



The Genome-Wide Study of Human Social Behavior and Its Application in Sociology

Peter T. Tanksley¹, Ryan T. Motz¹, Rachel M. Kail¹, J. C. Barnes¹ and Hexuan Liu^{1,2*}

¹ School of Criminal Justice, University of Cincinnati, Cincinnati, OH, United States, ² Institute for Interdisciplinary Data Science, University of Cincinnati, Cincinnati, OH, United States

Recent years have seen a push for the integration of modern genomic methodologies with sociological inquiry. The inclusion of genomic approaches promises to help address long-standing issues in sociology (e.g., selection effects), as well as open up new avenues for future research. This article reviews the substantive findings of behavior genetic/genomic research, both from the recent past (e.g., twin/adoption studies, candidate gene studies) and from contemporary genomic analyses. The article primarily focuses on modern genomic methods available to sociologists (e.g., polygenic score analysis) and their various applications for answering sociological questions. The article concludes by considering a number of areas to which genomic researchers and sociologists should pay close attention if a consilience between genomic methods and sociological research is to be fully realized.

OPEN ACCESS

Edited by:

Emma Sprooten,
Radboud University, Netherlands

Reviewed by:

Magdalena Zadkowska,
University of Gdansk, Poland
Maurizio Meloni,
Deakin University, Australia

*Correspondence:

Hexuan Liu
liu2h6@ucmail.uc.edu

Specialty section:

This article was submitted to
Evolutionary Sociology and
Biosociology,
a section of the journal
Frontiers in Sociology

Received: 28 February 2019

Accepted: 07 June 2019

Published: 26 June 2019

Citation:

Tanksley PT, Motz RT, Kail RM,
Barnes JC and Liu H (2019) The
Genome-Wide Study of Human Social
Behavior and Its Application in
Sociology. *Front. Sociol.* 4:53.
doi: 10.3389/fsoc.2019.00053

Keywords: social science genomics, sociogenomics, genome-wide association, polygenic scores, heritability, candidate genes, gene-environment interaction, gene-environment correlation

INTRODUCTION

Social scientists have long been interested in understanding what role biology plays in human social traits and behaviors. With the completion of the Human Genome Project at the turn of the twenty-first century and the subsequent advent of genome-wide data and methods, researchers are addressing this question in unprecedented ways. Genome-wide analyses, and particularly the polygenic score (PGS) analysis, have already made significant contributions in the health and life sciences such as medicine and epidemiology, and even the social sciences such as psychology and economics. The use of genomic methods in sociology, albeit growing, has not yet been fully realized. In other words, we agree with Conley and Fletcher (2017, p. 11) that, "...the social genomics revolution is just getting started." This review will examine the substantive findings of research that has explored the application of genomic approaches to sociological questions.

The purpose of this article is to review the past and current findings that have bearing on sociogenomics work. We will also offer thoughts and predictions about the future direction(s) of sociogenomics research. Thus, this article begins by providing a brief description of the behavior genetic research that preceded sociogenomics, beginning with twin/adoption studies and candidate gene studies. The bulk of the paper will focus on the contemporary era and, in particular, on genome-wide association (GWA) and PGS studies, with special focus turned toward the substantive findings and possible uses of these methods in sociological research. We will then conclude by considering some of the current and future issues that sociogenomics research will need to address in the coming years.

This article hopes to encourage familiarity with sociogenomic methods among sociological researchers. We also wish to establish the importance of answering a central question in sociogenomics: what role does biology play in social traits and behaviors? On this point, we believe that the eminent biologist E.O. Wilson was remarkably prescient in his book, *Consilience* (Wilson, 1999), in outlining the necessary steps to attaining the answer to that central question of biology and behavior. His words will provide a guide and inspiration for the sections that follow.

“The clarification of norms of reaction and heritability... is the crucial first step toward unbraiding the roles of heredity and environment in human behavior...”

-E.O. Wilson, *Consilience* (Wilson, 1999, p. 155–156)

BACKGROUND

Some of the earliest work on investigating the role of biology in human social behavior began in the 1900s and used the approach known as “variance partitioning” (Tabery, 2014). The variance partitioning approach evaluates the heritability (i.e., the proportion of variance explained by genetics) of traits or behaviors, typically with the use twin and adoption studies. These family-based studies provided a natural experiment wherein the actual amount of genetic relatedness between individuals could be known (e.g., 50 and 100% for fraternal and maternal twins, respectively, and 0% for non-biological siblings). Understanding the genetic relatedness of family members allowed researchers to partition the variation in social traits and behaviors in terms of their genetic (i.e., heritability) and environmental contributions.

Overall, twin research supported the hypothesis that genes play a crucial role in the development of human traits and behaviors. This support was strong enough that Turkheimer (2000) set forth three laws of behavior genetics: (1) “all human behavior is heritable,” and (2) “the effect of being raised in the same family is smaller than the effect of genes,” but (3) “a substantial portion of the variation in complex human behavioral traits is not accounted for by the effects of genes or families” (p. 160). These laws concisely and accurately summarized the conclusions of the first era of behavior genetic research. By evaluating over 2,500 publications and over 17,000 traits, Polderman et al. (2015) demonstrated that *all* human traits are heritable. Overall, ~50% of the variation in phenotypes (i.e., observable traits or behaviors) was explained by genetic influences including characteristics such as temperament and conduct disorder. Furthermore, Polderman et al. (2015) indicated that additive genetic influence was the best explanation for 69% of traits, suggesting that twin studies were an effective and efficient method to evaluate the influence of genes on trait variance.

“The logical next step is the location of the genes that affect behavior.”

- E.O. Wilson, *Consilience* (Wilson, 1999, p. 155–156)

Toward the end of the first era of behavior genetic research, scholars aimed to go beyond simply addressing *how much* genes

matter in behavior (i.e., the variance partitioning approach), and instead sought to answer the deeper questions of *which genes* and *why*—known as the “mechanism elucidation” approach (Tabery, 2014). The second era was launched under the impression that there were only a few genetic mechanisms with large effects that, when combined, would account for much of the heritability of an outcome. Researchers came to use the “candidate gene” approach, wherein they would identify a gene of interest (i.e., the genetic “candidate”) *a priori* and conduct statistical analyses to see if any variants of the gene (i.e., alleles) were associated with the behavioral outcome of interest.

