transgenic and knockout techniques are used to determine the function of a gene. Animal studies involve various experimental control, including specific measurements of outcomes (e.g., duration and intensity of aggression), assessments of time between stimuli and outcomes, specific environments in which the experiments take place. In contrast, human outcome measures are typically self-reported, and tend to lack specificity (e.g., when, where, how etc.). These differences in scientific methods may result in barriers to apply findings from animal models to humans. Moreover, the mixed linear model approach does not allow genetically related individuals and repeated measures, leading to a reduction of the effective sample size. Also, because of the relatively small sample size, we have to dichotomize the social-environmental variables (if there were more categories, the G×E analysis would require a much larger sample to have sufficient statistical power), which might result in some loss of information. With more samples, future research might replicate the analyses in this study using more refined socioenvironmental measures.

Despite these limitations, our study makes important contributions to social sciences. It underscores the significance of the dialogue between the biological and social sciences. Social scientists traditionally have assumed homogeneous human nature at birth and focused on social structural influences on individuals. However, there is growing evidence that the social environment modifies gene expression (Morgan *et al.* 2002; Norman *et al.* 2012), and genetic variability, in turn, affects individuals' responses to the environment (Freese 2008). Increasingly available molecular genetic data in large-scale datasets (e.g., Add Health, the Fragile Families Study, and the Health and Retirement Study) enable social scientists to investigate how socioenvironmental factors shape human behavior through moderating genetic effects. The conceptual framework and methodology in this study can be expanded to study other behavioral and social consequences of the complex interplay of multi-genetic and multi-environmental factors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research uses data from the National Longitudinal Study of Adolescent Health (Add Health), a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on

the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis. A National Science Foundation (NSF) grant to Guang Guo supported the genotyping of the aggression-related genes in the genetic sample of Add Health (NSF's Human and Social Dynamics program BCS-0826913). Special acknowledgment is due Kirk Wilhelmsen of the Genetics Department, Patricia Basta of the Bio-Specimen Process Center, Jason Luo of the Mammalian Genotyping Center, and the Odum Institute at the University of North Carolina at Chapel Hill. We received important assistance in SNP selection and the analysis of HGDP data from David Goldman and his Neurogenetics lab at NIAAA. Many thanks go to Kathleen Harris, François Nielsen, Carl Roberts, Brandon Wagner, the editor of Social Forces, and three anonymous reviewers for their helpful comments on the manuscript. We are grateful to the Carolina Population Center (R24 HD050924) for general support.

References

Barr, Christina S.; Goldman, David; Suomi, Stephen J.; Dee Higley, J.; Newman, Timothy K., et al. Rearing Condition and Rh5-Httlpr Interact to Influence Limbic-Hypothalamic-Pituitary-Adrenal Axis Response to Stress in Infant Macaques. Biological Psychiatry. 2004; 55:733–738. [PubMed: 15039002]

Beaver, Kevin M.; Wright, John Paul; DeLisi, Matt; Vaughn, Michael G. Desistance from Delinquency: The Marriage Effect Revisited and Extended. Social Science Research. 2008; 37:736–752. [PubMed: 19086113]

Benjamin, Daniel J.; Johannesson, Magnus; Peter Visscher, M.; Cesarini, David; van der Loos, Matthijs JHM., et al. The Genetic Architecture of Economic and Political Preferences. Proceedings of the National Academy of Sciences. 2012; 109:8026–8031.

Bennett AJ, Higley JD, Lesch KP, Heils A, Long JC, et al. Early Experience and Serotonin Transporter Gene Variation Interact to Influence Primate Cns Function. Molecular Psychiatry. 2002; 7:118–122. [PubMed: 11803458]

Boardman, Jason D.; Benjamin Domingue, W.; Jason Fletcher, M. How Social and Genetic Factors Predict Friendship Networks. Proceedings of the National Academy of Sciences. 2012; 109:17377–17381.

Case, Anne; Fertig, Angela; Paxson, Christina. The Lasting Impact of Childhood Health and Circumstance. Journal of Health Economics. 2005; 24:365–389. [PubMed: 15721050]

Caspi, Avshalom; Moffitt, Terrie E. Gene-Environment Interactions in Psychiatry: Joining Forces with Neuroscience. Nature Reviews Neuroscience. 2006; 7:583–590.

