ORIGINAL RESEARCH

A General Test for Gene–Environment Interaction in Sib Pair-based Association Analysis of Quantitative Traits

**Sophie van der Sluis · Conor V. Dolan ·
Michael C. Neale · Danielle Posthuma**

Abstract Several association studies support the hypothesis that genetic variants can modify the influence of environmental factors on behavioral outcomes, i.e., G × E interaction. The case-control design used in these studies is powerful, but population stratification with respect to allele frequencies can give rise to false positive or false negative associations. Stratification with respect to the environmental factors can lead to false positives or false negatives with respect to environmental main effects and G × E interaction effects as well. Here we present a model based on Fulker et al. (1999) and Purcell (2002) for the study of G × E interaction in family-based association designs, in which the effects of stratification can be controlled. Simulations illustrate the power to detect genetic and environmental main effects, and G × E interaction effects for the sib pair design. The power to detect interaction was studied in eight different situations, both with and without the presence of population stratification, and for categorical and continuous environmental factors. Results show that the power to detect genetic and environmental main effects, and G × E interaction effects, depends on the allele frequencies and the distribution of the environmental moderator. Admixture effects of realistic effect size lead only to very small stratification effects in the G × E component, so impractically large numbers of sib pairs are required to detect such stratification.

Keywords Gene by environment interaction ·
Genetic association · Sib pair-based designs

Introduction

Several studies have demonstrated that genetic variants may modify the influence of environmental factors on behavioral outcomes, or, equivalently, that environmental factors modify the effects of genes (e.g., Caspi et al. 2002, 2003; Foley et al. 2004; Huizinga et al. 2004; Yaffe et al. 2000). Recently, for example, Lasky-Su et al. (2007) reported SNP-by-socioeconomic status interaction with respect to attention hyperactivity deficit (ADHD) symptom count in and around the BDNF gene. Although some of these studies may be subject to methodological limitations (Eaves 2006), gene by environment interaction (G × E) should be considered in genetic association studies.

Most genetic association studies are based on a case-control design. While case-control designs for genetic association are powerful, they suffer from potential effects of population stratification, leading to false positives or negatives (e.g., Cardon and Bell 2001; Posthuma et al. 2004). Family-based designs, which compare genetically related subjects, are therefore preferred. Fulker et al. (1999) proposed a design for association analysis of quantitative traits in sib pair data using maximum-likelihood variance-components procedures. They showed that

S. van der Sluis (✉) · M. C. Neale · D. Posthuma
Biological Psychology, Vu University Amsterdam, Van der
Boechorststraat 1, Room 2B-37, Amsterdam 1081 BT,
The Netherlands
e-mail: s.van.der.sluis@psy.vu.nl

C. V. Dolan
Department of Psychology, FMG, University of Amsterdam,
Roeterstraat 15, Amsterdam 1018 WB, The Netherlands

M. C. Neale
Departments of Psychiatry and Human Genetics, Virginia
Institute of Psychiatry and Behavioral Genetics, Virginia
Commonwealth University, Richmond, VA, USA

the design is robust against spurious association stemming from population stratification, because the association effect is decomposed into a *within*-family effect and a *between*-family effect. The within-family effect is free of the potential effects of population stratification, because sibling pairs are drawn from the same family, and thus from the same genetic stratum. This design was extended by Neale et al. (1999) to include covariates, and by Abecasis et al. (2000a) to include multiple sibs, and parental information. The Fulker model now forms the basis for widely used statistical packages such as QTDT (Abecasis et al. 2000a, b).

Just like the association between genotypes and phenotypes, the associations between the environment and a phenotype, and between the G × E interaction and a phenotype are susceptible to the effects of population stratification. If two populations with (a) different allele frequencies, (b) different environmental frequencies (categorical environmental measure) or different environmental means (continuous environmental measure), and (c) different phenotypic means, are mixed, spurious environmental effects and spurious interaction effects can result, in addition to spurious allelic effects. In the sib pair design, it is therefore expedient to decompose into orthogonal between- and within-family effects (1) the allelic association; (2) the main effect of the environment; and (3) the G × E interaction.

In the present paper, we extend the sib pair model proposed by Fulker et al. (1999) to include environmental main effects and G × E interaction effects. The measured environmental variable may be either categorical or continuous. We report simulations carried out to investigate the statistical power to detect the presence of environmental main effects and G × E interaction effects for different effect sizes, different allele frequencies, and different environmental frequencies or means. In addition, we examine the statistical power to detect spurious G × E interaction due to population stratification.

Sib pair-based association including environmental effects and G × E interaction

We assume a diallelic marker with allele A_1 with frequency p, and allele A_2 with frequency $1 - p = q$, and genotypes A_1A_1, A_1A_2 and A_2A_2 with genotypic effects a, d and $-a$, respectively, (Fisher 1918; Falconer and Mackay 1996). For simplicity we assume throughout the paper that the marker under study is the actual quantitative trait locus (QTL), i.e., recombination fraction θ is zero. In reality, the genotypic value of a marker is unequal to zero only if the marker is the QTL, or if the marker is in linkage disequilibrium (LD) with the QTL. We assume that the observed trait value of an individual is a function of a major gene

effect (QTL), an additive polygenic genetic background effect, a shared familial or 'common environmental' effect, and an unshared, unique environmental effect (which also includes measurement error). Furthermore we assume that the effects of the additive polygenic genetic background, the common and unique environment, and the QTL are mutually uncorrelated, and that the additive polygenic genetic background effect and the environmental effects are normally distributed with mean zero.

If data from sibling pairs are available, the additive and dominance QTL effects may be partitioned into between- and within-family effects, as specified in Fulker et al. (1999), Abecasis et al. (2000a, b), and Posthuma et al. (2004). We will now introduce parameters for the effects of the environment, and for the interaction between genotype and environment.

Categorical environment

We assume that there are two environmental levels or conditions. The probability of being in either one of the environmental conditions is assumed not to depend on one's genotype, i.e., the correlation between genotype and environmental status (*rGE*) is zero. We also assume that the probability of being in either one of the environmental conditions is independent of the environmental condition of other family members.

We adopt the notation of van den Oord (1999), and model the environmental main effect (*e*) as the difference in the phenotypic means of environmental Conditions 1 and 2. To model the interaction effect, we assign interaction effect *i* to subjects with genotype A_1A_1 in Condition 2, interaction effect $-i$ to subjects with genotypes A_2A_2 in Condition 2, and interaction effect *c* to the heterozygotes A_1A_2 in Condition 2. Modeled as such, the interaction parameter *i* represents the difference between genotypic value *a* in Condition 2, and genotypic value *a* in Condition 1, after the main effect of the environmental condition has been taken into account. Similarly, interaction parameter $-i$ represents the difference between genotypic value $-a$ in Condition 2 and genotypic value $-a$ in Condition 1, after accounting for the environmental main effect. The interaction parameter *c* represents the difference between the dominance effect in Condition 2 and in Condition 1, once the main effect of the environment has been accounted for (see Mather and Jinks 1977, for a similar parameterization). For the purpose of illustration, the expected phenotypic means $\hat{y}_{kg}$ (i.e., the expected score of subjects in condition *k* with genotype *g*) are presented in Table 1 for the case of an environment with three levels.

Note that in the case of sib pair data (or data including multiple siblings and parents), various combinations of these means models are likely to be observed. When

Table 1 Expected phenotypic means for genotypic groups distinguished with respect to three environmental conditions

Genotype	Mean		
	Condition 1	Condition 2	Condition 3
A_1A_1	$\tau_i + a$	$\tau_i + a + e_1 + i_1$	$\tau_i + a + e_2 + i_2$
A_1A_2	$\tau_i + d$	$\tau_i + d + e_1 + c_1$	$\tau_i + d + e_2 + c_2$
A_2A_2	$\tau_i - a$	$\tau_i - a + e_1 - i_1$	$\tau_i - a + e_2 - i_2$

Note: τ_i denotes the family-specific intercept, a denotes the additive genotypic value, d denotes the dominance deviation for the heterozygotes, e_1 and e_2 denote the main effects for the environment (i.e., the in- or decrease of the phenotypic mean in Conditions 2 and 3 compared to Condition 1), i_1 and i_2 denote the G × E interaction effects for homozygotes (i.e., the difference between the genotypic values a in Conditions 2 and 3 compared to the genotypic value a in Condition 1, after the main effect for environment is accounted for), and c_1 and c_2 denote the G × E interaction for heterozygotes (i.e., the difference between the dominance effect in Conditions 2 and 3 compared to the dominance effect in Condition 1, once the main effect of the environment is accounted for)

family-data are available, the effects of the QTL, the environmental measure under study, and their interaction, may be further partitioned into between- and within-family effects. To illustrate, for sib pairs and a dichotomous environment, all possible combinations are presented in Tables 2–4.