Of particular interest is the phenomenon of gene-environment ($G \times E$) interaction. Simply put, a $G \times E$ interaction suggests the effect of the environment depends on genes or, the reverse, that the effect of genes depends on the environment (Shanahan and Hofer, 2005; Rutter, 2006; Boardman et al., 2013). Two of the most groundbreaking ($G \times E$) studies using candidate genes were performed by Caspi et al. (2002, 2003). In the first study (Caspi et al., 2002), the authors investigated the interactive relationship between variants of the *MAOA* gene and childhood maltreatment on antisocial behavior in adulthood. In the second study (Caspi et al., 2003), the authors investigated the interactive relationship between variants in the *5-HTTLPR* gene and stressful life events on depression. In both studies, the authors found that the *MAOA* and the *5-HTTLPR* candidate genes alone did not predict antisocial behavior and depression, respectively. Crucially, however, when the alleles of the candidate genes were interacted with their respective environmental exposures, significant relationships emerged. These findings suggest that some behavioral traits are co-dependent on (1) the possession of genetic risk and (2) exposure to environmental stressors in order for them to manifest.

The enthusiasm for the candidate gene era lost momentum, however, as subsequent studies using the same genetic variants and same outcomes in different samples came to inconsistent conclusions and oftentimes completely failed to replicate (e.g., Chabris et al., 2012). For example, Duncan and Keller (2011) performed a thorough review of the first tens year of $G \times E$ studies using candidate genes in psychiatry and found that while 96% of novel findings were significant, they only replicated 27% of the time. A number of reasons have been presented for why candidate gene studies suffered from inconsistent findings and failure to replicate (e.g., see Tabor et al., 2002); however, one of the key failures of the second era of behavior genetic research was the assumption that only a few genetic variants played a role in complex human traits and behaviors.

“Once the genes have been mapped on the chromosomes and their pathways of expression identified, their interaction with the environment can be more precisely traced.”

- E.O. Wilson, *Consilience* (Wilson, 1999, p. 155–156)

SOCIOGENOMICS

In the past decade, a new polygenic model has been adopted in the study of complex human traits. The polygenic model assumes

complex traits are the result of “...unknown but certainly multifarious, and interactive biological and social pathways, with large numbers of genetic loci” (i.e., “weak biologism”; Turkheimer, 1998, p. 787). By the early 2010s, GWA studies provided support for the polygenic model by locating hundreds of common genetic variants across the genome that bore significant associations with complex traits such as human height, body mass, and educational attainment (Wood et al., 2014; Locke et al., 2015; Lee et al., 2018). In 2015, Chabris et al. published an article wherein they proffered a “fourth law of behavior genetics” to supplement Turkheimer’s original three. The fourth law: “A typical human behavioral trait is associated with very many genetic variants, each of which accounts for a very small percentage of the behavioral variability” (Chabris et al., 2015, p. 305).

Genome-Wide Association Studies

Early GWA studies focused on diseases (e.g., age-related macular degeneration [AMD]) and other medically-relevant physical traits like body mass index and height. Only a short time later, however, researchers were able to conduct GWA studies on complex traits that were of interest to social scientists. Some recent GWA studies have examined traits such as intelligence (Savage et al., 2018), subjective well-being (Okbay et al., 2016a), attention-deficit/hyperactivity disorder (ADHD) (Demontis et al., 2019), aggression (Pappa et al., 2016), antisocial behavior (Tielbeek et al., 2017), tobacco, and alcohol use (Liu et al., 2019), and educational attainment (Rietveld et al., 2013a; Okbay et al., 2016b; Lee et al., 2018).

The study of educational attainment provides a prototypical example for how GWA studies progress in terms of gene discovery and explanatory power. The first GWA study of educational attainment (EA1; Rietveld et al., 2013a) had a sample size of ~126,000 individuals. The study reported finding three genome-wide significant single nucleotide polymorphisms (SNPs) that explained ~2% of the variation in years of education. Three years later, EA2 (Okbay et al., 2016b) boasted a sample of nearly 300,000 individuals and succeeded in discovering 74 significant SNPs that explained ~4% of the variation in education. The most recent study, EA3 (Lee et al., 2018), used a sample of 1.1 million individuals and discovered 1,274 genome-wide significant SNPs that explained 11–13% of the variation in years of education.

Although GWA is a method for gene-discovery, the large number and small effect sizes of the genetic variants identified by GWA studies largely preclude researchers from identifying specific causal pathways. Contrary to the initial assumption that GWA studies would find a small number of highly impactful SNPs, Plomin (2018) writes that “...what [GWA studies] found was gold dust, not nuggets” (p. 187). Though small, when the effects of SNPs are combined, a significant proportion of the variance in complex traits can be predicted. Plomin (2018) continues, “[e]ach speck of gold was not worth much, but scooping up handfuls of gold dust made it possible to predict genetic propensities of individuals” (p. 187).

Polygenic Scores & Their Utility for Sociology

In the past decade, the polygenic score (PGS) approach has emerged as a useful tool in research of complex social traits and behaviors (Belsky and Israel, 2014). A PGS is compound measure that aggregates the genetic effects on a particular outcome. Essentially, once a GWA study estimates the association between all of the genetic variants in the genome and an outcome of interest, these estimates can be summed for individuals, thereby creating individualized compound measure of genetic predisposition. In the following paragraphs, we will discuss four of the most common uses of PGS analysis in social science research, including: (1) genetic mediation, (2) genetic confounding, (3) gene-environment interaction, and (4) gene-environment correlation.

Genetic Mediation

One of the most fruitful areas of PGS-based research is the examination of how genetic associations are environmentally mediated. The PGS for educational attainment provides a characteristic example of this kind of analysis. A recent study by Belsky et al. (2018) examined intergenerational mobility in five longitudinal cohorts from three counties. The aim of the study was to test if higher PGSs for educational attainment translated into greater intergenerational mobility (i.e., greater educational and occupational outcomes than one’s parents). One of the key findings was that the PGS of mothers predicted their children’s educational attainment, even after controlling for the offspring’s own PGS. This finding indicates that maternal genes operate through environmental pathways in addition to genetic pathways (i.e., through genetic transmission). Interestingly, environmental mediation has been demonstrated for both transmitted and non-transmitted genetic material (i.e., genetic nurture). For instance, Kong et al. (2018) examined the effects of parental genetic material that was not passed down to their offspring and found that non-transmitted genetic material exerted a genetic effect on educational attainment that was about one third of the size of the child’s own PGS. This finding underscores the significant role of environmental mediation in genetic associations as non-transmitted genetic material can only operate through the environmental means.

Phenotypes do not exist in a vacuum: rather, a phenotype like educational attainment is best thought of as being embedded in an interactive network of associations with other phenotypes. Considering the many intermediate and subsequent phenotypes that would result from a genetic propensity for education, studies have shown that a higher PGS for educational attainment also predicts greater social mobility (Belsky et al., 2018; Liu, 2018), labor market earning (Papageorge and Thom, 2018), parental investment (Wertz et al., 2018), and cognitive performance (Lee et al., 2018).