Caspi, Avshalom; McClay, Joseph; Moffitt, Terrie E.; Mill, Jonathan; Martin, Judy, et al. Role of Genotype in the Cycle of Violence in Maltreated Children. Science. 2002; 297:851–854. [PubMed: 12161658]

Chabris, Christopher F.; Hebert, Benjamin M.; Benjamin, Daniel J.; Beauchamp, Jonathan P.; Cesarini, David, et al. Most Reported Genetic Associations with General Intelligence Are Probably False Positives. Psychological Science. 2012; 23:1314–1323. [PubMed: 23012269]

Chen E, Miller GE, Walker HA, Arevalo JM, Sung CY, et al. Genome-Wide Transcriptional Profiling Linked to Social Class in Asthma. Thorax. 2009; 64:38–43. [PubMed: 19001005]

Cole, Steven W.; Arevalo, Jesusa MG.; Takahashi, Rie; Sloan, Erica K.; Lutgendorf, Susan K., et al. Computational Identification of Gene-Social Environment Interaction at the Human Il6 Locus. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:5681–5686. [PubMed: 20176930]

Crabbe, John C. Genetic Contributions to Addiction. Annual Review of Psychology. 2002; 53:435–462.

Cullen, Francis T.; Unnever, James D.; Wright, John Paul; Beaver, Kevin M. Parenting and Self-Control. Out of Control: Assessing the General Theory of Crime. 2008:61–74.

Davies G, Luciano M, McGhee K, Lopez L, Gow AJ, et al. Genome-Wide Association Studies Establish That Human Intelligence Is Highly Heritable and Polygenic. Molecular Psychiatry. 2011; 16:996–1005. [PubMed: 21826061]

Ellis, Bruce J.; Thomas Boyce, W.; Belsky, Jay; Bakermans-Kranenburg, Marian J.; van Ijzendoorn, Marinus H. Differential Susceptibility to the Environment: An Evolutionary–Neurodevelopmental Theory. Development and Psychopathology. 2011; 23:7–28. [PubMed: 21262036]

Enoch, Mary-Anne; Shen, Pei-Hong; Xu, Ke; Hodgkinson, Colin; Goldman, David. Using Ancestry-Informative Markers to Define Populations and Detect Population Stratification. Journal of Psychopharmacology. 2006; 20:19–26. [PubMed: 16785266]

Foley, Debra L.; Eaves, Lindon J.; Wormley, Brandon; Silberg, Judy L.; Maes, Hermine H., et al. Childhood Adversity, Monoamine Oxidase a Genotype, and Risk for Conduct Disorder. Archives of General Psychiatry. 2004; 61:738–744. [PubMed: 15237086]

Fowler, James H.; Settle, Jaime E.; Christakis, Nicholas A. Correlated Genotypes in Friendship Networks. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108:1993–1997. [PubMed: 21245293]

Freese, Jeremy. Genetics and the Social Science Explanation of Individual Outcomes. The American Journal of Sociology. 2008; 114:S1–S35.

Gammie SC, Nelson RJ. Maternal Aggression Is Reduced in Neuronal Nitric Oxide Synthase-Deficient Mice. The Journal of neuroscience: the official journal of the Society for Neuroscience. 1999; 19:8027–8035. [PubMed: 10479702]

Gammie, Stephen C.; Olaghere-da Silva, Uade B.; Nelson, Randy J. 3-Bromo-7-Nitroindazole, a Neuronal Nitric Oxide Synthase Inhibitor, Impairs Maternal Aggression and Citrulline Immunoreactivity in Prairie Voles. Brain Research. 2000; 870:80–86. [PubMed: 10869504]

Gibson, Chris L.; Sullivan, Christopher J.; Jones, Shayne; Piquero, Alex R. “Does It Take a Village?” Assessing Neighborhood Influences on Children’s Self-Control. Journal of Research in Crime and Delinquency. 2010; 47:31–62.

Glazier, Anne M.; Nadeau, Joseph H.; Aitman, Timothy J. Finding Genes That Underlie Complex Traits. Science. 2002; 298:2345–2349. [PubMed: 12493905]

Gottfredson, Denise C. Schools and Delinquency. Cambridge University Press; 2001.