In the case of the sib pair association design, the phenotypic score y_{ijkg} (i.e., the observed score y of subject j from family i in condition k with genotype g) is modeled as:

$$\hat{y}_{ijkg} = \tau_i + a_b A_{bi} + a_w A_{wij} + d_b D_{bi} + d_w D_{wij} + e_{bk} E_k + e_{wk} E_k + i_{bkg} I_{kg} + i_{wkg} I_{kg} + \varepsilon_{ij}, \tag{1}$$

where τ_i is the family-specific intercept, and ε_{ij} the residual term, i.e., the part of the phenotypic score y_{ijkg} that is not explained by the measured QTL, the environmental measure, or the interaction between these two, and which may be due to background genetic, or background environmental effects, unmodeled interactions, or measurement error. The parameters a_b and a_w are the estimated between- and within-family additive genetic effects of the marker, which are weighted by the derived coefficients A_{bi} and A_{wij}, respectively. These coefficients are either -1, $-\frac{1}{2}$, 0, $\frac{1}{2}$ or 1, as calculated in the 7th and 8th column of Tables 2–4 (see Fulker et al. 1999). Parameters d_b and d_w are the estimated between- and within-family dominance genetic effects of the marker, which are weighted by the derived coefficients D_{bi} and D_{wij}, respectively. These coefficients are either 0, $\frac{1}{2}$ or 1, as calculated in the 9th and 10th column of Tables 2–4 (see Posthuma et al. 2004). Similarly, the parameters e_{bk} and e_{wk} represent the between- and within-family effects of environmental condition k, which are weighted by the derived coefficient E_k. This coefficient

Table 2 Both sibs in environmental Condition 1

Genotype		Genotypic effect		Environmental effect		Additive and interaction		Dominance and interaction		Environment		Partitioned effects	
Sib1	Sib2	Sib1	Sib2	Sib1	Sib2	Between mean	Within diff/2	Between mean	Within diff/2	Between mean	Within diff/2	Sib1 (B + W)	Sib2 (B − W)
A_1A_1	A_1A_1	a	a	0	0	a_b	0	0	0	0	0	a_b	a_b
A_1A_1	A_1A_2	a	d	0	0	$\frac{1}{2}a_b$	$\frac{1}{2}a_w$	$\frac{1}{2}d_b$	$-\frac{1}{2}d_w$	0	0	$\frac{1}{2}[a_b + a_w] + \frac{1}{2}[d_b - d_w]$	$\frac{1}{2}[a_b - a_w] + \frac{1}{2}[d_b + d_w]$
A_1A_1	A_2A_2	a	$-a$	0	0	0	a_w	0	0	0	0	a_w	$-a_w$
A_1A_2	A_1A_1	d	a	0	0	$\frac{1}{2}a_b$	$-\frac{1}{2}a_w$	$\frac{1}{2}d_b$	$\frac{1}{2}d_w$	0	0	$\frac{1}{2}[a_b - a_w] + \frac{1}{2}[d_b + d_w]$	$\frac{1}{2}[a_b + a_w] + \frac{1}{2}[d_b - d_w]$
A_1A_2	A_1A_2	d	d	0	0	0	0	d_b	0	0	0	d_b	d_b
A_1A_2	A_2A_2	d	$-a$	0	0	$-\frac{1}{2}a_b$	$\frac{1}{2}a_w$	$\frac{1}{2}d_b$	$\frac{1}{2}d_w$	0	0	$\frac{1}{2}[-a_b + a_w] + \frac{1}{2}[d_b + d_w]$	$-\frac{1}{2}[a_b + a_w] + \frac{1}{2}[d_b - d_w]$
A_2A_2	A_1A_1	$-a$	a	0	0	0	$-a_w$	0	0	0	0	$-a_w$	a_w
A_2A_2	A_1A_2	$-a$	d	0	0	$-\frac{1}{2}a_b$	$-\frac{1}{2}a_w$	$\frac{1}{2}d_b$	$-\frac{1}{2}d_w$	0	0	$-\frac{1}{2}[a_b + a_w] + \frac{1}{2}[d_b - d_w]$	$\frac{1}{2}[-a_b + a_w] + \frac{1}{2}[d_b + d_w]$
A_2A_2	A_2A_2	$-a$	$-a$	0	0	$-a_b$	0	0	0	0	0	$-a_b$	$-a_b$

The main effect for the environmental and the G × E interaction are zero in both sibs

is either $-\frac{1}{2}$, 0, $\frac{1}{2}$ or 1, as calculated in the 11th and 12th column of Tables 2–4. The parameters i_{bkg} and i_{wkg} represent the between- and within-family effects of the interaction of genotype g and environmental Condition k, which are weighted by the derived coefficient I_{kg}. This coefficient is either $-\frac{1}{2}$, 0, $\frac{1}{2}$ or 1, as calculated in the 7th to 10th column of Tables 2–4.

Continuous environment

If the environmental measure is continuous in nature, rather than categorical, the model as presented in Eq. 1, is altered as follows. The between and within-family environmental parameters e_b and e_w are simply weighted by the subject's score on the continuous environmental measure, E_j, just as the genotype-dependent between and within-family interaction parameters i_{bg} and i_{wg}. The continuous environmental measure is now modeled as a continuous moderator, in the manner proposed by Purcell (2002). In the case of a continuous environmental measure, the phenotypic score y_{ijg} is modeled as:

$$\hat{y}_{ijg} = \tau_i + a_b A_{bi} + a_w A_{wij} + d_b D_{bi} + d_w D_{wij} + e_b E_j + e_w E_j + i_{bg} E_j + i_{wg} E_j + \varepsilon_{ij}.$$

(2)

Given these additional effects of the environment and the G × E interaction, the variance-covariance matrix for siblings j and m of the ith family, Σ_i, is given as:

$$\sum{}_i = \left\{ \begin{array}{ll} \sigma_{QTL-A}^2 + \sigma_{QTL-D}^2 + \sigma_{ENV}^2 + \sigma_{INT}^2 + \sigma_s^2 + \sigma_u^2 & \text{if } j = m \\ \hat{\pi}_{ijk}\sigma_{QTL-A}^2 + \hat{Z}_{ijk}\sigma_{QTL-D}^2 + \sigma_s^2 & \text{if } j \neq m \end{array} \right\}$$

(3)

where σ_{QTL-A}^2 is the variance due to the additive genetic effect of the marker, σ_{QTL-D}^2 is the variance due to the dominance effects of the marker, σ_{ENV}^2 is the variance due to the measured environmental indicator, and σ_{INT}^2 is the variance due to the interaction between the marker and the environmental measure. After all these measured effects are accounted for, σ_s^2 denotes the residual sibling resemblance, which is due to shared alleles other than the QTL alleles under study, shared environmental effects other than the measured environmental variable under study, or covariance between these two sources. Finally, σ_u^2 denotes all variance that is not shared by siblings from the same family, and which is due to unshared alleles, and unshared environmental effects. The covariance between the phenotypic scores of siblings equals the additive and dominance QTL variance, weighted by $\hat{\pi}_{ijk}$ (the estimated

proportion of alleles that siblings j and m from family i share IBD, i.e. $p(IBD = 2) + \frac{1}{2} p(IBD = 1)$) and $\hat{Z}_{ijk}$ (the probability of complete IBD sharing between siblings j and m, i.e., $p(IBD = 2)$), respectively. Because we assumed the environmental effect under study to be unrelated to genotype (i.e. $rGE = 0$) or to family membership, the environmental effect and the interaction effect only contribute to the variance through σ_{ENV}^2 and σ_{INT}^2, but not to the covariance between siblings j and m. Note that in practice, σ_{ENV}^2, σ_{INT}^2, and σ_u^2 cannot be estimated individually (i.e., only the sum of them can be estimated). Note also that when the marker under study is indeed the actual QTL, as is assumed throughout this paper, and the environmental measure is an accurate reflection of the true environmental moderator, the expected variance–covariance matrix Σ_i reduces to

$$\sum{}_i = \left\{ \begin{array}{ll} \sigma_s^2 + \sigma_u^2 & \text{if } j = m \\ \sigma_s^2 & \text{if } j \neq m \end{array} \right\}$$

(4)

(Fulker et al. 1999), because the family variance–covariance matrices Σ_i are formed *conditionally* on the marker genotypes of the siblings, and *conditionally* on their environmental status. Conditionally on the marker genotype and the environmental status of the siblings, there no longer is any variation in marker genotype or environmental status, so these variables no longer explain any variance. The effects of the marker and the environment are then modeled via the mean structure only, per Eqs. 1 and 2.

