

Genes, Environments, and Developmental Research: Methods for a Multi-Site Study of Early Substance Abuse

E. Jane Costello,^{1*} Lindon Eaves,² Patrick Sullivan,³ Martin Kennedy,⁴ Kevin Conway,⁵ Daniel E. Adkins,⁶ A. Angold,¹ Shaunna L. Clark,⁶ Alaattin Erkanli,¹ Joseph L. McClay,⁶ William Copeland,¹ Hermine H. Maes,² Youfang Liu,⁷ Ashwin A. Patkar,¹ Judy Silberg,² and Edwin van den Oord²

¹Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA

²Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA

³Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

⁴Department of Pathology, University of Otago, North Dunedin, New Zealand

⁵Division of Epidemiology, Services and Prevention Research, National Institute on Drug Abuse, Bethesda, MD, USA

⁶Center for Biomarker Research and Personalized Medicine, Virginia Commonwealth University School of Pharmacy, Richmond, VA, USA

⁷Thurston Arthritis Research Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

The importance of including developmental and environmental measures in genetic studies of human pathology is widely acknowledged, but few empirical studies have been published. Barriers include the need for longitudinal studies that cover relevant developmental stages and for samples large enough to deal with the challenge of testing gene–environment–development interaction. A solution to some of these problems is to bring together existing data sets that have the necessary characteristics. As part of the National Institute on Drug Abuse-funded Gene-Environment-Development Initiative, our goal is to identify exactly which genes, which environments, and which developmental transitions together predict the development of drug use and misuse. Four data sets were used of which common characteristics include (1) general population samples, including males and females; (2) repeated measures across adolescence and young adulthood; (3) assessment of nicotine, alcohol, and cannabis use and addiction; (4) measures of family and environmental risk; and (5) consent for genotyping DNA from blood or saliva. After quality controls, 2,962 individuals provided over 15,000 total observations. In the first gene–environment analyses, of alcohol misuse and stressful life events, some significant gene–environment and gene–development effects were identified. We conclude that in some circumstances, already collected data sets can be combined for gene–environment and gene–development analyses. This greatly reduces the cost and time needed for this type of research. However, care must be taken to ensure careful matching across studies and variables.

■ **Keywords:** genomewide association study, environment, development, multi-site

The focus of the next generation of gene–environment research will add development into the equation and focus upon gene–environment–development interactions (Rose & Dick, 2010, pp. 1854–1855).

It is clear from twin studies that the relative significance of genetic and environmental factors changes across some stages of development, notably childhood and adolescence (reviewed in Dick, 2011). Indeed, some have suggested that the first reason for failures to replicate gene–environment findings may be the different developmental status of the samples. For example, Cole and colleagues (2011, p. 1174) have argued that one explanation for the evolutionary main-

tenance of genetic features that are maladaptive late in life is that they must be selectively advantageous earlier in life

RECEIVED 17 August 2012; ACCEPTED 8 October 2012.

ADDRESS FOR CORRESPONDENCE: Professor E. J. Costello, Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Suite 22, Brightleaf Square, 905 West Main Street, Durham, NC 27701, USA. E-mail: jcostell@psych.duhs.duke.edu

*Authors are listed in the following order: The lead investigators at the study sites, followed by the rest of the research team in alphabetical order.

‘or else they would have been eradicated from the gene pool by adverse selection’. But there are many genetically driven changes with a strong developmental component, such as puberty or menopause, or the timing and synchronization of myelination in the central nervous system (van Ijzendoorn et al., 2011).

Grappling with gene–environment–development in the same study is important not only empirically but also methodologically:

We have known for decades that failure to incorporate both genetic and environmental factors in a joint analysis will weaken the observed associations between a true risk factor and disease occurrence. Because the pools of susceptible and non-susceptible persons are mixed, the observed associations tend to be shifted toward the null . . . Theoretically, if we are able to measure gene–environment interactions, we should sharpen our measurements of effects in subsets of the population and even potentially increase our statistical power in measuring such effects. (Khoury & Wacholder, 2009, p. 228)

The same argument can be applied to the importance of incorporating developmental measures (e.g., Chiang et al., 2011).

Longitudinal studies with repeated, prospective assessments using standardized measures of phenotype and envirotypes offer several opportunities for improved data quality compared with the standard methods of genetic case-control studies, which tend to rely on lifetime retrospective assessments using multiple diagnosticians and uncertain control over the uniformity of diagnostic criteria. In a longitudinal study, diagnoses are likely to be more consistent and there is likely to be less recall bias in participants’ psychiatric histories, which means fewer false negatives. Both of these qualities increase statistical power in hypothesis testing (Anastasi, 1950; Luan et al., 2001; Wong et al., 2003, 2004). The timing of environmental events relative to the onset of a disorder is also likely to be more accurate and less vulnerable to ‘seeking after meaning’ (Spatola et al., 2011).

The fourth reason to explore the use of longitudinal studies for genetics is that each participant provides multiple measures of phenotype and envirotypes. This increases the total number of observations, even after the corrections necessary to deal with the wave-to-wave correlations within individuals (Dunlap, 1996).

Despite theoretical discussions on the importance of Genome-Wide Association Studies (GWAS)-based gene–environment studies (sometimes called Gene–Environment-Wide Interactions Studies (GEWIS; Khoury & Wacholder, 2009), and even some discussion on developmental GEWIS (Lenroot & Giedd, 2011; Rose & Dick, 2010), there are few published results of GEWIS analyses in the behavioral sciences. The few studies of developmental

effects on gene–environment interplay have focused on individual genes such as BDNF (Casey et al., 2009); none has yet published uses of a genome-wide approach to psychiatric disorders (Lenroot & Giedd, 2011).

It is not hard to see why developmental GEWIS studies are so rare. The costs of initiating and maintaining large longitudinal studies are extremely high, and it is important that subjects be assessed using similar measures of both phenotype and environment over key developmental stages. Clinical studies rarely cover sufficiently long periods of time, so epidemiological samples are needed. Birth cohorts or other life-course epidemiologic studies, such as the nascent US National Children’s Study, are potential sources.