Gottfredson, Michael R.; Hirschi, Travis. A General Theory of Crime. Stanford University Press; 1990.

Gunter, Chris; Dhand, Ritu. Human Biology by Proxy. Nature. 2002; 420:509–509.

Guo, Guang; Roettger, Michael E.; Cai, Tianji. The Integration of Genetic Propensities into Social-Control Models of Delinquency and Violence among Male Youths. American Sociological Review. 2008; 73:543–568.

Guo, Guang; Roettger, Michael E.; Shih, Jean C. Contributions of the Dat1 and Drd2 Genes to Serious and Violent Delinquency among Adolescents and Young Adults. Human Genetics. 2007; 121:125–136. [PubMed: 17120049]

Hagan, John; Foster, Holly. S/He’s a Rebel: Toward a Sequential Stress Theory of Delinquency and Gendered Pathways to Disadvantage in Emerging Adulthood. Social Forces. 2003; 82:53–86.

Hannon, Lance. Poverty, Delinquency, and Educational Attainment: Cumulative Disadvantage or Disadvantage Saturation? Sociological Inquiry. 2003; 73:575–594.

Harris, Kathleen Mullan; Halpern, Carolyn Tucker; Haberstick, Brett C.; Smolen, Andrew. The National Longitudinal Study of Adolescent Health (Add Health) Sibling Pairs Data. Twin Research and Human Genetics. 2013; 16:391–398. [PubMed: 23231780]

Hirschhorn, Joel N.; Daly, Mark J. Genome-Wide Association Studies for Common Diseases and Complex Traits. Nature Reviews Genetics. 2005; 6:95–108.

Hirschi, Travis. Causes of Delinquency. Los Angeles, CA: University of California; 1969.

Jaffee SR, Price TS. Gene-Environment Correlations: A Review of the Evidence and Implications for Prevention of Mental Illness. Molecular Psychiatry. 2007; 12:432–442. [PubMed: 17453060]

Kim-Cohen J, Caspi A, Taylor A, Williams B, Newcombe R, et al. Maoa, Maltreatment, and Gene-Environment Interaction Predicting Children’s Mental Health: New Evidence and a Meta-Analysis. Molecular Psychiatry. 2006; 11:903–913. [PubMed: 16801953]

Lander ES, Botstein D. Mapping Complex Genetic Traits in Humans: New Methods Using a Complete Rflp Linkage Map. Cold Spring Harbor Symposia on Quantitative Biology. 1986; 51:49–62.

Lander ES, Schork NJ. Genetic Dissection of Complex Traits. Science. 1994; 265:2037–2048. [PubMed: 8091226]

Lee, S Hong; DeCandia, Teresa R.; Ripke, Stephan; Yang, Jian; Sullivan, Patrick F., et al. Estimating the Proportion of Variation in Susceptibility to Schizophrenia Captured by Common Snps. Nature Genetics. 2012; 44:247–U35. [PubMed: 22344220]

Lee, Sang Hong; Wray, Naomi R.; Goddard, Michael E.; Visscher, Peter M. Estimating Missing Heritability for Disease from Genome-Wide Association Studies. The American Journal of Human Genetics. 2011; 88:294–305.

Maxson, Stephen C. The Genetics of Offensive Aggression in Mice. In: Kim, KY., editor. Handbook of Behavior Genetics. New York, NY: Springer New York; 2009. p. 301-316.

Maxson, Stephen C.; Canastar, Andrew. Conceptual and Methodological Issues in the Genetics of Mouse Agonistic Behavior. Hormones and Behavior. 2003; 44:258–262. [PubMed: 14609547]

Meldrum, Ryan Charles. Beyond Parenting: An Examination of the Etiology of Self-Control. Journal of Criminal Justice. 2008; 36:244–251.

Miczek KA, Maxson SC, Fish EW, Faccidomo S. Aggressive Behavioral Phenotypes in Mice. Behavioural Brain Research. 2001; 125:167–181. [PubMed: 11682108]

Moechars D, Gilis M, Kuipéri C, Laenen I, Van Leuven F. Aggressive Behaviour in Transgenic Mice Expressing App Is Alleviated by Serotonergic Drugs. Neuroreport. 1998; 9:3561–3564. [PubMed: 9858360]

Morgan, Drake; Nader, Michael A.; Grant, Kathleen A.; Donald Gage, H.; Mach, Robert H., et al. Social Dominance in Monkeys: Dopamine D2 Receptors and Cocaine Self-Administration. Nature Neuroscience. 2002; 5:169–174.