In the variance-components approach, the means and variances of related individuals are modeled simultaneously, as a function of the set of parameters θ which equals $\theta = \{\tau_i, a_b, a_w, d_b, d_w, e_b, e_w, i_{bg}, i_{wg}, \sigma_s^2, \sigma_u^2\}$, if the marker under study is the QTL. Maximum likelihood estimation can be used to obtain parameter estimates, and likelihood ratio tests can be used to test specific constraints on the parameters (Azzelini 1996). For example, one can test whether the regression weight for the between-family additive genetic marker effect, a_b, is equal to the within-family additive genetic marker effect, a_w, the idea being that a_b only differs from a_w when population stratification significantly influences the results of the test for genetic association.

Sib pair association models including a measured environment and G × E interaction effects can readily be

Table 3 Both sibs in environmental Condition 2

Genotype		Genotypic effect		Environmental effect		Additive and interaction		Dominance and interaction		Environment		Partitioned effects	
Sib1	Sib2	Sib1	Sib2	Sib1	Sib2	Between mean	Within diff/2	Between mean	Within diff/2	Between mean	Within diff/2	Sib1 (B + W)	Sib2 (B − W)
A_1A_1	A_1A_1	$a + i$	$a + i$	e	e	$a_b + i_b$	0	0	0	e_b	0	$[a_b + i_b] + e_b$	$[a_b + i_b] + e_b$
A_1A_1	A_1A_2	$a + i$	$d + c$	e	e	$\frac{1}{2}(a_b + i_b)$	$\frac{1}{2}(a_w + i_w)$	$\frac{1}{2}(d_b + c_b)$	$-\frac{1}{2}(d_w + c_w)$	e_b	0	$\frac{1}{2}[a_b + a_w + i_b + i_w] + \frac{1}{2}[d_b - d_w + c_b - c_w] + e_b$	$\frac{1}{2}[a_b - a_w + i_b - i_w] + \frac{1}{2}[d_b + d_w + c_b + c_w] + e_b$
A_1A_1	A_2A_2	$a + i$	$-a - i$	e	e	0	$a_w + i_w$	0	0	e_b	0	$[a_w + i_w] + e_b$	$-[a_w + i_w] + e_b$
A_1A_2	A_1A_1	$d + c$	$a + i$	e	e	$\frac{1}{2}(a_b + i_b)$	$-\frac{1}{2}(a_w + i_w)$	$\frac{1}{2}(d_b + c_b)$	$\frac{1}{2}(d_w + c_w)$	e_b	0	$\frac{1}{2}[a_b - a_w + i_b - i_w] + \frac{1}{2}[d_b + d_w + c_b + c_w] + e_b$	$\frac{1}{2}[a_b + a_w + i_b + i_w] + \frac{1}{2}[d_b - d_w + c_b - c_w] + e_b$
A_1A_2	A_1A_2	$d + c$	$d + c$	e	e	0	0	$d_b + c_b$	0	e_b	0	$[d_b + c_b] + e_b$	$[d_b + c_b] + e_b$
A_1A_2	A_2A_2	$d + c$	$-a - i$	e	e	$-\frac{1}{2}(a_b + i_b)$	$\frac{1}{2}(a_w + i_w)$	$\frac{1}{2}(d_b + c_b)$	$\frac{1}{2}(d_w + c_w)$	e_b	0	$-\frac{1}{2}[a_b - a_w + i_b - i_w] + \frac{1}{2}[d_b + d_w + c_b + c_w] + e_b$	$-\frac{1}{2}[a_b + a_w + i_b + i_w] + \frac{1}{2}[d_b - d_w + c_b - c_w] + e_b$
A_2A_2	A_1A_1	$-a - i$	$a + i$	e	e	0	$-(a_w + i_w)$	0	0	e_b	0	$-[a_w + i_w] + e_b$	$[a_w + i_w] + e_b$
A_2A_2	A_1A_2	$-a - i$	$d + c$	e	e	$-\frac{1}{2}(a_b + i_b)$	$-\frac{1}{2}(a_w + i_w)$	$\frac{1}{2}(d_b + c_b)$	$-\frac{1}{2}(d_w + c_w)$	e_b	0	$-\frac{1}{2}[a_b + a_w + i_b + i_w] + \frac{1}{2}[d_b - d_w + c_b - c_w] + e_b$	$-\frac{1}{2}[a_b - a_w + i_b - i_w] + \frac{1}{2}[d_b + d_w + c_b + c_w] + e_b$
A_2A_2	A_2A_2	$-a - i$	$-a - i$	e	e	$-(a_b + i_b)$	0	0	0	e_b	0	$-[a_b + i_b] + e_b$	$-[a_b + i_b] + e_b$

The main effect for the environmental and the G × E interaction are non-zero in both sibs

Table 4 Sibs in different environmental conditions: sib1 in Condition 1 and sib2 in Condition 2

Genotype		Genotypic effect		Environmental effect		Additive and interaction		Dominance and interaction		Environment		Partitioned effects	
Sib1	Sib2	Sib1	Sib2	Sib1	Sib2	Between mean	Within diff/2	Between mean	Within diff/2	Between mean	Within diff/2	Sib1 (B + W)	Sib2 (B − W)
A_1A_1	A_1A_1	a	$a + i$	0	e	$a_b + \tfrac{1}{2} i_b$	$-\tfrac{1}{2} i_w$	0	0	$\tfrac{1}{2} e_b$	$-\tfrac{1}{2} e_w$	$[a_b + \tfrac{1}{2} i_b - \tfrac{1}{2} i_w] + \tfrac{1}{2}[e_b - e_w]$	$[a_b + \tfrac{1}{2} i_b + \tfrac{1}{2} i_w] + \tfrac{1}{2}[e_b + e_w]$
A_1A_1	A_1A_2	a	$d + c$	0	e	$\tfrac{1}{2} a_b$	$\tfrac{1}{2} a_w$	$\tfrac{1}{2} d_b + \tfrac{1}{2} c_b$	$-\tfrac{1}{2} d_w - \tfrac{1}{2} c_w$	$\tfrac{1}{2} e_b$	$-\tfrac{1}{2} e_w$	$\tfrac{1}{2}[a_b + a_w] + \tfrac{1}{2}[d_b - d_w + c_b - c_w] + \tfrac{1}{2}[e_b - e_w]$	$\tfrac{1}{2}[a_b - a_w] + \tfrac{1}{2}[d_b + d_w + c_b + c_w] + \tfrac{1}{2}[e_b + e_w]$
A_1A_1	A_2A_2	a	$-a - i$	0	e	$-\tfrac{1}{2} i_b$	$a_w + \tfrac{1}{2} i_w$	0	0	$\tfrac{1}{2} e_b$	$-\tfrac{1}{2} e_w$	$[a_w - \tfrac{1}{2} i_b + \tfrac{1}{2} i_w] + \tfrac{1}{2}[e_b - e_w]$	$[-a_w - \tfrac{1}{2} i_b + \tfrac{1}{2} i_w] + \tfrac{1}{2}[e_b + e_w]$
A_1A_2	A_1A_1	d	$a + i$	0	e	$\tfrac{1}{2} a_b + \tfrac{1}{2} i_b$	$-\tfrac{1}{2} a_w - \tfrac{1}{2} i_w$	$\tfrac{1}{2} d_b$	$\tfrac{1}{2} d_w$	$\tfrac{1}{2} e_b$	$-\tfrac{1}{2} e_w$	$\tfrac{1}{2}[a_b - a_w + i_b - i_w] + \tfrac{1}{2}[d_b + d_w] + \tfrac{1}{2}[e_b - e_w]$	$\tfrac{1}{2}[a_b + a_w + i_b + i_w] + \tfrac{1}{2}[d_b - d_w] + \tfrac{1}{2}[e_b + e_w]$
A_1A_2	A_1A_2	d	$d + c$	0	e	0	0	$d_b + \tfrac{1}{2} c_b$	$-\tfrac{1}{2} c_w$	$\tfrac{1}{2} e_b$	$-\tfrac{1}{2} e_w$	$[d_b + \tfrac{1}{2} c_b - \tfrac{1}{2} c_w] + \tfrac{1}{2}[e_b - e_w]$	$[d_b + \tfrac{1}{2} c_b + \tfrac{1}{2} c_w] + \tfrac{1}{2}[e_b + e_w]$
A_1A_2	A_2A_2	d	$-a - i$	0	e	$-\tfrac{1}{2} a_b - \tfrac{1}{2} i_b$	$\tfrac{1}{2} a_w + \tfrac{1}{2} i_w$	$\tfrac{1}{2} d_b$	$\tfrac{1}{2} d_w$	$\tfrac{1}{2} e_b$	$-\tfrac{1}{2} e_w$	$-\tfrac{1}{2}[a_b - a_w + i_b - i_w] + \tfrac{1}{2}[d_b + d_w] + \tfrac{1}{2}[e_b - e_w]$	$-\tfrac{1}{2}[a_b + a_w + i_b + i_w] + \tfrac{1}{2}[d_b - d_w] + \tfrac{1}{2}[e_b + e_w]$
A_2A_2	A_1A_1	$-a$	$a + i$	0	e	$\tfrac{1}{2} i_b$	$-a_w - \tfrac{1}{2} i_w$	0	0	$\tfrac{1}{2} e_b$	$-\tfrac{1}{2} e_w$	$-[a_w - \tfrac{1}{2} i_b + \tfrac{1}{2} i_w] + \tfrac{1}{2}[e_b - e_w]$	$[a_w + \tfrac{1}{2} i_b + \tfrac{1}{2} i_w] + \tfrac{1}{2}[e_b + e_w]$
A_2A_2	A_1A_2	$-a$	$d + c$	0	e	$-\tfrac{1}{2} a_b$	$-\tfrac{1}{2} a_w$	$\tfrac{1}{2} d_b + \tfrac{1}{2} c_b$	$-\tfrac{1}{2} d_w - \tfrac{1}{2} c_w$	$\tfrac{1}{2} e_b$	$-\tfrac{1}{2} e_w$	$-\tfrac{1}{2}[a_b + a_w] + \tfrac{1}{2}[d_b - d_w + c_b - c_w] + \tfrac{1}{2}[e_b - e_w]$	$-\tfrac{1}{2}[a_b - a_w] + \tfrac{1}{2}[d_b + d_w + c_b + c_w] + \tfrac{1}{2}[e_b + e_w]$
A_2A_2	A_2A_2	$-a$	$-a - i$	0	e	$-a_b - \tfrac{1}{2} i_b$	$\tfrac{1}{2} i_w$	0	0	$\tfrac{1}{2} e_b$	$-\tfrac{1}{2} e_w$	$-[a_b + \tfrac{1}{2} i_b - \tfrac{1}{2} i_w] + \tfrac{1}{2}[e_b - e_w]$	$[-a_b - \tfrac{1}{2} i_b - \tfrac{1}{2} i_w] + \tfrac{1}{2}[e_b + e_w]$