Another solution to this problem is to bring together multiple data sets to conduct joint analyses or meta-analysis. This approach depends crucially on the ability to combine the data across studies. Even before genetic analyses can begin, it is necessary to develop and test methods for harmonizing data across studies (Bookman et al., 2011; Cornelis et al., 2010; Fortier et al., 2011).

The National Institute on Drug Abuse (NIDA) and the National Cancer Institute (NCI) recognized both the promise and the problems of developmental GEWIS when they wrote the following in the Request for Applications:

Over many years, NIDA, other NIH Institutes, and other organizations have funded numerous high-quality longitudinal and developmental studies that contain a wealth of data from individuals who are at risk for, or are in the course of development, progression, and desistance of substance abuse and related phenotypes. . . . The GEDI [Gene-Environment-Development Initiative] seeks to build on this substantial public investment by soliciting applications that integrate environmental and developmental variables with genotypic information in order to permit comprehensive model-building and hypothesis testing for determining genetic, environmental, and developmental contributions to substance abuse and related phenotypes. (NIH/NIDA; R01DA024413)

NIDA and NCI hoped to take already existing materials and see if they could be woven into something that, if created from scratch, would have taken 20 years and untold millions of dollars. If successful, GEDI would be a proof-of-concept that could lead perhaps to an expansion of the collaborative group of studies.

In summary, we report here on a proof-of-concept study to carry out gene–environment, gene–development, and gene–environment–development analyses (both parallel and meta-analytic) using longitudinal, population-based data sets with repeated measures over childhood, adolescence, and early adulthood, with DNA available or obtainable, with comparable measures of drug and alcohol use, abuse, and dependence, and also of key environmental exposures.

TABLE 1
Characteristics of the Studies

	GSMS	CCC	VTSABD	CHDS
Number of data waves	7–10	3	5	18
Two or more developmental periods?	Yes	Yes	Yes	Yes
Age range from start to most recent wave (years)	9–26	9–17	8–30	1–30
Number of participants recruited at baseline	1,420	920	913 twin pairs	1,265
Number of participants successfully genotyped	784	518	913 (632 unique families)	747
Representative population sample?	Yes	Yes	Yes (twins)	Yes

Note: GSMS = Great Smoky Mountains Study; CCC = Caring for Children and the Community; VTSABD = Virginia Twin Study of Adolescent Behavioral Development; CHDS = Christchurch Health and Development Study.

Materials and Methods

Common GEDI Study Characteristics

The data sets that make up the consortium have the following characteristics in common: (1) General population samples; (2) multiple waves of data collection across childhood, adolescence, and young adulthood; (3) detailed assessments of drug use, abuse, and dependence (substance use disorders (SUDs)) and drug abuse symptoms; (4) assessments of comorbid psychiatric disorders, diagnosed using the *Diagnostic and Statistical Manual* (American Psychiatric Association, 1994) and psychiatric symptom scores; (5) measures of a range of environmental exposures, including serious life events. Methods used to collect information on diagnoses, symptoms, and environmental factors are described first, followed by brief descriptions of each study. Table 1 presents a summary of similarities and differences. Further details can be found in study-specific publications cited below.

GEDI Samples

1. Virginia Twin Study on Adolescent Behavioral Development (VTSABD; Simonoff et al., 1997).

The VTSABD is a cohort-longitudinal study of twins born between 1974 and 1983, ascertained primarily through the state school system and participating private schools in Virginia. Of 1,894 putative twin pairs, 1,412 families (75%, 2,775 children) participated and were included in the first wave of data collection. Three subsequent waves of data collection occurred at approximately 1½-year intervals, and the fifth wave when participants were in their mid-twenties (see Table 1). The study was limited to subjects of European ancestry as insufficient numbers from other ancestry groups were ascertainable. Parents completed a similar assessment on both twins. After age 18, the twins alone were interviewed individually by telephone. Over 8,500 family interview sets (parents and twins 8–17, twins 18+) have been completed. Variable numbers of subjects completed each interview wave, and 2,289 (82%) of the Wave 1 sample have completed the fifth wave.

2. Great Smoky Mountains Study (GSMS; Costello et al. 1996, 1997).

Three cohorts of boys and girls, aged 9, 11, and 13 years at intake in 1993, were selected from a rural population of

some 20,000 children using a household equal probability design. A two-phase procedure was used for White and African-American youth to increase power by oversampling children at risk for psychiatric and SUDs. Parents (usually mothers) of the first stage random population sample completed a questionnaire about their child's behavioral problems. Of 4,195 subjects selected, 95% ($N = 3,896$) of parents completed the screen. All children scoring above a predetermined threshold (the top 25% of the total scores), plus a 10% random sample of the remaining 75%, were recruited for detailed interviews. Results can be back-weighted to population levels for analyses. Half of the sample consists of females, and 6% are African Americans, reflecting the population of the study area. The interviewed sample of white and African-American subjects was 1,070 (80% of those recruited). American Indian youth were oversampled (100%) because they are an understudied group known to be at high risk for stressful events, substance disorders, and mood disorders. Of 431 age-eligible children, 350 (81%, 49% girls) participated. Thus, the size of total GSMS sample is $1,070 + 350 = 1,420$. Data collection is complete for ages 9–26, and age 30 interviews are in progress. By age 26 a total of 9,858 interviews had been completed; the average number of interviews per subject was seven, and by age 26, 97.3% completed two or more interviews.

3. The Caring for Children in the Community Study (CCC; Angold et al., 2002).

This representative study of psychiatric illness and service use in African-American and White youth took place in four rural counties in the southeastern USA. The two-stage sampling design and methods are similar to those used in the GSMS. Of 4,500 youth randomly selected from the 17,117 9- to 17-year-olds in the public school's database, 3,613 (80.0%) were successfully contacted and agreed to complete the behavioral screen. Of the 1,302 selected to participate in the study, 920 (70.7%) interviews were completed. Because CCC was also the only study in GEDI to contain more than a very few African-American participants, these were omitted from the multi-site analyses.

4. Child Health and Development Study (CHDS; Fergusson & Horwood, 2001).

The CHDS is a longitudinal study of a birth cohort from New Zealand. The cohort was based on an unselected

sample of 1,265 consecutive births (635 males; 630 females) occurring in the Christchurch urban region in mid-1977. The cohort has been studied at birth, 4 months and 1 year of age, annual intervals to the age of 16 years, and again at ages 18, 21, 25, and 30 years. Sample retention rates were high throughout the study and at age 30 the study was still able to assess over 80% of the surviving cohort.