Murphy DL, Li Q, Engel S, Wichems C, Andrews A, et al. Genetic Perspectives on the Serotonin Transporter. Brain Research Bulletin. 2001; 56:487–494. [PubMed: 11750794]

Nelson RJ, Demas GE, Huang PL, Fishman MC, Dawson VL, et al. Behavioural Abnormalities in Male Mice Lacking Neuronal Nitric Oxide Synthase. Nature. 1995; 378:383–386. [PubMed: 7477374]

Newman, Timothy K.; Syagailo, Yana V.; Barr, Christina S.; Wendland, Jens R.; Champoux, Maribeth, et al. Monoamine Oxidase a Gene Promoter Variation and Rearing Experience Influences Aggressive Behavior in Rhesus Monkeys. Biological Psychiatry. 2005; 57:167–172. [PubMed: 15652876]

Norman, Greg J.; Hawkley, Louise C.; Cole, Steve W.; Berntson, Gary G.; Cacioppo, John T. Social Neuroscience: The Social Brain, Oxytocin, and Health. Social Neuroscience. 2012; 7:18–29. [PubMed: 21714746]

Ogawa S, Chester AE, Hewitt SC, Walker VR, Gustafsson JA, et al. Abolition of Male Sexual Behaviors in Mice Lacking Estrogen Receptors Alpha and Beta (Alpha Beta Erko). Proceedings of the National Academy of Sciences of the United States of America. 2000; 97:14737–14741. [PubMed: 11114183]

Ogawa, Sonoko; Lubahn, Dennis B.; Korach, Kenneth S.; Pfaff, Donald W. Behavioral Effects of Estrogen Receptor Gene Disruption in Male Mice. Proceedings of the National Academy of Sciences of the United States of America. 1997; 94:1476–1481. [PubMed: 9037078]

Pescosolido, Bernice A.; Perry, Brea L.; Long, J Scott; Martin, Jack K., Jr; Nurnberger, John I., et al. Under the Influence of Genetics: How Transdisciplinarity Leads Us to Rethink Social Pathways to Illness. American Journal of Sociology. 2008; 114:S171–S201.

Plomin, R.; DeFries, JC.; McClearn, GE.; McGuffin, P. Behavioral Genetics. Plomin, R., editor. New York: Worth Pubs; 2001.

Pratt, Travis C.; Turner, Michael G.; Piquero, Alex R. Parental Socialization and Community Context: A Longitudinal Analysis of the Structural Sources of Low Self-Control. Journal of Research in Crime and Delinquency. 2004; 41:219–243.

Pritchard JK, Stephens M, Donnelly P. Inference of Population Structure Using Multilocus Genotype Data. Genetics. 2000; 155:945–959. [PubMed: 10835412]

Rietveld, Cornelius A.; Agrawal, Arpana; Eriksson, Johan G.; Albrecht, Eva; Alizadeh, Behrooz Z., et al. Gwas of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. Science. 2013; 340:1467–1471. [PubMed: 23722424]

Robinson, Gene E.; Grozinger, Christina M.; Whitfield, Charles W. Sociogenomics: Social Life in Molecular Terms. Nature Reviews Genetics. 2005; 6:257–270.

Rutter, Michael; Moffitt, Terrie E.; Caspi, Avshalom. Gene-Environment Interplay and Psychopathology: Multiple Varieties but Real Effects. Journal of Child Psychology and Psychiatry, and Allied Disciplines. 2006; 47:226–261.

Sampson, Robert J.; Laub, John H. Crime in the Making: Pathways and Turning Points through Life. Harvard University Press; 1993.

Scordalakes, Elka M.; Rissman, Emilie F. Aggression in Male Mice Lacking Functional Estrogen Receptor Alpha. Behavioral Neuroscience. 2003; 117:38–45. [PubMed: 12619906]

Shanahan, Michael J.; Hofer, Scott M. Social Context in Gene-Environment Interactions: Retrospect and Prospect. The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences. 2005; 60(Spec No 1):65–76.