The main effect for the environmental and the G × E interaction are zero in sib1 and non-zero in sib2

implemented in the Mx software package[1] (Neale et al. 2003). Appendices I and II include example Mx-scripts for the case of sib pair data and a dichotomous environment or a continuous environment, respectively. Adaptation of these scripts for the modeling of more than two siblings, or categorical environments with multiple levels is quite straightforward. Extension of these scripts to include data from nuclear families (parents and offspring; Abecasis et al. 2000a) requires some modifications which are spelled out in the Mx-script provided by Posthuma et al. (2004) in their Appendix II. An example script for the modeling of data from monozygotic and dizygotic twin pairs is available online.[2] Note that whereas sib pair data only allow distinction between σ_s^2 and σ_u^2, twin data allow a more detailed decomposition of the background variance into variance due to additive genetic effects (σ_A^2), common environmental effects (σ_C^2) or dominance genetic effects (σ_D^2), and unshared environmental effects (σ_E^2).

Power calculations for the G × E model

We performed a series of simulation studies to investigate the power of the extended sib pair model to detect the G × E interaction effects. We considered both a dichotomous environmental measure and a continuous environmental measure. All simulations were based on simulated sibling pairs only, and the simulated marker was assumed to be the actual QTL. The power analyses are thus limited to the detection of effects on the means (association), not the variances (linkage).

Procedures

Simulations involved a diallelic marker locus with frequency p of the increaser allele A_1 being .5 or .2. Except where noted, QTL dominance effects were absent. In the case of a dichotomous environment, the frequency b_1 of Condition 1 was either .5 or .2. The continuous environmental measure was standard normally distributed, i.e., Environment $\sim N(0,1)$. Simulated environmental values were uncorrelated to the simulated genotypes (e.g., $rGE = 0$). The continuous phenotype was standard normally distributed when all measured allelic and environmental effects were zero. When these effects are not zero, the phenotypic mean and variance deviate from 0 and 1, respectively. The degree of deviation depends on their effect size.

The QTL effect, the environmental effect, and the interaction were manipulated so that in isolation, these factors each explained 1%, 2.5% or 5% of the total phenotypic variance in the total sample. In the simulations with a dichotomous environment, these effect sizes were determined for the case that both environmental conditions and alleles were evenly distributed (i.e., $b_1 = b_2 = .5$ and $p = q = .5$). Note, however, that the percentage of explained variance depends on the allele frequencies and the distribution of the environmental variable. For instance, if the parameters representing the genotypic effect of the QTL locus are chosen such that the locus explains 5% of the variance in the case that $p = q = .5$, this same locus (i.e., same genotypic values) only explains 3.3% of the variance in the case that $p = .2$. Likewise, an environmental effect that explains 5% of the variance if $b_1 = b_2 = .5$, explains only 3.3% of the variance if $b_1 = .2$. For the simulations including a continuous environment, effect sizes were determined for the case that $p = q = .5$.

Where noted, population stratification was generated by mixing two samples (A and B) of equal proportions, with different phenotypic means (μ_A and μ_B), and different marker allele frequencies ($p_A = .7$, $p_B = .3$). In the case of a dichotomous environment, environmental frequencies differed between samples A and B ($b_{A1} = .3$, $b_{B1} = .7$). In the case of a continuous environment, the environmental means differed between samples A and B ($\mu_{envA} = 0$, $\mu_{envB} = 2$). The phenotypic means of samples A and B were selected such that admixture accounted for 20% of the total phenotypic variance in the combined population, i.e., $(\mu_A - \mu_B)^2 / 4\sigma_{TOT}^2 = .20$ (see Abecasis et al. 2000a).[3] The mixture of these two samples with different phenotypic means, different allele frequencies, and different environmental frequencies or means, results in spurious allelic, spurious environmental, and spurious interaction effects. The emphasis in these simulations is thus on the detection of false positives, but false negatives are theoretically possible (e.g., Posthuma et al. 2004; Neale et al. 1999).

For all simulations, background variance was modeled such that, after accounting for the QTL-effect, the environmental main effect, and the interaction, 30% of the remaining variance was attributable to additive polygenic genetic effects (A), and 70% was due to non-shared environmental effects (E). Covariance between the sibs due to shared environmental components (C) was fixed to zero, so all resemblance between the sibs was due to genetic factors only (i.e., the QTL and other unidentified genes). Because A and C cannot be distinguished unless the sample includes

---

[1] The Mx program is freely available at http://www.vcu.edu/mx/.

[2] www.psy.vu.nl/u/s.van.der.sluis.

[3] Effect size is defined as the variance explained by the effect, divided by the total variance, i.e., $[\frac{1}{2}(\mu_A - \mu_G)^2 + \frac{1}{2}(\mu_B - \mu_G)^2] / \sigma_{TOT}^2$, where μ_A and μ_B are the means of the different subpopulations, and μ_G is the mean of the combined population, which is defined as $(\mu_A + \mu_B)/2 = \frac{1}{2}\mu_A + \frac{1}{2}\mu_B$. Substitution of μ_G with $\frac{1}{2}\mu_A + \frac{1}{2}\mu_B$ gives $(\mu_A - \mu_B)^2 / 4\sigma_{TOT}^2$.

monozygotic twins, in addition to regular siblings or dizygotic twins, the term σ_s^2 will include all the siblings' resemblance due to shared genes other than the QTL, and common environmental influences. The term σ_u^2 then includes all variance due to unidentified non-shared genes and non-shared environmental effects. Note that, in general, the power to detect the effects of interest increases as the residual sibling resemblance σ_s^2 increases, even if the exact nature of resemblance (genes or environment) cannot be distinguished. This is because, as a result of increasing σ_s^2, the non-shared component σ_u^2 decreases, and less unshared variance implies less "noise" (i.e., unexplained variance), which increases statistical power. The choice to fix shared environmental effects to zero in all simulations, thus results in conservative estimations of the power to detect the effects of interest.