Informed Consent in Each Study

Participants in all the studies gave consent for their DNA to be genotyped. However, depositing biological samples and genetic data in controlled-access biorepositories (e.g., dbGaP (Mailman et al., 2007) required a different level of consent. This was obtained for GSMS and VTSABD participants in year 1 of GEDI. Further consents were not required from CCC as the study was closed. CHDS subjects gave consent for genotyping only.

Each Institutional Review Board (IRB) had slightly different requirements for consent forms, but in general study participants were given the opportunity to consent to (1) completing only the assessment instruments; (2) assessment plus DNA collection for internal use only; or (3) assessment and DNA collection, the anonymized data to be put into a repository.

Blood Samples and Genotyping

VTSABD. Nine milliliters of blood were collected from each of the VTSABD participants in the first year of the GEDI study, that is, when subjects were aged 25 to 34. Blood and informed consent for genotyping and storage in dbGaP were obtained from 913 participants, of whom 281 were co-twins.

GSMS and CCC. Blood from the GSMS and CCC samples was collected at each assessment: 10 finger-stick samples were collected on specially prepared paper, dried, and shipped to the study laboratory, where they were stored at -23°C until they were assayed. A pilot study showed that even after 10 years of storage adequate DNA could be extracted from these samples. Most subjects (94%) provided at least one sample; the one collected as close to age 19 as possible was used for genotyping. Because there were so few African American participants for any of the data sets except CCC, the multi-site analyses excluded them, leaving 196 CCC and 784 GSMS participants with adequate genotype data and, in the case of GSMS, consent to deposit data in dbGaP. Since CCC was a closed study, no further consents were needed.

CHDS. Beginning in 2004 (at age 28), participants were asked for consent to provide saliva sample for DNA, and 918 (90% of the surviving cohort) consented. Consent for DNA collection was separate from consent for the rest of the study. In 2008–2009, participants were asked for consent for the GEDI multi-site GWAS, and 813 consented. Of these, 86%

provided peripheral blood samples, 8% provided saliva, and 6% provided buccal swabs (the latter proved not to provide samples of sufficient quality for genotyping). After quality control checks, good quality data were obtained on 747 participants. The New Zealand government does not permit the data to be deposited in dbGaP.

Blood samples from the VTSABD, GSMS, and CCC samples were sent to the Rutgers University Cell and DNA Repository for DNA extraction, and to the Genotyping Shared Resource at the Mayo Clinic Cancer Center for genotyping. DNA for the CHDS sample was prepared in New Zealand and also sent to Mayo for genotyping. DNA samples were randomized to plates within studies. All samples were genotyped using Illumina Human660W-Quad v1 DNA Analysis BeadChips. Quality control was carried out in the Department of Genetics at the University of North Carolina, Chapel Hill. In each of the four samples, single nucleotide polymorphisms (SNPs) with missing rate > 0.01 , minor allele frequency (MAF) < 0.05 , or extreme deviation ($p < 10^{-6}$) from the Hardy–Weinberg equilibrium (HWE) were removed from further analysis. Subjects with missing rate > 0.01 or unusual genome-wide homozygosity ($|\text{normalized homozygosity rate}| > 5$) were excluded. Sex was investigated using the no-call proportions of chrY SNPs and heterozygosity proportions for chrX SNPs. Mislabeled sex information was corrected after double-checking with the original data, and subjects with unexplainable results were deleted. In addition, pairwise identical-by-descent (IBD) estimation was evaluated to identify unexpected duplicates and relative pairs. We imputed SNP dosages in all samples using MACH (Liu et al., 2010b). The imputation reference was HapMap3 CEU (Utah residents with Northern and Western European ancestry from the CEPH collection) for subjects of European ancestry. All subjects with $pc1 < 0$ were grouped as ‘white’ and were imputed using HapMap3 CEU as reference; all subjects with $pc1 > 0$ were grouped together as ‘other’ and imputed using HapMap3 CEU + YRI as reference. Before imputation, the studies had between 496,000 and 515,000 SNPs that passed quality control. After imputation, all studies had over 1,193,000 total SNP values for analysis. The number of SNPs used in particular studies may differ based on cut-offs used for minor allele frequency.

Unobserved population admixture due to ancestry is a well-known confound in GWAS. To protect against false-positives due to ancestry, we extracted five principal components from each sample to correspond with ancestral and cryptic population stratification. To improve the efficiency of the principal components analysis (PCA) for control of population stratification, a subset of independent SNPs was selected using the PLINK function – independent-effect with proper parameters (window size = 50, the number of SNPs to shift the window at each step = 5, and the VIF threshold = 2). PCA was applied to the selected SNPs

using the smartpca module of EigenSoft (Price et al., 2008). Between 77,155 and 79,517 SNPs were used for each of the samples that we analyzed in the ancestry PCA. All genotyping and quality control (QC) was done blind to phenotype.

The VCU samples included in the analysis were all White, and HapMap3 CEU was used as the reference data to do the imputation for all the subjects from this data set. Most of the New Zealand samples were also White, although some were either Maori or mixed Maori and White. Since the number of Maori was small, we ignored the Maori's Asian genetic background and used HapMap3 CEU samples as the reference data to run the imputation for all subjects. The imputation quality will not be perfect for Maori, but we used Rsq (the imputation quality score from MACH) to remove badly imputed SNPs.

The Duke samples contained a range of populations, including White, African American, American Indian, Hispanic, and Asian, but most were either White or Black. For the imputation purpose, we split the samples into two groups. All the subjects with $pc1 > 0$ were grouped as 'Black' and were imputed using HapMap3 CEU + YRI as reference. All the subjects with $pc1 < 0$ were grouped as 'White' and were imputed using HapMap3 CEU as reference.