Shanahan, Michael J.; Vaisey, Stephen; Erickson, Lance D.; Smolen, Andrew. Environmental Contingencies and Genetic Propensities: Social Capital, Educational Continuation, and Dopamine Receptor Gene Drd2. American Journal of Sociology. 2008; 114(Suppl):S260–S286.

Shih JC, Thompson RF. Monoamine Oxidase in Neuropsychiatry and Behavior. American Journal of Human Genetics. 1999; 65:593–598. [PubMed: 10441564]

Simons, Ronald L.; Lei, Man Kit; Beach, Steven RH.; Brody, Gene H.; Philibert, Robert A., et al. Social Environment, Genes, and Aggression: Evidence Supporting the Differential Susceptibility Perspective. American Sociological Review. 2011; 76:883–912.

Tung, Jenny; Barreiro, Luis B.; Johnson, Zachary P.; Hansen, Kasper D.; Michopoulos, Vasiliki, et al. Social Environment Is Associated with Gene Regulatory Variation in the Rhesus Macaque Immune System. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109:6490–6495. [PubMed: 22493251]

Turner, Michael G.; Piquero, Alex R.; Pratt, Travis C. The School Context as a Source of Self-Control. Journal of Criminal Justice. 2005; 33:327–339.

Vanyukov, Michael M.; Maher, Brion S.; Devlin, Bernie; Kirillova, Galina P.; Kirisci, Levent, et al. The Maoa Promoter Polymorphism, Disruptive Behavior Disorders, and Early Onset Substance Use Disorder: Gene-Environment Interaction. Psychiatric genetics. 2007; 17:323–332. [PubMed: 18075472]

Vinkhuyzen AAE, Luciano M, Payton A, Horan M, Ollier W, et al. Common Snps Explain Some of the Variation in the Personality Dimensions of Neuroticism and Extraversion. Translational Psychiatry. 2012; 2:e102. [PubMed: 22832902]

Visscher, Peter M.; Brown, Matthew A.; McCarthy, Mark I.; Yang, Jian. Five Years of Gwas Discovery. American journal of human genetics. 2012; 90:7–24. [PubMed: 22243964]

Wagner, Brandon; Li, Jiang; Liu, Hexuan; Guo, Guang. Gene-Environment Correlation: Difficulties and a Natural Experiment-Based Strategy. American Journal of Public Health. 2013; 103(Suppl 1):S167–73. [PubMed: 23927502]

Wikström, Per-Olof H.; Sampson, Robert J. Social Mechanisms of Community Influences on Crime and Pathways in Criminality. In: Lahey, B.; Moffitt, T.; Caspi, A., editors. The Causes of Conduct Disorder and Serious Juvenile Delinquency. New York, NY: Guilford; 2003. p. 118-148.

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, et al. Common Snps Explain a Large Proportion of the Heritability for Human Height. Nature Genetics. 2010; 42:565–569. [PubMed: 20562875]

Yang, Jian; Geoffrey Hayes, M.; Hill, William G.; Landi, Maria Teresa; Alonso, Alvaro, et al. Genome Partitioning of Genetic Variation for Complex Traits Using Common Snps. Nature Genetics. 2011a; 43:519–525. [PubMed: 21552263]

Yang, Jian; Hong Lee, S.; Goddard, Michael E.; Visscher, Peter M. Gcta: A Tool for Genome-Wide Complex Trait Analysis. The American Journal of Human Genetics. 2011b; 88:76–82.

Biographies

Hexuan Liu is currently a PhD candidate in the Department of Sociology at the University of North Carolina at Chapel Hill. His research interests include health disparity, delinquency and crime, quantitative methods, and the integration of genetics and social sciences.

Yi Li is a doctoral candidate in the Department of Sociology at the University of North Carolina at Chapel Hill. His research primarily focuses on the integration of genetics with sociological studies of delinquency and crime, health, marriage, gender, and inequality. He is also interested in causality and quantitative methods.

Guang Guo is Dr. George and Alice Welsh Distinguished Professor in the Department of Sociology at the University of North Carolina at Chapel Hill. His research focuses on the integration of sociology with genetics and epigenetics in the studies of fundamental sociological issues such as social and health behavior in humans, production of social stratification, and bio- ancestry and social construction of racial and ethnic identity. He has recently published in the American Sociological Review, American Journal of Sociology, Demography, Sociological Methodology, Social Science Research, and PLoS One.