All data simulations were performed in the R program,[4] and exact data simulation was used for all analyses (van der Sluis et al. 2008; Bollen and Stine 1993; Dolan et al. 2005). Exact data simulation can be used when sufficient summary statistics are available in theory, i.e., when all information present in the raw data can be summarized sufficiently in the variance covariance matrix Σ, and the means vector μ. Exact data simulation implies the simulation of raw data that are transformed to fit the true model exactly. Consequently, when the true model is fitted to these data, all parameter estimates used to simulate the data are recovered exactly. Subsequently, the constrained, nested (wrong) model is fitted to the data, in which parameters of interest are fixed to zero, or constrained to be equal. Minus twice the difference in the log-likelihoods of the true model and the nested model asymptotically equals the non-centrality parameter λ of the non-central χ^2-distribution, with df equal to the difference in the number of parameters estimated. This non-centrality parameter can subsequently be used to calculate the sample size N required for a chosen power level, given a chosen critical value α (Saris and Satorra 1993).[5]

The results of power analyses based on exact data simulation equal exactly the results obtained through the analysis of (population or expected) summary statistics Σ and μ. Also, as in power calculations based on summary statistics, these results are asymptotically similar to results obtained through Monte Carlo simulation (depending on the number of runs, and the sample size N used in the Monte Carlo procedure). In contrast to Monte Carlo simulation, however, exact data simulation obviates the

requirement to replicate the analyses in different runs because the quasi-randomly generated data are transformed to fit the true model exactly. Exact data simulation is therefore not only easy to perform but also computationally light compared to Monte Carlo simulation, which is why we chose to use exact data simulation here. We refer to Van der Sluis et al. (2008) for an extensive discussion on exact data simulation.

Given non-centrality parameter λ, the Mx program computes the total sample size that would be required, given the reported proportion of subjects in each distinguishable group, to reject the tested hypothesis at various power levels, ranging from .25 to .99. Here, we focus on the conditions required for a power of 80%. For all statistical tests, α was chosen to equal .05.

Patterns of G × E interaction

The power to detect G × E interaction was studied given eight different patterns of interactions (see also Van den Oord 1999; Khoury et al. 1988, 1993). These eight designs are illustrated in Fig. 1 for a dichotomous environment. $Design_{(i)}$ concerns the situation that all effects are zero except the interaction effect for the homozygotes. As a result, the phenotypic means are equal across genotypes in Condition 1, but they are increased or decreased in the homozygotes in the second environmental condition. $Design_{(i = c)}$ represents a variation on $Design_{(i)}$; here the interaction effect in the heterozygotes is also assumed to be non-zero. More specifically, the interaction effect in the heterozygotes is set to equal the effect in the A_1 homozygotes (i.e., 'complete interaction dominance'). The phenotypic mean of the heterozygotes (A_1A_2) therefore equals the phenotypic mean of the group with genotype A_1A_1 in both the first and the second environmental condition. $Design_{(i,e)}$ applies when the environmental main effect and the interaction effect in the homozygotes are greater than zero. $Design_{(i,a)}$ is a function of a non-zero allelic effect (A_1 being the increaser allele), and a non-zero interaction effect. As a result, the phenotypic means of the three genotypic groups differ in Condition 1, and fan out even more in Condition 2. $Design_{(i,a,d)}$ is a variation on $Design_{(i,a)}$, with the difference that complete genetic dominance is present under environmental Condition 1, while the interaction effect in the heterozygotes remains zero. As a consequence, the phenotypic means in the groups with genotype A_1A_1 and A_1A_2 are equal in Condition 1, but differ in Condition 2 due to different interaction effects. In $Design_{(i,a,e)}$, allelic effects, environmental main effects and interaction effects are non-zero, and dominance is absent for all effects. $Design_{(-i,a)}$ is a variation on $Design_{(i,a)}$, where both allelic effects and interaction effects are non-zero. For $Design_{(-i,a)}$, however, the signs of the interaction

Fig. 1 Different patterns of genotype-environment interaction. $Design_{(i)}$: interaction effect for homozygotes, no main effects; $Design_{(i\,=\,c)}$: interaction effect for homozygotes and heterozygotes, with interaction effect heterozygotes equal to effect A_1 homozygotes, no main effects; $Design_{(i,e)}$: interaction effects homozygotes, and main effect environment; $Design_{(i,a)}$: interaction effect homozygotes, and QTL effect; $Design_{(i,a,d)}$: interaction effect homozygotes, and main effect QTL including dominance; $Design_{(i,a,e)}$: interaction effects homozygotes, and main effects environment and QTL; $Design_{(-i,a)}$: reversed interaction effects homozygotes, and main effect QTL; $Design_{(-i,a,e)}$: reversed interaction effects homozygotes, and main effects environment and QTL

Table 5 Number of sib pairs required to detect main effects of QTL and environment, and G × E interaction effects of different effect sizes, in the context of different allele frequencies, and different types of environments (categorical versus continuous) for power of .80 with α = .05 when all other effects are 0

Effect size		Zero effects freely estimated			Zero effects fixed to zero		
		Environment			Environment		
		$b_1 = .2$	$b_1 = .5$	Continuous	$b_1 = .2$	$b_1 = .5$	Continuous
Main effect QTL[a]							
1%	$p = .2$	3,607	1,441	737	600	600	600
2.5%		1,400	561	288	235	235	235
5%		664	266	142	116	116	116
1%	$p = .5$	2,290	921	473	385	385	385
2.5%		890	359	185	151	151	151
5%		422	171	92	75	75	75
Main effect environment[b]							
1%	$p = .2$	3,864	2,468	2,581	600	384	386
2.5%		1,509	964	1,010	228	146	152
5%		716	458	490	110	71	75
1%	$p = .5$	1,157	743	770	600	384	386
2.5%		452	291	302	228	146	152
5%		216	139	147	110	71	75
Interaction[a]							
1%	$p = .2$	1,111	710	725	149	186	278
2.5%		415	266	289	57	70	111
5%		186	120	142	27	33	55
1%	$p = .5$	706	454	463	117	296	378
2.5%		265	171	186	44	70	151
5%		119	77	92	21	33	75

[a] These power calculations assume 2 degrees of freedom

[b] These power calculations assume 1 degree of freedom

effects are reversed, resulting in crossing lines. As a consequence, the group with the highest phenotypic mean in environmental Condition 1, has the lowest phenotypic mean in environmental Condition 2, and vice versa. $Design_{(-i,a,e)}$ resembles $Design_{(-i,a)}$, except that in addition, environmental main effects are non-zero as well.

Results

All tables with results of power analyses (Tables 5–7) show the number of sib pairs required for a power of 80% given α = .05; non-centrality parameters are not reported here but are available online.[6]

To start with, we studied the power to detect specific effects in the situation where all other effects are zero. The simulated data included either a main effect for the QTL, or a main effect of the environment, or a G × E interaction effect (i.e., interaction in the absence of main effects).

Within this context, we studied the effects on the power of allele frequencies, the scale of the environmental measures (dichotomous or continuous), and in the case of a dichotomous environment, the frequencies of the environmental conditions. Knowledge of the power to detect isolated effects of given effect sizes, provides a useful guide to subsequent analyses, where interaction effects are tested in the presence of other effects. Data were simulated such that the specific effects explained 1%, 2.5% or 5% of the variance when $p = .5$ and, if applicable, $b_1 = .5$. Note that these simulations included no population stratification. All between and within parameters could thus be constrained to be equal without loss of fit (given the exact data simulation, this implies $\chi^2 = 0$ for all tests concerning admixture effects). Recall that the background variance (i.e., the variance not due to the marker under study, the environmental measure under study, or their interaction) was simulated to consist of 30% additive polygenic genetic effects (σ_s^2) and 70% environmental effects not-shared by the siblings (σ_u^2). In addition, note that in determining the power to detect the effects of interest, we first fitted the full model, i.e., the

---

[6] http://www.psy.vu.nl/u/s.van.der.sluis.