Genetic Confounding

A myriad of sociological studies have explored how social factors influence individual behaviors and outcomes. Due to the non-experimental nature of most sociological data, however, selection is a source of constant concern. One important source of selection

is human genetics; because roughly all of human behavior is heritable (i.e., influenced by genes; see Polderman et al., 2015), any association between an environmental factor (e.g., delinquent peers) and outcome (e.g., delinquent behavior) *may* be partially attributable to individual's specific genetic markup (Barnes et al., 2014). Advances in modern GWA studies and PGS analysis provide sociologists the opportunity to address some of the unexplained heterogeneity in past research and correct for this kind of selection.

As an example, intergenerational transmission of educational attainment is a central theme in sociological research. One critical challenge in this research is that parents and children share both their living environment and half of their DNA. Because of that, the observed parent-child association in education is both social and genetic. Without considering the genetic influence, an estimate of the socio-environmental influence is likely biased. Using a polygenic score based on the study of Rietveld et al. (2013a), Conley et al. (2015) estimated the genetic confounding effect in the intergenerational transmission of education. They found that genetic factors account for about one-sixth of the observed parent-child association in education. Liu (2018) revisited the issue by using a better powered polygenic score on educational attainment constructed based on the study of Okbay et al. (2016b). His results showed that around one-fifth of intergenerational transmission was accounted for by genetics. These analyses demonstrate not only how genetic contributions to social outcomes can be quantified and controlled for, but also how the inclusion of PGSs can help examine and rule out possible sources of genetic confounding.

Gene-Environment Interaction

Arguably one of the most interesting uses of PGS analysis for sociologists is the investigation of gene-environment ($G \times E$) interactions. It is useful to think of $G \times E$ interactions as functioning two ways. First, genes may be thought of as moderators of environmental associations, wherein the effects of an environmental exposure are conditional upon an individual's genetic makeup. For example, Guo et al. (2015a) hypothesized that, for college students, the effect of peer influence on binge drinking behavior may vary depending on a student's genetic propensity for alcohol consumption. Using data from the College Roommate Study (ROOM), the authors found that only college students with a moderate level of genetic propensity for consuming alcohol were significantly influenced by a roommate who drank—students of high and low levels of genetic liability were not affected by the influence of peers. While traditional sociological perspectives (e.g., social learning theory) might have predicted a positive and monotonic relationship between peer influence and binge drinking, the inclusion of biologically relevant variables allowed for the identification of a pattern otherwise inexplicable by pure sociological means. In another study, Perry (2016) found evidence for a three-way $G \times E$ interaction model, namely higher levels of social integration predicted reduced risk of nicotine dependent among men with high genetic susceptibility to nicotine cravings, yet the protective effect of social integration is substantially reduced among women.

A second way to view $G \times E$ interactions is to think of environmental exposures moderating the effects of genetic propensities on later outcomes. Schmitz and Conley (2016), for example, examined the effect of the $G \times E$ interaction of a PGS for smoking and Vietnam-era military service on later smoking behavior and health outcomes. Using data from the Health and Retirement Study (HRS), the authors found that veterans (i.e., compared to non-veterans) who possessed PGSs that put them at high risk for tobacco use (i.e., 1–2 standard deviations above the mean) were 57–71% more prone to be smokers later in life and smoke 18–27 more cigarettes per day. Interestingly, rates of smoking later in life did not substantively differ across veterans and non-veterans if they possessed low-medium levels of genetic risk for smoking. This suggests that the environmental exposure of military service only acted as a moderator for individuals who already possessed high genetic predisposition for tobacco use. This again demonstrates how $G \times E$ interactions may be employed to more fully explain patterns of unobserved heterogeneity in sociological research.

Gene-Environment Correlation

In addition to $G \times E$ interactions, gene-environment correlations (rGE) has increasingly drawn the attention of social scientists. A rGE occurs when an individual's environmental exposure depends on his/her genotype (Plomin et al., 1977; Scarr and McCartney, 1983; Jaffee and Price, 2007; Fletcher and Conley, 2013; Wagner et al., 2013). For instance, Christakis and Fowler (2014) studied the friendship networks of youths in the National Longitudinal Adolescent to Adult Health (Add Health) Study and found a number of significant correlations between members of friend groups at the genetic level. In particular, they found that friends were significantly more homogenous on genes for olfactory function (i.e., their sense of smell) and heterogenous for immune function genes (i.e., compared to random non-friends). This finding provides significant nuance to social network research in that it suggests individuals may sort themselves into groups based on social homophily (e.g., they prefer the same smells) as well as genetic complementarity (e.g., diversity of genetic immune function in groups may help prevent the spread of disease).

Determining the strength of rGE is an ongoing area of research. For example, Domingue et al. (2018), also using the Add Health, found that friends showed significant signs of homophily. Friends were more similar (compared to random non-friends in the Add Health) in terms of their genetics for BMI and educational attainment. However, after controlling for genetic similarity at the school level (i.e., keeping comparisons within schools), the overall genome-wide similarity among friends was halved, as was the correlation among friends for the educational attainment PGS (the correlation for BMI was relatively unchanged). This finding suggests that social sources of stratification may be responsible for some amount of rGE occurring in society. This study also found that, despite the fact that assortment at the social level accounted for much of the genetic similarity of friends, substantive social-genetic effects still occurred among friend groups. For example, friends' genetics

still significantly predicted the educational attainment of group members, even after controlling for genetic similarity.

Genomic-Relatedness-Matrix Restricted Maximum Likelihood Estimation (GREML)

Genomic-relatedness-matrix restricted maximum likelihood estimation is another innovative approach that handles large-scale genotype data. The basic idea of GREML is to estimate a matrix of genetic relationships among individuals who are not socially related in the sense of being family members. Based on the matrix, the genomic contribution (i.e., collective influence of all genetic variants across the whole genome) can be estimated. This method has also been used to estimate the genomic contribution to human height (Yang et al., 2010), BMI (Yang et al., 2011), schizophrenia (Lee et al., 2012). Social scientists have employed it to investigate the genetic architecture of intelligence (Davies et al., 2011; Chabris et al., 2012), personality traits (Vinkhuyzen et al., 2012), subjective well-being (Rietveld et al., 2013b), and economic and political preferences (Benjamin et al., 2012).