Environmental exposures. Each study collected extensive information on individual, family, and community risk for psychopathology. For the first analyses we selected a measure of exposure to potentially stressful life events (SLE) because several candidate-gene-based gene–environment studies have demonstrated significant gene–environment interplay using this measure (Karg et al., 2011). At each assessment, participants in all studies were asked to indicate whether they had experienced any of several potentially stressful life events such as losing a friend, or moving. Each study provided count variables of the total number of stressful life events recently. The period assessed varied by study from the previous 12 months to the previous 3 months. Stressful life event terms were centered to the study mean to reduce multi-collinearity with interaction terms. Although studies employed different primary periods, the parameter estimates for the association between stressful life events and substance-related outcomes were similar across studies. The study will only focus on effects that are robust to modest between-study differences in measurement or period assessed.

Data Analysis

Data analytic methods for each substance (cannabis, alcohol, and nicotine) varied, and will be described in the empirical reports. However, there are some general principles that we discuss here.

The overall goal is to determine which, if any, of the measured or imputed SNPs contribute to the explained variance in substance involvement after controlling for ancestry, sex,

and age. We expect that an SNP may contribute to substance involvement directly, via a main effect of the SNP on substance involvement, but that an SNP may also have a heterogeneous effect across individuals due to differential environmental exposures, including time. The degree to which genetic information influences substance involvement across the lifespan may vary over time as a function of time-specific life circumstances.

One of the issues that strongly influences the success of multi-site analyses is that of identifying measures of phenotypes or environmental factors that are comparable across data sets. For example, in the analyses described below, a factor score measuring alcohol involvement was estimated using Mplus 6.0 (<http://www.statmodel.com/>) from measures of quantity of use, frequency of use, and symptom counts common to all the data sets.

One of the unique features of GEDI is the rich developmental data inherent in each of the samples that allows us to investigate if there are genetic variants that influence substance use in a key period of development. As mentioned above, each of the samples covers a different age range (although they overlap across some ages) and each empirical paper takes a different approach to data harmonization and handling the developmental piece. As our first step toward data harmonization for GEDI, in consultation with our colleagues at Duke's Social Science Research Institute, we adopted the 'transform and recode' procedure most commonly used in harmonization studies (Bath et al., 2010). First, a key member of each study team is tasked to achieve consensus regarding whether it is possible to find variables (and associated response categories) that have the same 'face value'. Next, a new harmonized variable is created for each 'comparable' existing variable set by applying the transform and recode procedure to one or both of the original study measures such that existing codes for categories can be merged and relabeled in each study depending on the precise wording and ordering of the categories (Fortier et al., 2011).

For example, two primary, longitudinal measures of overtime alcohol consumption were generated to study the main effects of alcohol consumption on genetic variants: (1) a mixed model which explicitly models the developmental alcohol consumption trajectory spanning adolescence and early adulthood (ages 12–30), and (2) a simple mean of alcohol consumption (drinks per week) using repeated measures collected across adolescence (ages 12–21) for each individual. We selected these two specifications because (1) the trajectory outcome was indicated to be the best fitting longitudinal model, taking advantage of all repeated measures, and (2) the mean adolescent consumption outcome provided a simpler summary of individuals' drinking behavior, and thus provides greater continuity to existing literature (Agrawal et al., 2009, 2012; Grant et al., 2009). The harmonization methods used focused in this first instance on measures that are relatively constant in meaning across

development, such as number of drinks per week. With the help of our colleagues from the Data Harmonization team at Duke, we will then tackle measures that may change either content or meaning across development; for example, the content of self-regulation differs from one age to another, as does the developmental significance of an alcoholic drink (Sung et al., 2004).

Power for Analysis of Gene–Environment Interplay

In principle, biostatistical methods for testing for genetic association and gene–environment interplay do not differ from those for testing any other association, interaction, or correlation. The problem, of course, is the vast number of SNPs and environments (van den Oord, 2002) and the importance of controlling for false discoveries; that is, concluding that a marker affects an outcome when in reality it does not. We use an approach to control false discoveries based on the false discovery rate (FDR; Benjamini & Hochberg, 1995). In comparison to controlling a family-wise error rate – for example, the Bonferroni correction – the FDR (1) provides a better balance between the competing goals of finding true effects versus controlling false discoveries, (2) results in comparable standards for declaring significance across studies because it does not directly depend on the number of tests, and (3) is relatively robust against correlated tests (Borden et al., 1987; Fernando et al., 2004; Korn et al., 2004; Sabatti et al., 2003; Tsai et al., 2003). The FDR is commonly used in many high-dimensional applications and has also been applied successfully in the context of GWAS (Beecham, 2009; Lei, 2009; Liu, 2009). We chose the FDR threshold of 0.1 for declaring genome-wide significance (van den Oord & Sullivan, 2003), which means that on average 10% of the SNPs declared significant are expected to be false discoveries. Operationally (Black, 2004), the FDR is controlled using q -values that are FDRs calculated using the p -values of the markers as thresholds for declaring significance (Storey, 2003). It is important to note that performing many GWAS analyses does not present a problem for the FDR because it controls the expected ratio of false to all discoveries. Thus, when many GWAS are performed the number of false positives will increase and so will the number of true positives. The expected ratio of false positives to all discoveries will therefore remain 0.1, our threshold for declaring genome-wide significance.

Results, Cross-Validation, and Replication

The results of the first set of papers focus on the main effect of the SNP and the interaction term between the SNP and SLE exposure. The first three papers, currently under review, focus on the problem of alcohol use, number of cigarettes per week, and any cannabis use in the past 3 months. In each case, the environmental factor used was a measure of severe life events developed to be the same for all data sets.

Gene-Environment-Development Initiative takes two approaches for testing the validity of the results: cross-validation and replication. For the former we present the analyses separately for each data set and compare size and direction of effects across studies. This provides a more powerful test than the standard replication study because it involves complete genome-wide comparisons rather than simply comparisons of a few sites selected from one data set. The disadvantage is that individual data sets are necessarily smaller than the combined GEDI data set. We are therefore working to find other data sets with which we can carry out standard replication studies: comparing results on the ‘top hits’ from GEDI. There are few other data sets with the characteristics of the studies included in GEDI (multiple measures across adolescence and early adulthood of both substance use and abuse, and relevant environmental risk factors), but we have identified three with whom we are currently working (the Minnesota Twin and Family Study; Derringer et al., 2008), Finn Twin (Pagan et al., 2006), and the Center for Education and Drug Abuse Research (CEDAR) sample (Tarter & Vanyukov, 1994), with other collaborations under development.