Table 6 Number of sib pairs required to detect G × E interaction effects in eight different conditions (see Fig. 1) for power of .80 with $\alpha = .05$[a]

	Environment					
	$b_1 = .2$		$b_1 = .5$		Continuous	
	$p = .2$	$p = .5$	$p = .2$	$p = .5$	$p = .2$	$p = .5$
Design$_{(i)}$	410	261	263	169	289	185
Design$_{(i = c)}$	144	175	93	113	102	124
Design$_{(i,e)}$; Design$_{(i,a)}$; Design$_{(i,a,d)}$; Design$_{(i,a,e)}$; Design$_{(-i,a)}$; Design$_{(-i,a,e)}$	410	261	263	169	289	185

[a] All tests assume two degrees of freedom

Table 7 Number of sib pairs required to detect spurious (H1:B = W vs. H0: B ≠ W) and genuine (H1:B ≠ W = 0 vs. H0: B ≠ W ≠ 0) G × E interaction effects in eight different conditions for power of .80 with $\alpha = .05$[a]

	Dichotomous environment		Continuous environment	
	B = W	B ≠ W = 0	B = W	B ≠ W = 0
Design$_{(i)}$	4,745	274	2,543	184
Design$_{(i = c)}$	4,129	192	3,277	126
Design$_{(i,e)}$	3,832	274	4,000	184
Design$_{(i,a)}$	3,470	274	2,071	184
Design$_{(i,a,d)}$	3,658	274	2,146	184
Design$_{(i,a,e)}$	2,914	274	3,045	184
Design$_{(-i,a)}$	6,922	274	3,235	184
Design$_{(-i,a,e)}$	5,318	274	5,563	184

B = W, test for presence of population stratification; B ≠ W = 0, test for significance of within-family interaction effect only

[a] All tests assume two degrees of freedom

model including *all* effects, both zero and non-zero effects. Subsequently we fitted the model in which only the parameters of interest were constrained to zero.

The results are presented in the first three columns of Table 5. With respect to the main effects of the QTL, all tests have 2 degrees of freedom (df), as parameters for both additive and dominance allelic effects are constrained to zero. The power is greatest when $p = q = .5$, and when the environment is a continuous measure. A more uneven distribution of alleles is detrimental to the power to detect allelic effects, as is an uneven distribution of environmental conditions in the case of a dichotomous environmental measure. Interestingly, the distribution of the environmental variable influences the power to detect the QTL main effect, even though association between the environment and the phenotype is absent. These results are consistent with those in Table 6 of Neale et al. (1999).

All tests for environmental main effects have one degree of freedom. As can be seen from the first three columns of Table 5, the power to detect main effects of the environment is somewhat lower when the environmental effects are continuous, compared to a dichotomous environment

with equally distributed conditions. The power to detect environmental main effects is lowest when both the alleles and the environmental conditions are unevenly distributed. Evidently, the allele frequencies influence the power to detect the environmental main effect when the genotypic effects are estimated freely, even though association between the QTL and the phenotype is absent.

All tests for interaction effects are 2 df-tests as both the interaction effects in the homozygotes and the heterozygotes are constrained to zero. The first three columns of Table 5 show that the power to detect interaction effects is greatest when both the allele frequencies and the environmental frequencies are evenly distributed. The power to detect interaction in the context of a continuous environment is only slightly lower.

In conclusion, if alleles are approximately evenly distributed, representative samples of about 200–400 sibling pairs are sufficient to detect main effects for the QTL or the environment, or interaction effects with effect sizes as small as 2.5–5% of the variance.

For illustrative purposes, the last three columns of Table 5 show the sample sizes required to detect the isolated effects with a power of 80% when all zero-effects are actually fixed to zero. As knowledge about which effects are actually zero is usually absent in practice, this is not a realistic situation. It does however illustrate two interesting points. First, the power to detect the effects of interest is much better in the context of a more constrained model. Practically, this implies that the order in which constraints are imposed on the model, may determine the probability to detect effects. This is something to bear in mind when deciding on model fitting procedures. Second, we previously noted that the power to detect effects (e.g., a QTL main effect, an environmental main effect) depends on the distribution of other variables in the model (e.g., the environmental variable, allele frequencies), even when these other variables are not actually associated with the phenotype under study. Naturally, this effect disappears when these zero-effects are excluded from the model.

Subsequently, we examined the power to detect genuine G × E interaction effects in the eight different designs

distinguished by van den Oord (1999, see Fig. 1). For these simulations, parameter values for all non-zero effects were chosen such that in isolation, these effects would explain 2.5%. However, in the case of a dichotomous environment, the presence of other effects influences the percentage of variance explained by the G × E interaction. Using the same parameter values, the actual percentage of variance explained by the G × E interaction varied from 2.1% for Design$_{(i,a,e)}$, to 3.5% for Design$_{(i = c)}$. Also, as is well known in the context of ANOVA analysis, interaction effects can show up as main effects. In this case, the interaction effects show up as allelic main effects when the environment is dichotomous. Consequently, for Design$_{(i)}$ through Design$_{(i,a,e)}$, the main effects of the QTL deviated from zero, with effect sizes ranging from 2.5% (Design$_{(i)}$) to 8.6% (Design$_{(i,a)}$). For Design$_{(-i,a)}$ and Design$_{(-i,a,e)}$ on the other hand, the QTL main effect explained 0% of the variance as the actual effect of the QTL was nullified entirely by the reversed interaction effect. The G × E interaction only turned up as a main environmental effect in Design$_{(i = c)}$. In all cases that the main effect of the environment was specifically modeled to be larger than zero (Design$_{(i,e)}$, Design$_{(i,a,e)}$ and Design$_{(-i,a,e)}$), the effect was slightly lower than 2.5% (2.3, 2.0 and 2.4, respectively) due to the presence of the G × E interaction effect. Again, the background variance was simulated to consist of 30% additive polygenic genetic effects (σ_s^2)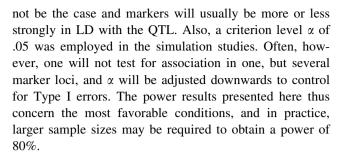, and 70% environmental effects not-shared by the siblings (σ_u^2) in all conditions. These simulations included no population stratification, so all between and within parameters could be constrained to be equal without loss of fit.

The results of these simulations are in Table 6. All tests are 2 df-tests, as both interaction effects for the homozygotes and the heterozygotes are constrained to zero. Note that, irrespective of the allele frequencies, and the measurement scale of the environment, the power to detect interaction effects is higher for Design$_{(i = c)}$ than for Design$_{(i)}$. This makes sense, because the heterozygous group only contributes to the power to detect G × E interaction if the heterozygous interaction effect deviates from zero (Design$_{(i = c)}$ and not Design$_{(i)}$). Note also that the power to detect the interaction in the context of complete interaction dominance (Design$_{(i = c)}$) is higher given $p = .2$ than given $p = .5$. This is because the distribution of the informative genotypic groups is more even in the case of $p = .2$ (i.e., $A_1A_1 + A_1A_2$ vs. A_2A_2: .51:.49) than in the case of $p = .5$ (i.e., $A_1A_1 + A_1A_2$ vs. A_2A_2: .75:.25), which increases the power to detect the effects of interest.

The power to detect the interaction effect is not influenced by the presence or absence of an environmental main effect (Design$_{(i,e)}$ versus Design$_{(i)}$, and Design$_{(i,a,e)}$ and Design$_{(-i,a,e)}$ versus Design$_{(i,a)}$, Design$_{(i,a,d)}$ and Design$_{(-i,a)}$). This is understandable, given that the environmental main

effect only influences the phenotypic means of the genotypic groups, but not the *differences* in phenotypic means between the genotypic groups. The environmental main effect may thus be viewed as a constant, which does not influence the power to detect interaction.

The presence or absence of a main effect of the QTL also has no influence on the power to detect G × E interaction. (To assure that this finding was not due to the size of the allelic effect, additional analyses including a larger allelic effect, explaining 10% and 20% of the variance rather than 2.5%, were run, which showed similar results.)