GREML has also been utilized in $G \times E$ interaction and rGE research. Guo et al. (2015b), for example, conducted a GREML analysis to test if genomic influence on BMI differs by historical period. They found that the genomic influence on BMI was substantially and significantly larger during the obesity epidemic than before in the United States. Using GREML methods, other studies found that the genetic influence was greater among adolescents who lived in adverse social conditions than those in favorable social conditions (Li et al., 2015; Liu et al., 2015). Deary et al. (2012) conducted a bivariate GREML analysis of intelligence at different life stages. As a result, they found a fairly large significant genetic correlation between intelligence in adolescence (age 11) and in late adulthood (age 65–78). In another study, Boardman et al. (2015) estimated genetic correlations among education, BMI, depression, and self-rated health. They showed that observed correlations between education and depression and between education and self-rated health were largely attributable to common genetic factors.

“When many such interactions have been defined, the whole can be braided back again to attempt a more complete picture of mental development.”

—E.O. Wilson, *Consilience* (Wilson, 1999, p. 155–156)

FUTURE HURDLES AND PROMISING AVENUES

This paper has presented an overview of the progress of sociogenomics research, from heritability, and twin studies to candidate genes and finally to the modern era of sociogenomic research, and along the way we have highlighted some of the ways social science researchers have used these techniques to reach new insights. In this final section, we will present future hurdles and promising avenues that we believe are facing sociogenomic research.

Missing Heritability

Heritability (i.e., the trait variation explainable by genetics) can be estimated from both twin/adoption models and genomic methods; however, family-based trait heritabilities have often dwarfed those estimated from genomic data. For example, family-based heritability of height (a highly biological trait) has been estimated to be as high as 0.80 (i.e., 80% genetic), while SNP heritability estimates have, until recently, only been estimated at less than half that level (Manolio et al., 2009). The discordance between family- and SNP-based estimates has been dubbed the problem of “missing heritability.” The significance of missing heritability for sociogenomics, and genomic research more broadly, is that it signals an incompleteness, either in the method (i.e., GWA and GREML) or in understanding the underlying genetics. A number of possible explanations of missing heritability have been posited, including the inability of GWA analysis to detect the effects of rare genetic variants (i.e., GWA is designed to focus on common variants) or to account for structural variants (i.e., DNA sequences that vary in their number, location, or orientation) (Manolio et al., 2009; see also Eichler et al., 2010).

Although the problem of missing heritability has persisted since the advent of genomic research on human traits and behaviors, recent developments in genomic methods suggest that missing heritability may indeed be recoverable. Wainschtein et al. (2019), using whole-genome data and a variation on the GREML method (discussed above), estimated the heritability of height in a sample of 21,620 unrelated individuals to be 79% (i.e., an estimate consistent with previous family-based estimates). The authors concluded that most of the missing heritability in height was the product of rare variants, which are not detectable using standard genomic data and methods. The ability to assess the contribution of rare variants has begun to close the heritability gap for highly biological traits like height; however, it is possible that different genetic mechanisms are responsible for the missing heritability in more complex social traits such as educational attainment (see Eichler et al., 2010). Time alone will tell if rare variants, or some other heretofore unassessed mechanism, can explain what drives the missing heritability of complex social traits.

Population Stratification

Although genomic methods have important contributions to make to a wide field of sociological inquiry, current methods are subject to a significant methodological constraint that limits their application to individuals of genetically diverse populations. Modern genomic methodologies (e.g., GWA and PGS studies) are largely limited to individuals with European ancestry due to an evolutionary phenomenon known as population stratification. Population stratification occurs when systematic differences in allele frequencies (i.e., the prevalence of particular versions genes) exist between two subgroups within a larger population. These systematic differences between groups can be used to identify an individual’s genetic ancestry; however, if a particular phenotype also breaks along ancestral lines (e.g., chopstick use in Asian countries), population stratification can lead researchers to assume that these systematic differences are

causally related to the phenotype itself (e.g., the “chopstick gene”; Hamer and Sirota, 2000).

Due to the potential bias introduced by population stratification, most GWA studies limit their samples to individuals of European descent. Restriction to European populations occurs for, at least, two reasons: (1) genome-wide data on non-European populations is rare (see Popejoy and Fullerton, 2016) and (2) GWA studies are highly dependent on large sample sizes in order to boost the signal of causal SNPs above the genomic noise (for a discussion on GWA studies sample sizes and discovery rates see Visscher et al., 2017). The consequence of these facts is that few GWA studies have assessed non-European samples. Duncan et al. (2018), for example, reported that exclusively European samples were used in 67% of all GWA studies, Asian samples in 19%, and African samples in only 3.8%. The concentration of GWA studies on European samples is especially problematic with regard to PGS construction: a PGS derived from European-based summary statistics may not be reliably applied to individuals from non-European populations.

Duncan et al. (2018) used European-based summary statistics on height, BMI, and schizophrenia to construct PGSs for samples of African and European descent. When comparing the effect sizes, the authors observed that the PGSs for the African ancestry sample were only around 36% as large as the effect sizes for the European sample. Part of the reason European-based PGSs do not function well for individuals of African descent is that their genetic architecture is not wholly overlapping. Sherman et al. (2019) compared the fully sequenced genomes of 910 individuals of African descent to the reference human genome and found that the African genomes possess around 10% more DNA than does the current human reference genome (the authors did note that the functional significance of these additional regions of DNA are still unknown).

Together, the paucity of non-European based GWA studies and the limited transferability of PGSs to non-European samples means that modern genomic research is (perhaps unavoidably) ignoring numerous segments of the population. These limitations suggest that genomic methodologies may produce a new form of inequality (i.e., inequality in genomic research) between groups of different ancestry in the population (Bustamante et al., 2011; Popejoy and Fullerton, 2016; West et al., 2017). This issue not only has implications for sociological research endeavors that seek to use genomic methods, but also for work being done in fields like medicine and public health (Petrovski and Goldstein, 2016).

Variance Quantitative Trait Loci

One promising avenue of genomic research for identifying $G \times E$ interactions are the current efforts to identify variance quantitative trait loci (vQTLs). As with most regression-based models, GWA studies are designed to discover SNPs that are associated with the mean level of a phenotype; vQTLs, on the contrary, are SNPs that are associated with the variance in a phenotype's expression. Although identifying variance effects across the genome has proven difficult (see Yang et al., 2012; Conley et al., 2018), a recent variance GWA (vGWA) study

was conducted using ~350,000 individuals in the UK Biobank that identified 75 genome-wide significant vQTLs associated with variability in body mass index (BMI) (Wang et al., 2019). What is more, these vQTLs were shown to produce significant $G \times E$ interactions with five BMI-related environmental factors (i.e., sex, age, physical activity, sedentary behavior, and smoking).