Further Analyses

The next stage in the program of data analysis is to broaden it to include gene-based analyses (Neale & Sham, 2004), pathway analyses (Wang et al., 2007), and polygenic risk score analyses (Purcell et al., 2009). Gene-based analyses test whether any genes harbor an excess of SNPs with small p -values. Such analyses must account for both gene length and linkage disequilibrium between SNPs (see VEGAS, Liu et al., 2010a, for one example). Pathway analyses similarly test for an enrichment of SNPs with low p -values in genes involved in specific functional pathways (such as those in the Gene Ontology and Kyoto Encyclopedia of Genes and Genomes databases). Optimal approaches must account for varying gene size and SNP density, linkage disequilibrium within and between genes, and overlapping genes with similar annotations (see INRICH; Lee et al., 2012, for one example). Finally, the polygenic risk score analyses test a polygenic basis for the phenotype by looking at the variance accounted for by a given set of top SNPs determined by a p -value threshold (e.g., .005, .01, .10, or .25). In the first step, the sample is partitioned into discovery and replication sets. Parameter estimates, derived in the discovery sample, are used as weights to calculate scores in the replication set. Subsequently, a regression is performed on the disease state in the replication set from the polygenic score and then p -values and pseudo r^2 values are presented (see, for example, International Schizophrenia Consortium 2009). In each case we propose to analyze the individual data sets, and also to perform a meta-analysis of the entire group. In these cases we are applying analytic approaches already in use in

other GWAS studies to the GEDI program, but by focusing on the model term related to the interaction between the environmental exposure and SNP status, for example, this standard approach allows us to address a novel outcome – genes that moderate the association between the exposure and the outcome.

Next Steps: Candidate Gene Selection for Next-Generation Sequencing

In addition to GWAS, we will employ targeted capture (Gnirke et al., 2009) and massively parallel next-generation sequencing (McKernan et al., 2009) to exhaustively determine all genetic variations at selected genomic loci with evidence for involvement in SUD etiology. This approach uses a solution-based capture method (Gnirke et al., 2009), where genomic DNA from each subject is mixed with a ‘library’ of synthetic oligonucleotides, designed to be complementary to the genomic regions of interest. Molecular tags on these oligonucleotides allow them to be pulled out of solution, bringing the bound, complementary genomic DNA with them. This ‘captured’ DNA from each individual is labeled with a unique identifier and sequenced using next-generation, ultra-high throughput technology (Smith et al., 2010). This approach uses similar methods to exome sequencing (Ng et al., 2009), but here we will sequence the entire genomic region of each gene of interest, rather than just the exons, in order to capture all relevant variation.

We selected regions to include in our targeted capture library as follows:

1. GEDI GWAS findings for smoking and alcohol. For each SNP showing significant association at the genome-wide level (q -value < 0.1), we targeted the genomic region encompassing it (± 25 kb) for sequencing, plus any genes that fall within this 50-kb window. For SNPs showing ‘potentially interesting’ associations (q -values 0.1–0.2), we chose only those that fell within 25 kb of a gene, based on the principle that potentially interesting associations are more likely to be real if they are in, or close to, a gene. These criteria led us to select 17 loci covering 3.5 Mb of genomic DNA sequence.
2. Genes identified through published GWAS alcohol (Schumann et al., 2011) and smoking GWAS meta-analyses (Liu et al., 2010a; Thorgeirsson et al., 2010), plus gene nominations by expert colleagues (17 loci covering 2.5 Mb).
3. All human alcohol and aldehyde dehydrogenases, the key enzymes involved in alcohol metabolism (Edenberg, 2007; 28 loci covering 1.3 Mb).
4. Reward system genes, including dopaminergic (Di Chiara & Imperato, 1988), opioid (Le Merrer et al., 2009), and cannabinoid (Solinas et al., 2007) receptors and related metabolic genes (16 loci covering 1.09 Mb).
5. All remaining human nicotinic acetylcholine receptors not already selected (12 loci covering 0.42 Mb).
6. Additional priority genes close to GWAS hits (4 loci covering 0.35 Mb).
7. Prioritized candidate genes. We compiled three lists of candidate genes, the first based on previous associations in the literature with SUDs, the second comprising all known human genes involved in absorption, distribution, metabolism, and excretion (ADME) of drugs (www.pharmaadme.org), and the third included all human neuroactive ligand receptors from the KEGG database (Kanehisa & Goto, 2000; Kanehisa et al., 2011). We ranked genes by the number of times they co-occurred in the literature with the search terms ‘smoking’, ‘alcohol’, or ‘cannabis’, and filled the remainder of our targeted capture library with the top-ranked genes from this list (31 loci covering 1.1 Mb).

After removing overlap and collapsing neighboring genes into single loci, our selection encompassed 121 unique loci, covering a total of 10.2 Mb. However, human genomic DNA includes repetitive elements, including RNA and DNA transposons, which constitute approximately 45% of the human genome (Lander et al., 2001). These contribute no useful sequence information, because they are difficult to align to unique positions. After elimination of these repetitive elements, our final library encompassed approximately 5.5 Mb. We are currently sequencing this library in 1,000 individuals selected from the VTSABD and CHDS.

Discussion

As noted earlier, there is an abundance of literature recommending that genomics move in the direction of genome-wide gene–environment–development analyses, but very few data – in fact, we have found no empirical ‘developmental GEWIS’ (Khoury & Wacholder, 2009) studies so far. The few developmental gene–environment studies published have used a candidate gene approach (e.g., Adkins et al., 2012; Casey et al., 2009; Cole et al., 2011), and carry many of the limitations that have long plagued these studies (Sullivan et al., 2001).