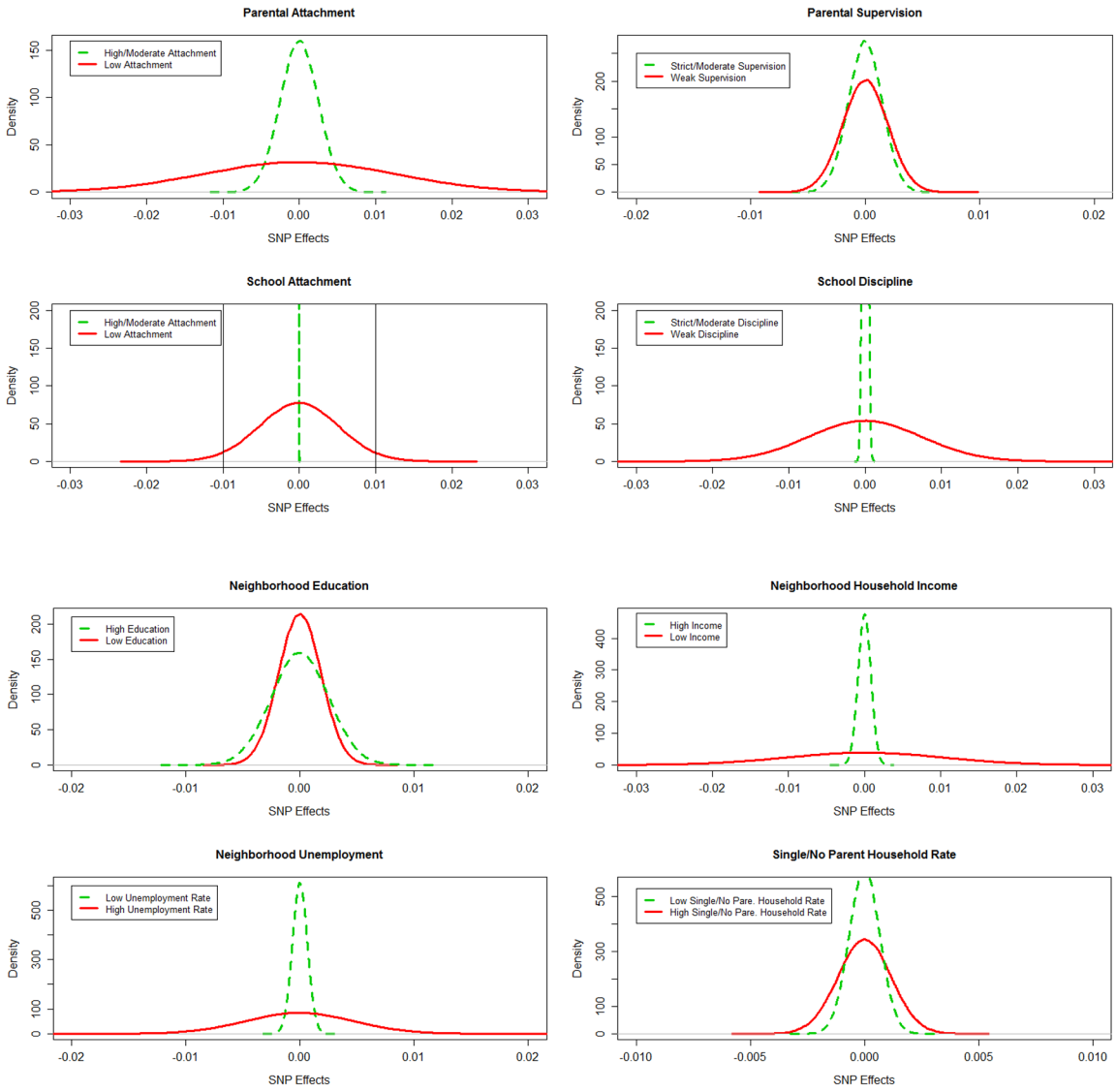

Figure 1.

Individual SNP Effects on Serious Delinquency. See the test results in Table 4.