Developments in the area of vQTLs are exciting because they offer researchers the opportunity to examine certain models of interaction that do not readily avail themselves to empirical tests. For instance, researchers have been particularly interested in the differential susceptibility hypothesis of $G \times E$ interaction (see Belsky and Pluess, 2009) that states that some individuals are more susceptible to their environment (i.e., doing worse in bad environments and better in supportive ones) compared to others. This model of $G \times E$ interaction has large implications because it suggests that improving environmental conditions or relocating individuals to more supportive environments could produce disproportionately positive behavioral responses in some. Yet support for the hypothesis has usually been only partial in nature due to the high level of difficulty in conducting a comprehensive test (i.e., a full test would require an interaction analysis of a plasticity factor [i.e., typically a candidate gene] and an environmental exposure that ranges from highly negative to highly positive). With the advent of vQTL research, however, a convincing test of the differential susceptibility hypothesis that fully leverages modern genomic methods may have finally been made a realistic possibility. vQTL research is still in its infancy, so researchers will need to wait some time until vQTLs for complex traits like educational attainment are identified and made available.

Phenotypic Annotation

The aim of GWA studies is to estimate the GWA with a specific phenotype; however, the estimates produced by GWA studies represent the average effects of SNPs nested within individuals, all of whom have unique sets of environmental exposures. GWA studies results, and the PGSs derived from them, do not possess information about the mechanistic pathways through which genes exert their influence. This is one area in which sociologists can help improve the state of the research; the rich theoretical traditions of social causation in sociology provide sociogenomics research with a map of environmental associations that are likely to have biological underpinnings, that may mediate the genetic associations, and that could moderate them (Freese, 2018).

The process of mapping the interconnections between associated phenotypes has recently been termed “phenotypic annotation” (Belsky and Harden, 2019), a phrase that mirrors “genotypic annotation” (i.e., the process of mapping biological functions onto sets of specific genes). Phenotypic annotation is one area in which sociologists, in particular, have significant and substantive contributions to make to sociogenomics research. Relying on their wealth of past theory and research, sociologists can inform the phenotypic annotation of socioenvironmental networks that have been established empirically. By using this knowledge, sociogenomics research can proceed to test how far flung genetic effects disperse along the networks of phenotypic associations.

Epigenetics

Epigenetics is the study of gene regulation (i.e., which genes are read and used for protein synthesis), and one of the most exciting findings to emerge from modern genomics research is that the epigenome is responsive to the environment (e.g., Pembrey et al., 2006). Thus, social scientists have been especially excited to incorporate epigenetics into their work because it provides an opportunity to show how environments might “get under the skin” (Landecker and Panofsky, 2013; Meloni, 2014; Lock, 2015; Harris and McDade, 2018; Meloni et al., 2018). To put it briefly—which will require us to simplify and skip important details (see for more details Deichmann, 2016)—epigenetics is the study of how biological mechanisms can turn genes “on” and “off.” If a gene is turned “on,” it is available to be transcribed and translated, meaning it can have an impact on protein synthesis. If a gene is turned “off,” that gene has essentially been silenced such that it no longer has an impact.

While a thorough discussion of biological mechanisms that determine how genes are regulated is beyond the scope of this paper (but interested readers are encouraged to see National Institute of Health, 2019; see also Szyf and Bick, 2013), what is important for our purposes is that social and environmental factors likely impact the epigenetic process and, thus, it may be possible to identify those relationships through empirical study. For example, imagine two identical twins. On day one of their life, the two twins will be identical in their genome and their epigenome (Fraga et al., 2005). But, as their lives go on and they begin to accumulate unique experiences, their epigenomes will start to diverge. The twin who is exposed to a traumatic event will have certain genes turned on/off that are not affected in the unexposed twin. This is how our environments might get under our skin.

It is probably obvious, at this point, why this area of research has received so much hype (Deichmann, 2016). What may not be as obvious, though, are the theoretical, methodological, and statistical challenges that are involved with epigenetics research. Let us start by recalling one of the complications that has emerged from GWA studies. Specifically, low levels of statistical power exist for identifying any given genetic signal. Now, consider that epigenetic mechanisms work to regulate genes. If the signal from the gene is weak and statistically difficult to identify, then the signal for the epigenome is likely to be even weaker and even more difficult to statistically identify. Epigenetic research thus requires large sample sizes in order to identify reliable signals. Although we are hesitant to say what “large” means, we feel it is safe to say that epigenetics research will likely require sample sizes that are even larger than those required to reliably identify genetic signals with GWA methods.

Nonetheless, there several studies that have drawn on epigenetic data in social science research. One such study was conducted by Beach et al. (2013). Using a sample of $N = 155$ women, these authors found that exposure to childhood sexual abuse was associated with epigenetic change. Moreover, the epigenetic differences appeared to mediate the effect of childhood sexual abuse on later-in-life antisocial behavior. In another example, Lei et al. (2015) relied on a sample of $N = 99$ women and found evidence to suggest that exposure to high-crime neighborhoods had an effect on methylation patterns

(i.e., epigenetic markers) of the promoter region of the 5-*HTT* candidate gene and that this gene was linked to later depression risk. Taken together, these studies suggest there may be something to the idea that our social environments can indeed get under our skin and impact human behavior in ways that challenge the traditional view that biological and social factors can be considered separate spheres of influence. But low statistical power may be a concern.

Also, more recent evidence complicates this narrative by drawing attention to several practical and methodological challenges. Sugden et al. (2019) uncovered an intricate relationship between epigenetic markers and tobacco smoking. Of course, it has long been recognized that tobacco smoking leads to biological changes, some of which may ultimately result in outcomes such as increased risk of cancer. But there are other, more subtle changes, that occur when one smokes tobacco. Sugden et al. (2019) found evidence that changes in smoking behaviors were linked to changes in epigenetic markers across the human genome. Thus, researchers hoping to study the link between an environmental exposure and epigenetic modifications should consider taking into account the participants’ smoking behaviors. If smoking is ignored, it may confound any epigenetic signals that are identified.

Finally, a similar word of caution was expressed by another group of researchers (Marzi et al., 2018) who drew on data from a longitudinal cohort of youth from the United Kingdom ($N = 2,232$). This research team tested for an association between personal victimization experiences and epigenetic markers. They conducted an epigenome-wide analysis (EWA) that attempted to differentiate participants who had experienced victimization and those who had not. If the idea that our environments can become “embedded” and “get under our skin” is right, then we should expect to see epigenetic markers for extremely stressful environmental exposures like victimization. But their analysis failed to identify any epigenetic markers among victims. Marzi et al. (2018, p. 517) concluded with a recommendation that we believe is worth repeating here: “We need to come to terms with the possibility that epigenetic epidemiology is not yet well-matched to experimental, non-human models in uncovering the biological embedding of stress.”