Thus, we undertook GEDI partly as a proof of concept of the feasibility of a developmental GEWIS. Despite the lack of empirical data, most discussions in the literature take a gloomy view of the feasibility of development GEWIS, concentrating on the large sample sizes needed and the unreliability of measures of the envirotypes (Thomas, 2010). We acknowledge these problems unreservedly. On the other hand, there are other aspects of the situation that should be considered. First, estimates of sample sizes tend to be based on experience with early genetic studies that have collected cases and controls using what is often very unreliable data: lifetime psychiatric histories, or clinical diagnoses from hundreds or thousands of different clinicians. These

methods result not only in false positives but also in considerable numbers of false negatives (cases included non-cases because subjects have forgotten past episodes of illness). Both of these types of error inflate the sample size needed. In their article on estimating the size of gene–environment interactions in the presence of measurement error (Wong et al., 2004), and related articles (Luan et al., 2001; Wong et al., 2003), Wong and colleagues (2004) pointed out that accuracy in measuring all the related elements – genotype, phenotype, and exposure – critically affect the sample size needed for a given power. For example, ‘the difference between unreliable (correlation with true score = 0.4) and reliable ($r = 0.7$) measurements corresponds to a 20-fold difference in sample size’ (Moffitt et al., 2005, p. 476). Furthermore, Wong et al. (2003, p. 54) noted that ‘improving the measurement can be achieved by taking repeated measurements’, thus the longitudinal studies preferred for developmental analyses will also increase their power to test hypotheses.

Advantages of the GEDI consortium are that it includes only data sets with longitudinal and repeated assessments of subjects taken across the period of adolescence and young adulthood. We can expect subjects to be less vulnerable to either false remembering or false forgetting than those in studies using lifetime retrospective data. The use of repeated assessments also means that even a relatively small number of subjects yield a large number of person-observations (e.g., the 1,420 GSMS subjects yielded 9,858 person-observations by age 26). Even after controlling for non-independence of observations, this approach substantially increases the effective sample size and therefore power. Third, the studies used reliable assessments of symptoms and diagnoses created using a single taxonomy (DSM-IV) and highly structured diagnostic algorithms. Fourth, the studies used reasonably similar measures of key environmental risk factors.

The results of our first analyses (Copeland et al., 2012) appear to provide tentative empirical evidence that the combination of prospective, longitudinal assessment and careful attention to data harmonization can, to some extent, compensate for modest sample sizes. However, regardless of these benefits, there remains an acute need to build a broader and more inclusive consortium of qualifying longitudinal data sets. Our strongest recommendation from this experiment is for the creation of an international ‘developmental dbGaP’ of such data sets to optimize power in future investigations of environmentally and developmentally contingent genetic effects on behavioral outcomes.

Acknowledgments

Work on this project was carried out with support from NIDA (R01DA024413). Dr Adkins was supported by K01MH093731, and Dr Copeland by K23MH080230. We

are grateful to all the study participants who contributed to this work.

References

- Ackermann, K. H., Adams, N., Adler, C., Ahammed, Z., Ahmad, S., Allgower, C., . . . Zubarev, A. N. (2001). Elliptic flow in Au + Au collisions at square root(S) NN = 130 GeV. *Physical Review Letters*, *86*, 402–407.
- Adkins, D. E., Daw, J. K., McClay, J. L., & van den Oord, E. J. C. G. (2012). The influence of five monoamine genes on trajectories of depressive symptoms across adolescence and young adulthood. *Development and Psychopathology*, *24*, 267–285.
- Agrawal, A. N., Freedman, D., Cheng, Y. C., Lin, P., Shaffer, J. R., Sun, Q., . . . GENEVA Consortium. (2012). Measuring alcohol consumption for genomic meta-analyses of alcohol intake: Opportunities and challenges. *American Journal of Clinical Nutrition*, *95*, 539–547.
- Agrawal, A. J., Grant, D., Littlefield, A., Waldron, M., Pergadia, M. L., Lynskey, M. T., . . . Heath, A. C. (2009). Developing a quantitative measure of alcohol consumption for genomic studies on prospective cohorts. *Journal of Studies on Alcohol and Drugs*, *70*, 157–168.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed., DSM-IV). Washington, DC: APA.
- Anastasi, A. (1950). The concept of validity in the interpretation of test scores. *Journal of Psychology and Educational Measures*, *10*, 67–78.
- Angold, A., Erkanli, A., Farmer, E. M., Fairbank, J. A., Burns, B. J., Keeler, G., & Costello, E. J. (2002). Psychiatric disorder, impairment, and service use in rural African American and White youth. *Archives of General Psychiatry*, *59*, 893–901.
- Bath, P. A., Deeg, D., & Poppelaars, J. (2010). The harmonisation of longitudinal data: A case study using data from cohort studies in The Netherlands and the United Kingdom. *Ageing & Society*, *30*, 1419–1437.
- Beecham, G. W., Martin, E. R., Li, Y.-J., Slifer, M. A., Gilbert, J. R., Haines, J. L., & Pericak-Vance, M. A. (2009). Genome-wide association study implicates a chromosome 12 risk locus for late-onset Alzheimer disease. *American Journal of Human Genetics*, *84*, 36–43.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, *57*, 289–300.
- Black, M. A. (2004). A note on the adaptive control of false discovery rates. *J. R. Statist. Soc. B*, *66*(2), 297–304.
- Bookman, E. B., McAllister, K., Gillanders, E., Wanke, K., Balshaw, D., Rutter, J., . . . Birnbaum, L. S., for the NIH G × E Interplay Workshop participants. (2011). Gene-environment interplay in common complex diseases: Forging an integrative model—recommendations from an NIH workshop. *Genetic Epidemiology*, doi:10.1002/gepi.20571 (Epub ahead of print).
- Borden, K. A., Brown, R. T., Jenkins, P., & Clingerman, S. R. (1987). Achievement attributions and depressive symptoms