Note: (1) Individual SNP effects are plotted along the horizontal axis and the effects' density along the vertical axis. (2) All densities follow a normal distribution with a mean of 0 (the density for high/moderate school attachment does not appear normal due to its small variance). (3) A greater spread of the distribution suggests a larger proportion of SNPs with relatively large effects on serious delinquency. Above figures show there are more SNPs with relatively large effects under most low-social-control conditions (solid lines) than high/moderate-social-control conditions (dashed lines). For example, for individuals poorly attached to school at Wave I, approximately 7 percent of the 403 SNPs have an effect size greater than 0.01 on serious delinquency at Wave III (the area under the curve and not in between the vertical lines), whereas for those who were highly/moderately attached to school, none of the SNPs fall into that area.

Table 1

Variable Description

Variable Name	Description	Mean or Proportion	SD
Delinquency and Violence			
Delinquency	Serious Delinquency Score, Wave III	.691	1.751
Violence	Violence Score, Wave III	.381	1.097
Demographics			
Bio-ancestry (Europe)	European bio-ancestry score	.699	.397
Bio-ancestry (Africa)	African bio-ancestry score	.184	.351
Bio-ancestry (Asian)	European bio-ancestry score	.117	.259
Age	Respondent's age at the time of Wave III	21.949	1.709
Female	Respondent's gender	.514	
PVT < 90	Verbal IQ less than 90 at Wave I	.223	
PVT 90 to 110	Verbal IQ between 90 and 110 at Wave I	.467	
PVT >110	Verbal IQ greater than 90 at Wave I	.272	
PVT missing	Missing on IQ score at Wave I	.038	
West	Lives in West state at Wave I	.164	
Midwest	Lives in Midwest state at Wave I	.317	
South	Lives in Southern state at Wave I	.362	
Northeast	Lives in Northeast state at Wave I	.152	
Region missing	Missing on region	.005	
Family SES			
High school or higher	Parent has at least high school education at Wave I	.840	
No high school	Parent has less than high school education at Wave I	.112	
Parent education Missing	Missing on parent education at Wave I	.048	
Family Structure			
Two biological parents	Lives with both biological parents at Wave I	.617	
Parenting Factors			
High/moderate parental attachment	High/moderate emotional attachment to resident parents at Wave I	.785	
Low parental attachment	Low emotional attachment to resident parents at Wave I	.211	
Parental attachment missing	Missing on emotional attachment to resident parents	.004	
Strict/moderate parental supervision	Strict/moderate parental supervision at Wave I	.796	
Weak parental supervision	Weak parental supervision at Wave I	.187	
Parental supervision missing	Missing on parental supervision	.016	
School Factors			
High/moderate school attachment	High/moderate emotional attachment to school at Wave I	.687	
Low school attachment	Low emotional attachment to school at Wave I	.291	
School attachment missing	Missing on school attachment	.021	
Strict/moderate school discipline	Strict/Moderate school discipline at Wave I	.471	
Low school discipline	Weak school discipline at Wave I	.264	
School discipline missing	Missing on school attachment	.264	
Neighborhood			

Variable Name	Description	Mean or Proportion	SD
High/moderate education	Respondent lives in higher education blocks at Wave I	.664	
Low education	Respondent lives in lower education blocks at Wave I	.330	
Education missing	Missing on education	.007	
High/moderate income	Respondent lives in higher income blocks at Wave I	.662	
Low income	Respondent lives in lower income blocks at Wave I	.331	
Income missing	Missing on income	.007	
Low/moderate unemployment rate	Respondent lives in blocks with lower unemployment rate at Wave I	.653	
High unemployment Rate	Respondent lives in blocks with higher unemployment rate at Wave I	.326	
Unemployment rate missing	Missing on unemployment rate	.020	
Low/moderate single/no parent household rate	Respondent lives in blocks with lower single/no-parent household rate at Wave I	.656	
High single/no parent household rate	Respondent lives in blocks with higher single/no-parent household rate at Wave I	.328	
Single/no parent household missing	Missing on single/no-parent household rate	.016	

Table 2

The Collective Genetic Contribution of 403 SNPs to Serious Delinquency and Violence and Standard Errors