Finally, we studied the power to detect population stratification with respect to the interaction component of the model. As described above, we mixed two subsamples of equal proportions, which differed with respect to allele frequencies ($p_A = .7$, $p_B = .3$), and environmental distribution (in case of a dichotomous environmental measure, $b_{A1} = .3$, $b_{B1} = .7$; in case of a continuous environmental measure, $\mu_{envA} = 0$, $\mu_{envB} = 2$), choosing phenotypic means of the subsamples such that the admixture accounted for 20% of the total phenotypic variance in the combined sample. When these admixture proportions were used to simulate data in which the actual effects (allelic, environmental and interaction) were zero, spurious allelic, environmental, and interaction effects were observed in the combined sample due to the admixture. For the dichotomous environment, the between family effects deviated from the within family effects, with the stratification effect being largest for the allelic effects ($N = 184$ for 80% power), intermediate for the environmental effects ($N = 1,465$ for 80% power), and smallest for the interaction effects ($N = 7,028$ for 80% power). For the continuous environment, the between family effects also deviated from the within family effects, with the stratification effect being largest for the environmental effects ($N = 278$ for 80% power), medium for the allelic effects ($N = 576$ for) and smallest for the interaction effects ($N = 3,233$ for 80% power). It is clear that very large numbers of sib pairs are required to detect stratification effects in the interaction component. It is also noteworthy that the allele frequencies in the subsamples determine how the spurious G × E interaction is expressed. With the present admixture settings ($p_A = .7$, $p_B = .3$, i.e., contrasting allele frequencies), spurious G × E is only apparent with respect to the interaction parameter for the heterozygous group, while the interaction parameter for the homozygous group obtained in the combined sample does not deviate from its actual value in the subsamples. However, if the allele frequencies in the subsamples are not contrasting (e.g., $p_A = .3$, $p_B = .5$), both interaction parameters for the heterozygous and homozygous groups are informative about spurious interaction.

Given population stratification, we again considered the eight different interaction designs (Fig. 1) to study (a) the power to detect stratification effects with respect to the interaction component of the model (tests with 2 df as both homozygote and heterozygote interaction effects are constrained to be equal within and across families) and (b) the power to detect genuine interaction effects (tests with 2 df as both homozygote and heterozygote interaction effects are constrained to be zero within families, while the between-family effects are freely estimated). For all conditions, the background variance was again simulated to consist of 30% additive polygenic genetic effects (σ_s^2), and 70% environmental effects not-shared by the siblings (σ_u^2).

The results are presented in Table 7. Besides confirming the observation that prohibitively large samples of sib pairs are required to detect spurious interaction (B = W), it is shown that the power to detect the spurious interaction due to population admixture varies across the eight differentiated subtypes. Overall, the power to detect spurious interaction is somewhat higher when the environment is continuous in nature, but the sample sizes required to detect stratification with respect to the interaction effect are prohibitively large in all simulated scenarios.

An indication of the power to detect the genuine interaction effect is obtained by freely estimating the between-family effect, while the within-family effect is constrained to be zero (B ≠ W = 0). The results in Table 7 show that the power to detect G × E effects on the means is about as large as one would expect given the previous results presented in Table 6, and the distribution of the genotypes in the mixed population (i.e., freq(A_1A_1) = $(.7^2 + .3^2)$/2 = .29; freq(A_1A_2) = $((2 * .3 * .7) + (2 * .3 * .7))$/2 = .42; freq($A_2A_2$) = $(.3^2 + .7^2)$/2 = .29).

Discussion

In this paper, the family-based association design was extended to include G × E interaction effects and environmental main effects. Power calculation showed that allele frequencies, and characteristics of environment (e.g., measurement level, and in the case of a categorical environment, the frequencies of the environmental conditions) affect the power to detect G × E interaction. Relatively small interaction effects, explaining 2.5–5% of the phenotypic variance in the total sample, can be detected with reasonably small sample sizes (200–400 sib pairs, respectively), if alleles are evenly distributed. The power to detect main effects and interaction effects generally is reasonable, particularly when all zero-effects are removed from the model first.

Throughout the paper, we assumed that the marker locus under study is the actual QTL. In practice, this will often

not be the case and markers will usually be more or less strongly in LD with the QTL. Also, a criterion level α of .05 was employed in the simulation studies. Often, however, one will not test for association in one, but several marker loci, and α will be adjusted downwards to control for Type I errors. The power results presented here thus concern the most favorable conditions, and in practice, larger sample sizes may be required to obtain a power of 80%.

Modeling measured environmental effects in association studies is standard (e.g., Caspi et al. 2002, 2003; Foley et al. 2004; Huizinga et al. 2006; Lasky-Su et al. 2007; Yaffe et al. 2000). The use of the extended sib pair model in such studies has the advantage of controlling for population stratification, and excluding spurious main effects of the QTL and the environment, and, given sufficiently large sample size, spurious interaction effects. This extension can be implemented readily in packages such as Mx (Appendices I and II), or, in case of a categorical environmental factor, in SPSS (Beem and Boomsma 2006).

Some caveats are in order. First, it has often been shown that non-normality can result in spurious interaction effects (e.g., Boomsma and Martin 2002; Martin 1999; Purcell 2002; van den Berg et al. 2007; van der Sluis et al. 2006). However, the actual presence of G × E can also render the distribution non-normal (e.g., Purcell 2002; van der Sluis et al. 2006), resulting in the problem that non-normality of the data can either *indicate* the presence of G × E (i.e., G × E being the source of the non-normality) or *mimic* the presence of G × E (i.e., non-normality due to e.g. censoring or poor scaling of the phenotypic measure). The model presented here is equally susceptible to this phenomenon.

Although there is no ready solution to this problem, researchers should at least investigate alternative reasons for the non-normality of their data than the presence of G × E (e.g., poor measurement scale, selective sampling, etc.). As has been argued before (e.g., Martin 1999; van der Sluis et al. 2006), transformation of the data is no solution as it will remove both spurious and genuine G × E from the data.

Here we presented a model with measured genotypes and a measured environment. If these measured variables are indeed the ones involved in the G × E interaction, and thus the ones causing the heteroscedasticity, then accounting for these measures (i.e., modeling their effects) should render the remaining variance (as summarized in Eq. 4) homoscedastic. In a previous paper (van der Sluis et al. 2006), marginal maximum likelihood showed to be useful in the detection of heteroscedasticity. If heteroscedasticity is present before modeling the genotypic and environmental effects, but absent when these effects are controlled for, then this can be taken to indicate that the

heteroscedasticity was due to the interaction between the locus and environment under study. Yet, if the heteroscedasticity is still present, this can mean (a) that the heteroscedasticity is caused by scaling problems in the instrument used to measure the phenotype, or (b) that G × E interaction is present but the genes and environment controlled for are not the ones involved in this interaction, or only 'rough approximations' of the actual gene/environment involved (e.g., a poorly designed environmental measure, or a marker that is only slightly in LD with the actual QTL). Important in this context is an issue discussed by Eaves (1984) in the light of plant studies, that the genes that control average performance (i.e., main effects) may not be the genes that control the sensitivity to the environment (i.e., the genes involved in the interaction effect, giving rise to the heteroscedasticity, see Berg et al. 1989, for a similar distinction between 'level' and 'variability' genes). Within a design as discussed here, where both genes and environment are measured entities, level and variability genes can be distinguished. This distinction may be important in understanding the biological basis of the G × E interaction.

Second, throughout this paper, we assumed that the environmental measure is independent of genotype and family membership. Using so-called family-level environmental measures, i.e., environmental measures that are by definition equal for all siblings within a family, is problematic in the sib pair design discussed here, because the decomposition in between and within family environmental effects (e_b vs. e_w) depends on siblings that are discordant with respect to the environmental measure (see Tables 2–4). The use of family-level environmental measures thus excludes the possibility to test for stratification effects in family-level environmental components, such as socioeconomic status, divorce status of the parents, domestic violence, and loss of a parent. However, stratification with respect to the allelic effects and the interaction effect can still be controlled for, and one can still test the significance of the interaction effect, and allelic and environmental main effects. In this context it is important to note that there is ample debate about whether genuine family-level environmental measures actually exist. For example, the fact that divorce status of the parents is necessarily equal for siblings from the same family does not necessarily imply that this event has similar effects on

the siblings, or is experienced in exactly the same manner by all siblings. We refer to Turkheimer et al. (2005) for an extensive discussion of this subject.

Third, the model presented so far does not account for the presence of gene-environment correlation (rGE). rGE represents the genetic liability to experience different environmental events, or the genetic control of exposure to different environments (e.g., Kendler and Eaves 1986; Plomin et al. 1977). Genetic factors have been found to substantially influence individual differences in, for example, the likelihood of experiencing stressful life events, lack of social support, participation in leisure activities, martial status, and age of first sexual intercourse (see Rutter and Silberg 2002 for a review). The finding that so many diverse 'environmental' measures are under genetic control, suggests that the present sib pair model may prove to be of limited use. Extension of this model to include the possibility to test and account for rGE is therefore desirable. For now, we advise researchers to test for the presence of rGE before they proceed. For instance, one can test whether the genotypic groups differ with respect to their environmental mean (ANOVA), or, if the environmental measure is categorical, with respect to the distribution of subjects across environmental conditions (χ^2 test for equal frequencies). If differences with respect to the environmental measure are absent, one can proceed with the extended sib pair model as presented here.