- in attention deficit disorder and normal children. *Journal of School Psychology*, 25, 399–404.
- Casey, B. J., Glatt, C. E., Tottenham, N., Soliman, F., Bath, K., Amso, D., . . . Lee, F. S. (2009). Brain-derived neurotrophic factor as a model system for examining gene by environment interactions across development. *Neuroscience*, 164, 108–120.
- Chiang, M. C., McMahon, K. L., de Zubicaray, G. I., Martin, N. G., Hickie, I., Toga, A. W., . . . Thompson, P. M. (2011). Genetics of white matter development: A DTI study of 705 twins and their siblings aged 12 to 29. *Neuroimage*, 54, 2308–2317.
- Cole, S. W., Arevalo, J. M., Manu, K., Telzer, E. H., Kiang, L., Bower, J. E., . . . Fuligni, A. J. (2011). Antagonistic pleiotropy at the human IL6 promoter confers genetic resilience to the pro-inflammatory effects of adverse social conditions in adolescence. *Developmental Psychology*, 47, 1173–1180.
- Copeland, W., Gottfredson, N., Adkins, D. E., Angold, A., Clark, S. L., Erkanli, A., . . . Costello, E. J. (2012). Stressful life events and alcohol use: A longitudinal G×E GWAS meta-analysis. Manuscript submitted for publication.
- Cornelis, M. C., Agrawal, A., Cole, J. W., Hansel, N. N., Barnes, K. C., Beaty, T. H., . . . GENEVA Consortium. (2010). The Gene, Environment Association Studies consortium (GENEVA): Maximizing the knowledge obtained from GWAS by collaboration across studies of multiple conditions. *Genetic Epidemiology*, 34, 364–372.
- Costello, E. J., Angold, A., Burns, B. J., Stangl, D. K., Tweed, D. L., Erkanli, A., & Worthman, C. M. (1996). The Great Smoky Mountains Study of Youth: Goals, designs, methods, and the prevalence of DSM-III-R disorders. *Archives of General Psychiatry*, 53, 1129–1136.
- Costello, E., Farmer, E., Angold, A., Burns, B. J., & Erkanli, A. (1997). Psychiatric disorders among American Indian and White youth in Appalachia: The Great Smoky Mountains Study. *American Journal of Public Health*, 87, 827–832.
- Derringer, J., Krueger, R. F., McGue, M., & Iacono, W. G. (2008). Genetic and environmental contributions to the diversity of substances used in adolescent twins: a longitudinal study of age and sex effects. *Addiction*, 103, 1744–1751.
- Di Chiara, G., & Imperato, A. (1988). Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proceedings of the National Academy of Sciences USA*, 85, 5274–5278.
- Dick, D. M. (2011). Developmental changes in genetic influences on alcohol use and dependence. *Child Development Perspectives*, 5, 223–230.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, 1, 170–177.
- Edenberg, H. J. (2007). The genetics of alcohol metabolism: Role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Research & Health*, 30, 5–13.
- Fergusson, D., & Horwood, L. (2001). The Christchurch health and development study: Review of findings on child and adolescent mental health. *Australian and New Zealand Journal of Psychiatry*, 35, 287–296.
- Fernando, R., Nettleton, D., Southey, B. R., Dekkers, J. C. M., Rothschild, M. F., & Soller, M. (2004). Controlling the proportion of false positives in multiple dependent tests. *Genetics*, 166, 611–619.
- Fortier, I., Doiron, D., Little, J., Ferretti, V., L'Heureux, F., Stolk, R. P., . . . International Harmonization Initiative. (2011). Is rigorous retrospective harmonization possible? Application of the DataSHaPER approach across 53 large studies. *International Journal of Epidemiology*, 40, 1314–1328.
- Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E. M., Brockman, W., . . . Nusbaum, C. (2009). Solution hybrid selection with ultra-long oligonucleotides for massively parallel-targeted sequencing. *Nature Biotechnology*, 27, 182–189.
- Grant, J. D., Agrawal, A., Bucholz, K. K., Madden, P. A. F., Pergadia, M. L., . . . Heath, A. C. (2009). Alcohol consumption indices of genetic risk for alcohol dependence. *Biological Psychiatry*, 66, 795–800.
- International Schizophrenia Consortium. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, 460, 748–752.
- Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28, 27–30.
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., & Tanabe, M. (2011). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research*, 40, D109–D114.
- Karg, K., Burmeister, M., Shedden, K., & Sen, S. (2011). The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: Evidence of genetic moderation. *Archives of General Psychiatry*, 68, 444–454.
- Khoury, M. J., & Wacholder, S. (2009). Invited commentary: From genome-wide association studies to gene-environment-wide interaction studies – Challenges and opportunities. *American Journal of Epidemiology*, 169, 227–230; discussion 234–225.
- Korn, E., Troendle, J., McShane, L. M., & Simon, R. (2004). Controlling the number of false discoveries: Application to high-dimensional genomic data. *Journal of Statistical Planning and Inference*, 124, 379–398.
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., . . . International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409, 860–921.
- Lee, P. H., O'Dushlaine, C., Thomas, B., & Purcell, S. M. (2012). INRICH: Interval-based enrichment analysis for genome-wide association studies. *Bioinformatics*, 28, 1797–1799.
- Le Merrer, J., Becker, J. A., Befort, K., & Kieffer, B. L. (2009). Reward processing by the opioid system in the brain. *Physiological Reviews*, 89, 1379–1412.
- Lei, S.-F., Yang, T.-L., Tan, L.-J., Chen, X.-D., Guo, Y., Guo, Y.-F., . . . Deng, H.-W. (2009). Genome-wide association scan for stature in Chinese: evidence for ethnic specific loci. *Human Genetics*, 125(1), 1–9.
- Lenroot, R. K., & Giedd, J. N. (2011). Annual research review: Developmental considerations of gene by environment