CONCLUSION

In his book, *Consilience*, E.O. Wilson envisioned a unification of all branches of knowledge including biology, the humanities, and the social sciences. Likewise, our object in this paper has been the consilience of modern genomic research with the discipline of sociology. Sociogenomics research has experienced its own “jumping together”—the literal meaning of “consilience”—of its two historical approaches perspectives: variance partitioning and mechanism elucidation (Tabery, 2014). Modern genomic methods consist of the construction of variance partitioning polygenic scores that are themselves derived from the mechanism-elucidation process of gene-discovery through GWA studies. In the words of Wilson (1999, p. 12), “[w]e are approaching a new age of synthesis, when the testing of consilience is the greatest of

all intellectual challenges.” With the advent and growing robustness of genomic methodologies, sociologists are in an enviable position to adopt these tools and integrate them into their research.

REFERENCES

- Barnes, J. C., Boutwell, B. B., Beaver, K. M., Gibson, C. L., and Wright, J. P. (2014). On the consequences of ignoring genetic influences in criminological research. *J. Crim. Justice* 42, 471–482. doi: 10.1016/j.jcrimjus.2014.08.003
- Beach, S. R. H., Brody, G. H., Lei, M. K., Gibbons, F. X., Gerrard, M., Cutrona, C. E., et al. (2013). Impact of child sex abuse on adult psychopathology: a genetically and epigenetically informed investigation. *J. Fam. Psychol.* 27, 3–11. doi: 10.1037/a0031459
- Belsky, D. W., Domingue, B. W., Wedow, R., Arseneault, L., Boardman, J. D., Caspi, A., et al. (2018). Genetic analysis of social-class mobility in five longitudinal studies. *Proc. Natl. Acad. Sci. U. S. A.* 115, E7275–E7284. doi: 10.1073/pnas.1801238115
- Belsky, D. W., and Harden, K. P. (2019). Phenotypic annotation: using polygenic scores to translate discoveries from genome-wide association studies from the top down. *Curr. Direct. Psychol. Sci.* 28, 82–90. doi: 10.1177/0963721418807729
- Belsky, D. W., and Israel, S. (2014). Integrating genetics and social science: genetic risk scores. *Biodemo. Soc. Biol.* 60, 137–155. doi: 10.1080/19485565.2014.946591
- Belsky, J., and Pluess, M. (2009). Beyond diathesis stress: differential susceptibility to environmental influences. *Psychol. Bull.* 135, 885–908. doi: 10.1037/a0017376
- Benjamin, D. J., Cesarini, D., van der Loos, M. J., Dawes, C. T., Koellinger, P. D., Magnusson, P. K., et al. (2012). The genetic architecture of economic and political preferences. *Proc. Natl. Acad. Sci. U. S. A.* 109, 8026–8031. doi: 10.1073/pnas.1120666109
- Boardman, J. D., Daw, J., and Freese, J. (2013). Defining the environment in gene-environment research: lessons from social epidemiology. *Am. J. Public Health* 103, S64–S72. doi: 10.2105/AJPH.2013.301355
- Boardman, J. D., Domingue, B. W., and Daw, J. (2015). What can genes tell us about the relationship between education and health? *Soc. Sci. Med.* 127, 171–180. doi: 10.1016/j.socscimed.2014.08.001
- Bustamante, C. D., Burchard, E. G., and De La Vega, F. M. (2011). Genomics for the world: medical genomics has focused almost entirely on those of European descent. Other ethnic groups must be studied to ensure that more people benefit. *Nature* 475, 163–165. doi: 10.1038/475163a
- Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science* 297, 851–854. doi: 10.1126/science.1072290
- Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301, 386–389. doi: 10.1126/science.1083968
- Chabris, C. F., Hebert, B. M., Benjamin, D. J., Beauchamp, J., Cesarini, D., van der Loos, M., et al. (2012). Most reported genetic associations with general intelligence are probably false positives. *Psychol. Sci.* 23, 1314–1323. doi: 10.1177/0956797611435528
- Chabris, C. F., Lee, J. J., Cesarini, D., Benjamin, D. J., and Laibson, D. I. (2015). The fourth law of behavior genetics. *Curr. Dir. Psychol. Sci.* 24, 304–312. doi: 10.1177/0963721415580430
- Christakis, N. A., and Fowler, J. H. (2014). Friendship and natural selection. *Proc. Natl. Acad. Sci. U. S. A.* 111(Suppl. 3), 10796–10801. doi: 10.1073/pnas.1400825111
- Conley, D., Domingue, B. W., Cesarini, D., Dawes, C., Rietveld, C. A., and Boardman, J. D. (2015). Is the effect of parental education on offspring biased or moderated by genotype? *Sociol. Sci.* 2, 82–105. doi: 10.15195/v2.a6
- Conley, D., and Fletcher, J. M. (2017). The genome factor: What the social genomics revolution reveals about ourselves, our history, and the future. *NJ: Princeton University Press.* doi: 10.1515/9781400883240
- Conley, D., Johnson, R., Domingue, B., Dawes, C., Boardman, J., Siegal, M. L. (2018). A sibling method for identifying vQTLs. *PLoS ONE* 13:e0194541. doi: 10.1371/journal.pone.0194541
- Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S. E., Liewald, D., et al. (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol. Psychiatry* 16, 996–1005. doi: 10.1038/mp.2011.85
- Deary, I. J., Yang, J., Davies, G., Harris, S. E., Tenesa, A., Liewald, D., et al. (2012). Genetic contributions to stability and change in intelligence from childhood to old age. *Nature* 482, 212–215. doi: 10.1038/nature10781
- Deichmann, U. (2016). Epigenetics: The origins and evolution of a fashionable topic. *Developmental Biology* 416, 249–254. doi: 10.1016/j.ydbio.2016.06.005
- Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., et al. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* 51, 63–63. doi: 10.1038/s41588-018-0269-7
- Domingue, B. W., Belsky, D. W., Fletcher, J. M., Conley, D., Boardman, J. D., and Harris, K. M. (2018). The social genome of friends and schoolmates in the national longitudinal study of adolescent to adult health. *Proc. Natl. Acad. Sci. U. S. A.* 115, 702–707. doi: 10.1073/pnas.1711803115
- Duncan, L., Shen, H., Gelaye, B., Ressler, K., Feldman, M., Peterson, R. E., et al. (2018). Analysis of polygenic score usage and performance in diverse human populations. *BioRxiv [Preprint]*. doi: 10.1101/398396
- Duncan, L. E., and Keller, M. C. (2011). A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am. J. Psychiatry* 168, 1041–1049. doi: 10.1176/appi.ajp.2011.11020191
- Eichler, E. E., Flint, J., Gibson, G., Kong, A., Leal, S. M., Moore, J. H., et al. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genetics* 11, 446–450. doi: 10.1038/nrg2809
- Fletcher, J. M., and Conley, D. (2013). The challenge of causal inference in gene-environment interaction research: leveraging research designs from the social sciences. *Am. J. Pub. Health* 103, S42–S45. doi: 10.2105/AJPH.2013.301290
- Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. U. S. A.* 102, 10604–10609. doi: 10.1073/pnas.0500398102
- Freese, J. (2018). The arrival of social science genomics. *Contemp. Sociol. J. Rev.* 47, 524–536. doi: 10.1177/0094306118792214a
- Guo, G., Li, Y., Wang, H., Cai, T., and Duncan, G. J. (2015a). Peer influence, genetic propensity, and binge drinking: A natural experiment and a replication. *Am. J. Sociol.* 121, 914–954. doi: 10.1086/683224
- Guo, G., Liu, H., Wang, L., Shen, H., and Hu, W. (2015b). The genome-wide influence on human bmi depends on physical activity, life-course, and historical period. *Demography* 52, 1651–1670. doi: 10.1007/s13524-015-0421-2
- Hamer, D. H., and Sirota, L. (2000). Beware the chopsticks gene. *Mol. Psychiatry* 5, 11–13. doi: 10.1038/sj.mp.4000662
- Harris, K. M., and McDade, T. W. (2018). The biosocial approach to human development, behavior, and health across the life course. *RSF J. Social Sci.* 4, 2–26. doi: 10.7758/RSF.2018.4.4.01
- Jaffee, S. R., and Price, T. S. (2007). Gene-environment correlations: a review of the evidence and implications for prevention of mental illness. *Mol. Psychiatry* 12, 432–442. doi: 10.1038/sj.mp.4001950
- Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsdottir, B. J., Young, A. I., Thorgeirsson, T. E., et al. (2018). The nature of nurture: Effects of parental genotypes. *Science* 359, 424–428. doi: 10.1126/science.aan6877
- Landecker, H., and Panofsky, A. (2013). From social structure to gene regulation, and back: A critical introduction to environmental epigenetics for sociology. *Ann. Rev. Sociol.* 39, 333–357. doi: 10.1146/annurev-soc-071312-145707
- Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghazian, O., Zacher, M., et al. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* 50, 1112–1121. doi: 10.1038/s41588-018-0147-3
- Lee, S. H., DeCandia, T. R., Ripke, S., Yang, J., Sullivan, P. F., Goddard, M. E., et al. (2012). Estimating the proportion of variation in susceptibility to schizophrenia captured by common snps. *Nat. Genet.* 44, 247–235. doi: 10.1038/ng.1108