	Serious Delinquency (Wave III)	Violence (Wave III)
Collective genetic contribution (Proportion of total variance explained by SNPs)	.007(.014)	.010(.015)
Intercept	7.853(6.596)	3.187(4.134)
Bio-ancestry (Europe)	--	--
Bio-ancestry (African)	.071(.173)	.029(.116)
Bio-ancestry (Asian)	−.226(.211)	−.166(.137)
Female	−.771(.089)***	−.481(.056)***
Age	−.445(.603)	−.149(.378)
Age^2	.008(.014)	.002(.009)
Parental education (below high school)	--	--
Parental education (high school or above)	.062(.148)	.020(.093)
Parental education missing	.425(.243)	.322(.152)*
Two biological parents	−.101(.095)	−.045(.059)
PVT < 90	.116(.120)	.136(.075)
PVT 90 to 110	--	--
PVT >110	.087(.109)	−.000(.068)
PVT missing	.161(.240)	.091(.151)
West	--	--
Midwest	−.080(.142)	.004(.089)
South	−.121(.140)	−.008(.088)
Northeast	.037(.158)	.050(.099)
Region missing	−.097(.676)	.103(.424)
N of persons	1422	1424

Note: The collective genetic contribution is estimated by mixed linear models. Models are fit using the genome-wide complex trait analysis (GCTA) software package developed by Yang et al. (2010).

* p≤ .05;

** p≤ .01;

*** p≤ .001 (two-tailed tests)

Table 3

The Collective Genetic Contribution of 403 SNPs to Serious Delinquency and Violence under High/Moderate-Social-Control

	Serious Delinquency				Violence			
	HMSC		LSC		HMSC		LSC	
	Collective Genetic Contribution (1)	Number of Persons (2)	Collective Genetic Contribution (3)	Number of Persons (4)	Collective Genetic Contribution (5)	Number of Persons (6)	Collective Genetic Contribution (7)	Number of Persons (8)
Parenting Factors								
Parental attachment	.002	1214	.010***	407	.004	1216	.051***	407
Parental supervision	.007	1234	.020***	367	.009	1235	.069***	369
School Factors								
School attachment	.000	1118	.031***	558	.000	1120	.041***	558
School discipline	.001	777	.053***	436	.005	779	.089***	437
Neighborhood Factors								
Education	.014	954	.019	468	.022	954	.021	470
Income	.005	943	.082***	477	.004	944	.068***	478
Unemployment	.003	925	.036***	477	.007	926	.035***	478
Single pare. rate	.003	925	.010***	482	.008	926	.004	483

(HMSC) and Low-Social-Control (LSC) Conditions

Note: For parenting factors, the collective genetic contribution is estimated by mixed linear models after controlling for bio-ancestry scores, gender, age, and age^2; for school factors, we control for bio-ancestry scores, gender, age, age^2, parents' education, family structure, PVT score, and region; for neighborhood factors, we control for bio-ancestry scores, gender, age, age^2, parents' education, family structure, and region.

* p≤ .05;

** p≤ .01;

*** p≤ .001 (Kolmogrov-Smimov test of whether the distribution of values in column 3 is greater than that in column 1, and that in column 7 is greater than that in column 5).

Table 4

Individual SNP Effects under High/Moderate-Social-Control and Low-Social-Control Conditions.

	Serious Delinquency (F Ratio)	Violence (F Ratio)
Parenting Factors		
Parental attachment	24.984***	173.957***
Parental supervision	1.777***	16.674***
School Factors		
School attachment	519085.700***	50790.600***
School discipline	702.769***	192.595***
Neighborhood Factors		
Education	.551	.290
Income	150.931***	171.582***
Unemployment	51.964***	13.916***
Single pare. rate	2.994***	.060

Note: The F ratio is the ratio of the variance of individual SNP effects under low-social-control conditions (solid lines in Figure 1) to the variance of individual SNP effects under high/moderate-social-control conditions (dashed lines in Figure 1).

* p≤ .05;

** p≤ .01;

*** p≤ .001 (F Test)

Table 5

Gene-Environment Correlation: Predict Socio-environmental Variables Using 403 SNPs

	Collective Genetic Contribution
Parenting Factors	
Parental attachment	.002
Parental supervision	.009
School Factors	
School attachment	.012
School discipline	.011
Neighborhood Factors	
High education	.010
High income	.025
Low unemployment	.031
Low single pare. rate	.022

* p≤ .05;

** p≤ .01;

*** p≤ .001 (Likelihood Ratio Test)