- interactions. *Journal of Child Psychology and Psychiatry*, *52*, 429–441.
- Liu, J., Zhang, Z., Bando, M., Itoh, T., Deardorff, M. A., Clark, D., . . . Krantz, I. D. (2009). Transcriptional dysregulation in *NIPBL* and cohesin mutant human cells. *PLoS Biol*, *7*(5): e1000119. doi:10.1371/journal.pbio.1000119
- Liu, J. Z., McRae, A. F., Nyholt, D. R., Medland, S. E., Wray, N. R., Brown, K. M., . . . Macgregor, S. (2010a). A versatile gene-based test for genome-wide association studies. *The American Journal of Human Genetics*, *87*, 139–145.
- Liu, J. Z., Tozzi, F., Waterworth, D. M., Pillai, S. G., Muglia, P., Middleton, L., . . . Marchini, J. (2010b). Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nature Genetics*, *42*, 436–440.
- Luan, J. A., Wong, M. Y., Day, N. E., & Wareham, N. J. (2001). Sample size determination for studies of gene-environment interaction. *International Journal of Epidemiology*, *30*, 1035.
- Mailman, M. D., Feolo, M., Jin, Y., Kimura, M., Tryka, K., Bagoutdinov, R., . . . Sherry, S. T. (2007). The NCBI dbGaP database of genotypes and phenotypes. *Nature Genetics*, *39*, 1181–1186.
- McKernan, K. J., Peckham, H. E., Costa, G. L., McLaughlin, S. F., Fu, Y., Tsung, E. F., . . . Blanchard, A. P. (2009). Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. *Genome Research*, *19*, 1527–1541.
- Moffitt, T. E., Caspi, A., & Rutter, M. (2005). Strategy for investigating interactions between measured genes and measured environments. *Archives of General Psychiatry*, *62*, 473–481.
- Neale, B. M., & Sham, P. C. (2004). The future of association studies: Gene-based analysis and replication. *American Journal of Human Genetics*, *75*, 353–362.
- Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C., . . . Shendure, J. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*, *461*, 272–276.
- Pagan, J. L., Rose, R. J., Viken, R. J., Pulkkinen, L., Kaprio, J., & Dick, D. M. (2006). Genetic and environmental influences on stages of alcohol use across adolescence and into young adulthood. *Behavior Genetics*, *36*, 483–497.
- Price, A. L., Butler, J., Patterson, N., Capelli, C., Pascali, V. L., Scarnicci, F., . . . Hirschhorn, J. N. (2008). Discerning the ancestry of European Americans in genetic association studies. *PLoS Genetics*, *4*, e236.
- Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., . . . International Schizophrenia Consortium (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, *460*, 748–752.
- Rose, R. J., & Dick, D. M. (2010). Commentary on Agrawal et al. (2010): Social environments modulate alcohol use. *Addiction*, *105*, 1854–1855.
- Sabatti, C., Service, S., & Freimer, N. (2003). False discovery rate in linkage and association genome screens for complex disorders. *Genetics*, *164*, 829–833.
- Schumann, G., Coin, L. J., Lourdasamy, A., Charoen, P., Berger, K. H., Stacey, D., . . . Elliott, P. (2011). Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (*AUTS2*) in the regulation of alcohol consumption. *Proceedings of the National Academy of Sciences, USA*, *108*, 7119–7124.
- Simonoff, E., Pickles, A., Meyer, J. M., Silberg, J. L., Maes, H. H., Loeber, R., . . . Eaves, L. J. (1997). The Virginia Twin Study of adolescent behavioral development: Influences of age, sex and impairment on rates of disorder. *Archives of General Psychiatry*, *54*, 801–808.
- Smith, A. M., Heisler, L. E., St Onge, R. P., Farias-Hesson, E., Wallace, I. M., Bodeau, J., . . . Nislow, C. (2010). Highly multiplexed barcode sequencing: An efficient method for parallel analysis of pooled samples. *Nucleic Acids Research*, *38*, e142.
- Solinas, M., Yasar, S., & Goldberg, S. R. (2007). Endocannabinoid system involvement in brain reward processes related to drug abuse. *Pharmacological Research*, *56*, 393–405.
- Spatola, C. A., Scaini, S., Pesenti-Gritti, P., Medland, S. E., Moruzzi, S., Ogliari, A., . . . Battaglia, M. (2011). Gene-environment interactions in panic disorder and CO₂ sensitivity: Effects of events occurring early in life. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *156B*, 79–88.
- Storey, J. (2003). The positive false discovery rate: A Bayesian interpretation and the q-value. *Annals of Statistics*, *31*, 2013–2035.
- Sullivan, P. F., Eaves, L. J., Kendler, K. S., & Neale, M. C. (2001). Genetic case-control association studies in neuropsychiatry. *Archives of General Psychiatry*, *58*, 1015–1024.
- Sung, M., AErkanli, A., Angold, A., & Costello, E. J. (2004). Effects of age at first substance use and psychiatric comorbidity on the development of substance use disorders. *Drug and Alcohol Dependence*, *75*, 287–299.
- Tarter, R. E., & Vanyukov, M. (1994). Alcoholism: A developmental disorder. *Journal of Consulting and Clinical Psychology*, *62*(6), 1096–1107.
- Thomas, D. (2010). Gene-environment-wide association studies: Emerging approaches. *Nature Reviews Genetics*, *11*, 259–272.
- Thorgeirsson, T. E., Gudbjartsson, D. F., Surakka, I., Vink, J. M., Amin, N., Geller, F., . . . Stefansson, K. (2010). Sequence variants at *CHRNA6* and *CYP2A6* affect smoking behavior. *Nature Genetics*, *42*, 448–453.
- Tsai, C. A., Hsueh, H. M., & Chen, J. J. (2003). Estimation of false discovery rates in multiple testing: Application to gene microarray data. *Biometrics*, *59*, 1071–1081.
- van den Oord, E. J. (2002). Association studies in psychiatric genetics: What are we doing? *Molecular Psychiatry*, *7*, 827–828.
- van den Oord, E. J., & Sullivan, P. (2003). False discoveries and models for gene discovery. *Trends in Genetics*, *19*, 537–542.
- van Ijzendoorn, M. H., Bakermans-Kranenburg, M. J., Belsky, J., Beach, S., Brody, G., Dodge, K. A., . . . Scott, S. (2011). Gene-by-environment experiments: A new approach to

- finding the missing heritability. *Nature Reviews Genetics*, 12, 881.
- Wang, K., Li, M., & Bucan, M. (2007). Pathway-based approaches for analysis of genomewide association studies. *The American Journal of Human Genetics*, 81, 1278–1283.
- Wong, M. Y., Day, N. E., Luan, J. A., Chan, K. P., & Wareham, N. J. (2003). The detection of gene–environment interaction for continuous traits: Should we deal with measurement error by bigger studies or better measurement? *International Journal of Epidemiology*, 32, 51.
- Wong, M. Y., Day, N. E., Luan, J. A., & Wareham, N. J. (2004). Estimation of magnitude in gene–environment interactions in the presence of measurement error. *Statistics in Medicine*, 23, 987–998.
-