AUTHOR CONTRIBUTIONS

HL, JB, and PT designed the paper. PT, RM, RK, JB, and HL wrote the paper.

- Lei, M. K., Beach, S. R., Simons, R. L., and Philibert, R. A. (2015). Neighborhood crime and depressive symptoms among african american women: genetic moderation and epigenetic mediation of effects. *Social Sci. Med.* 146, 120–128. doi: 10.1016/j.socscimed.2015.10.035
- Li, Y., Liu, H., and Guo, G. (2015). Does marriage moderate genetic effects on delinquency and violence? *J. Marr. Fam.* 77, 1217–1233. doi: 10.1111/jomf.12208
- Liu, H. (2018). Social and genetic pathways in multigenerational transmission of educational attainment. *Am. Sociol. Rev.* 83, 278–304. doi: 10.1177/0003122418759651
- Liu, H., Li, Y., and Guo, G. (2015). Gene by social-environment interaction for youth delinquency and violence: Thirty-nine aggression-related genes. *Soc. Forces* 93, 881–903. doi: 10.1093/sf/sou086
- Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D. M., Chen, F., et al. (2019). Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* 51, 237–244. doi: 10.1038/s41588-018-0307-5
- Lock, M. (2015). Comprehending the body in the era of the epigenome. *Curr. Anthropol.* 56, 151–177. doi: 10.1086/680350
- Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., Day, F. R., et al. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature* 518, 197–U401 doi: 10.1038/nature14177
- Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorf, L. A., Hunter, D. J., et al. (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747–753 doi: 10.1038/nature08494
- Marzi, S. J., Sugden, K., Arseneault, L., Belsky, D. W., Burrage, J., Corcoran, D. L., et al. (2018). Analysis of DNA methylation in young people: Limited evidence for an association between victimization stress and epigenetic variation in blood. *Am. J. Psychiatry.* 175, 517–529 doi: 10.1176/appi.ajp.2017.17060693
- Meloni, M. (2014). The social brain meets the reactive genome: Neuroscience, epigenetics and the new social biology. *Front. Hum. Neurosci.* 8:309. doi: 10.3389/fnhum.2014.00309
- Meloni, M., Cromby, J., Fitzgerald, D., and Lloyd, S. (eds.). (2018). *The Palgrave Handbook of Biology and Society*. London, UK: Palgrave Macmillan. doi: 10.1057/978-1-137-52879-7
- National Institute of Health (2019). *How Genes Work*. Available online at: <https://ghr.nlm.nih.gov/primer/howgeneswork/epigenome> (accessed February 27, 2019).
- Okbay, A., Baselmans, B. M. L., Rueedi, R., Svento, R., Kolcic, I., Fontana, M. A., et al. (2016a). Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat. Genet.* 48, 624–624. doi: 10.1038/ng.3552
- Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., et al. (2016b). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* 533, 539–542. doi: 10.1038/nature17671
- Papageorge, N. W., and Thom, K. (2018). *Genes, Education, and Labor Market Outcomes: Evidence From the Health and Retirement Study (No. w25114)*. Rochester, NY: National Bureau of Economic Research.
- Pappa, I., St. Pourcain, B., Benke, K., Cavadino, A., Hakulinen, C., Nivard, M. G., et al. (2016). A genome-wide approach to children's aggressive behavior: The eagle consortium. *Am. J. Med. Genet. B* 171, 562–572. doi: 10.1002/ajmg.b.32333
- Pembrey, M. E., Bygren, L. O., Kaati, G., Edvinsson, S., Northstone, K., Sjöström, M., et al. (2006). Sex-specific, male-line transgenerational responses in humans. *Euro. J. Hum. Genet.* 14, 159–166. doi: 10.1038/sj.ejhg.5201538
- Perry, B. L. (2016). Gendering genetics: social and biological contingencies in the protective effects of social integration for men and women. *Am. J. Sociol.* 121, 1655–1696. doi: 10.1086/685486
- Petrovski, S., and Goldstein, D. B. (2016). Unequal representation of genetic variation across ancestry groups creates healthcare inequality in the application of precision medicine. *Genome Biol.* 17:157. doi: 10.1186/s13059-016-1016-y
- Plomin, R. (2018). *Blueprint: How DNA Makes us Who We Are*. Cambridge, MA: MIT Press.
- Plomin, R., DeFries, J. C., and Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychol. Bull.* 84, 309–322. doi: 10.1037/0033-2909.84.2.309
- Polderman, T. J., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., et al. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* 47, 702–709. doi: 10.1038/ng.3285
- Popejoy, A. B., and Fullerton, S. M. (2016). Genomics is failing on diversity. *Nat. News* 538, 161–164. doi: 10.1038/538161a
- Rietveld, C. A., Cesarini, D., Benjamin, D. J., Koellinger, P. D., De Neve, J. E., Tiemeier, H., et al. (2013b). Molecular genetics and subjective well-being. *Proc. Natl. Acad. Sci.* 110, 9692–9697. doi: 10.1073/pnas.1222171110
- Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., et al. (2013a). Gwas of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* 340, 1467–1471. doi: 10.1126/science.1235488
- Rutter, M. (2006). *Genes and Behavior: Nature-Nurture Interplay Explained*. Hoboken, NJ: Blackwell Publishing.
- Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C. A., et al. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* 50:912. doi: 10.1038/s41588-018-0152-6
- Scarr, S., and McCartney, K. (1983). How people make their own environments: a theory of genotype → environment effects. *Child Dev.* 54, 424–435. doi: 10.1111/j.1467-8624.1983.tb03884.x
- Schmitz, L., and Conley, D. (2016). The long-term consequences of Vietnam-era conscription and genotype on smoking behavior and health. *Behavior genetics* 46, 43–58. doi: 10.1007/s10519-015-9739-1
- Shanahan, M. J., and Hofer, S. M. (2005). Social context in gene-environment interactions: Retrospect and prospect. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 60, 65–76. doi: 10.1093/geronb/60.Special_Issue_1.65
- Sherman, R. M., J., Forman, J., Antonescu, V., Puiu, D., Daya, M., Rafaels, N., et al. (2019). Assembly of a pan-genome from deep sequencing of 910 humans of african descent. *Nat. Genet.* 51:30. doi: 10.1038/s41588-018-0273-y
- Sugden, K., Hannon, E. J., Arseneault, L., Belsky, D. W., Broadbent, J. M., Corcoran, D. L., et al. (2019). Establishing a generalized polyepigenetic biomarker for tobacco smoking. *Transl. Psychiatry* 9, 1–12. doi: 10.1038/s41398-019-0430-9
- Szyf, M., and Bick, J. (2013). DNA methylation: a mechanism for embedding early life experiences in the genome. *Child Dev.* 84, 49–57. doi: 10.1111/j.1467-8624.2012.01793.x
- Tabery, J. (2014). *Beyond Versus: The Struggle to Understand the Interaction of Nature and Nurture*. Cambridge: MIT Press. doi: 10.7551/mitpress/9780262027373.001.0001
- Tabor, H. K., Risch, N. J., and Myers, R. M. (2002). Candidate-gene approaches for studying complex genetic traits: Practical considerations. *Nat. Rev. Genet.* 3, 391–397. doi: 10.1038/nrg796
- Tielbeek, J. J., Johansson, A., Polderman, T. J. C., Rautiainen, M. R., Jansen, P., Taylor, M., et al. (2017). Genome-wide association studies of a broad spectrum of antisocial behavior. *JAMA Psychiatry* 74, 1242–1250. doi: 10.1001/jamapsychiatry.2017.3069
- Turkheimer, E. (1998). Heritability and biological explanation. *Psychol. Rev.* 105, 782. doi: 10.1037/0033-295X.105.4.782-791
- Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Curr. Direct. Psychol. Sci.* 9, 160–164. doi: 10.1111/1467-8721.00084
- Vinkhuyzen, A. A., Pedersen, N. L., Yang, J., Lee, S. H., Magnusson, P. K., Iacono, W. G., et al. (2012). Common snps explain some of the variation in the personality dimensions of neuroticism and extraversion. *Transl. Psychiatry* 2:e102. doi: 10.1038/tp.2012.27
- Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., et al. (2017). 10 years of GWAS discovery: biology, function, and translation. *Am. J. Hum. Genet.* 101, 5–22. doi: 10.1016/j.ajhg.2017.06.005
- Wagner, B., Li, J., Liu, H., and Guo, G. (2013). Gene-environment correlation: Difficulties and a natural experiment-based strategy. *Am. J. Pub. Health* 103, S167–S173. doi: 10.2105/AJPH.2013.301415
- Wainschtein, P., Jain, D. P., Yengo, L., Zheng, Z., Cupples, L. A., Shadyab, A. H., et al. (2019). Recovery of trait heritability from whole genome sequence data. *BioRxiv [Preprint]*. doi: 10.1101/588020
- Wang, H., Zhang, F., Zeng, J., Wu, Y., Kemper, K. E., Xue, A., et al. (2019). Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the uk biobank. *BioRxiv [Preprint]*. doi: 10.1101/519538

- Wertz, J., Moffitt, T. E., Agnew-Blais, J., Arseneault, L., Belsky, D. W., Corcoran, D. L., et al. (2018). Using DNA from mothers and children to study parental investment in children's educational attainment. *BioRxiv [Preprint]*. doi: 10.1101/489781
- West, K. M., Blacksher, E., and Burke, W. (2017). Genomics, health disparities, and missed opportunities for the nation's research agenda. *JAMA* 317, 1831–1832. doi: 10.1001/jama.2017.3096
- Wilson, E. O. (1999). *Consilience: The Unity of Knowledge (Vol. 31)*. New York, NY: Vintage.
- Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., et al. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* 46, 1173–1186. doi: 10.1038/ng.3097
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., et al. (2010). Common snps explain a large proportion of the heritability for human height. *Nat. Genet.* 42, 565–569. doi: 10.1038/ng.608
- Yang, J., Loos, R. J., Powell, J. E., Medland, S. E., Speliotes, E. K., Chasman, D. I., et al. (2012). Fto genotype is associated with phenotypic variability of body mass index. *Nature* 490, 267–272. doi: 10.1038/nature11401
- Yang, J., Manolio, T. A., Pasquale, L. R., Boerwinkle, E., Caporaso, N., Cunningham, J. M., et al. (2011). Genome partitioning of genetic variation for complex traits using common snps. *Nat. Genet.* 43, 519–525. doi: 10.1038/ng.823

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Tanksley, Motz, Kail, Barnes and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.