Gene by environment interaction studies are relatively new and such studies are often characterized by difficulties concerning measurement and modeling (e.g., Eaves 2006). In general however, researchers seem to agree that studies aimed at revealing the sources of individual differences in specific qualities need to take G × E interaction into account, in order to arrive at a full account of individual differences (e.g., Caspi et al. 2006, Moffitt et al. 2005, 2006). Tests for G × E interaction are thus likely to become standard in future (association) studies.

Appendix 1

```
! Mx script for the conduction of sib-pair association
! models including the effects of measured genes (G), measured categorical environment (E)
! and the GxE interaction.
! Note that alleles need to be coded 1,2,… n,
! and environmental conditions need to be coded 1,2 … k

#define n 2          ! number of alleles
#define env 2        ! number of environmental states
#define nvar 1       ! univariate
#define nsib 2       ! sibshipsize
#define ngroups 6  ! G1= precalc group,
                     ! G2=data group
                     ! G3 - G6 constraint groups

G1: calculation group between and within effects
data calc Ngroups=6
Begin Matrices;
  A full 1 n free ;    ! additive genetic effects within
  C full 1 n free ;    ! additive genetic effects between
  D sdiag n n free ;   ! dominance deviations within
  F sdiag n n free ;   ! dominance deviations between

  H full 1 env free ;  ! environmental effect within
  J full 1 env free ;  ! environmental effect between

  E full 1 n free ;    ! int effect homozyg within
  G full 1 n free ;    ! int effect homozyg between

  S sdiag n n free ;   ! int effect  heterozyg within
  P sdiag n n free ;   ! int effect heterozyg between

  T full 1 nvar free;  ! grand mean

  I unit 1 n ;         ! unit vector

  Q lower nvar nvar free ! familial variance
  R lower nvar nvar free ! non-familial variance

End Matrices;

Fi H 1 1 J 1 1    ! first elements of H and J are fixed to zero
Va 0 H 1 1 J 1 1

Begin Algebra;
  W = ((A'@I)+(A@I')) + (D + D');        ! total within QTL effect
  B = ((C'@I)+(C@I')) + (F + F');        ! total between QTL effect
  X = ((E'@I)+(E@I')) + (S + S');        ! total within interaction effect
  Y = ((G'@I)+(G@I')) + (P + P');        ! total between interaction effect
End Algebra;

st .2 all
end

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

G2: datagroup
data NInput=8
rect file=file.dat

Labels
! all refer to allele so all11 is allele 1 sib 1 and all12 is allele 2 sib 1, etc.
! env1 and env2 are the environmental conditions of sib1 and 2 respectively
! and pheno1 and pheno2 are the continuous phenotypic scores of sib1 and 2 respectively

all11 all12 env1 pheno1 all21 all22 env2 pheno2

select
all11 all12 env1 pheno1
all21 all22 env2 pheno2;

definition_variables
all11  all12  env1 all21  all22  env2 ;

begin matrices;
  I unit nsib nsib fix ;
  K full 1 4 fix ;   ! first and second allele sib1, times 2, coded as 1 and 2
  L full 1 4 fix ;   ! first and second allele sib2, times 2, coded as 1 and 2
  U full 1 4 fix ;   ! env condition sib1, coded as 1 and 2
  O full 1 4 fix ;   ! env condition sib2, coded as 1 and 2

  S full 1 1 fix ;   ! contains nsib
  M full 1 env fix ; ! contains 0 and 1

  B computed n n = B1 ;    ! spurious and genuine genotypic effects
  W computed n n = W1;     ! genuine genotypic effects
  Y computed n n = Y1 ;    ! spurious and genuine interaction effects
  X computed n n = X1;     ! genuine interaction effects

  J full 1 env = J1 ;   ! spurious and genuine environmental effects
  H full 1 env = H1;    ! genuine environmental effects
  T full 1 nvar = T1 ;  !grand mean
  Q lower nvar nvar = Q1 ; ! familial variance
  R lower nvar nvar = R1 ; ! non-familial variance

End Matrices ;
```

```
matrix S nsib      ! sibship size = 2
                   ! Note that with varying sibship sizes, one can read in sibship-size
                   ! as a variable and use it as a definition variable
matrix M 0 1       ! weights matrix
matrix K 1 1 1 1   ! provide initial values otherwise matrix with 0's
matrix L 1 1 1 1   ! which is problematic in part-statement
matrix U 1 1 1 1
matrix O 1 1 1 1

Specify K all11 all12 all11 all12   ! genotypes sib1 to be used for \part
Specify L all21 all22 all21 all22   ! genotypes sib2 to be used for \part
Specify U 0 env1 0 env1    ! env sib 1
Specify O 0 env2 0 env2    ! env sib 2

Begin Algebra;
! between family effect; genetic part, env part and interaction part
  V = ( \part(B,K) + \part(B,L) + \part(J,U) + \part(J,O) +
      \part(Y,K)@\part(M,U) + \part(Y,L)@\part(M,O) ) % S ;

! within family effect; genetic part, env part and interaction part
  D = ( \part(W,K) + \part(W,L) + \part(H,U) + \part(H,O) +
      \part(X,K)@\part(M,U) + \part(X,L)@\part(M,O)  ) % S ;

End Algebra;

! mean is grand mean + between family effect which is equal for all sibs
! + within family effect which are deviations between sibs

Means T + V +((\part(W,K) +  \part(H,U) + \part(X,K)@\part(M,U))-D) |
       T + V +((\part(W,L) +  \part(H,O) + \part(X,L)@\part(M,O))-D) ;

Covariance
   Q*Q' + R*R' | Q*Q' _
         Q*Q' | Q*Q' + R*R' ;
end

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

G3: Constraint sum allelic effects = 0
constraint
Begin Matrices ;
  A full 1 n = A1 ;
  O zero 1 1 ;
End Matrices ;
Begin Algebra ;
  B = \sum(A) ;
End ALgebra ;
constraint B=O ;
option NO-output
End

G4: Constraint sum allelic effects = 0
constraint
Begin Matrices ;
  C full 1 n = C1 ;
  O zero 1 1 ;
End Matrices ;
Begin Algebra ;
  B = \sum(C) ;
End ALgebra ;
constraint B=O;
option NO-output
End

G5: Constraint sum interaction effects homozyg = 0
constraint
Begin Matrices ;
  E full 1 n = E1 ;
  O zero 1 1 ;
End Matrices ;
Begin Algebra ;
  B = \sum(E) ;
End ALgebra ;
constraint B=O;
option NO-output
End

G6: Constraint sum interaction effects homozyg = 0
constraint
Begin Matrices ;
  G full 1 n = G1 ;
  O zero 1 1 ;
End Matrices ;
```

```
Begin Algebra ;
  B = \sum(G) ;
End ALgebra ;
constraint B=O;
option NO-output
option multiple issat
End
```

!!!

```
save dich.mxs

! Test for spurious allelic association
get dich.mxs
eq A 1 1 1 C 1 1 1   ! because of constraint groups G3 and G4, A 1 1 2
eq D 1 2 1 F 1 2 1   ! and C 1 1 2 will be equal as well
end

! Test for spurious environmental association
get dich.mxs
eq H 1 1 2 J 1 1 2
end

! Test for spurious interaction association
get dich.mxs
eq E 1 1 1 G 1 1 1   ! because of constraint groups G5 and G6, E 1 1 2
eq P 1 2 1 S 1 2 1   ! and G 1 1 2 will be equal as well
end
```

```
! If the equality constraints tested above are tenable, then the significance of the allelic effects,
! the environmental effects, and the interaction effects can be tested in the context of the
! model in which these equality constraints are implemented.  One can then test the
! significance of allelic effects, environmental effects and interaction effects across families,
! i.e., both between and within effects are fixed to zero, and the deterioration of the fit of the
! model is indicative of the significance of the effects.
! If the equality constraints are not tenable, then only the within family effects (i.e., the genuine
! effects) are fixed to zero.
```

```
get dich.mxs
eq A 1 1 1 C 1 1 1   ! because of constraint groups G3 and G4, A 1 1 2
eq D 1 2 1 F 1 2 1   ! and C 1 1 2 will be equal as well
eq E 1 1 1 G 1 1 1   ! because of constraint groups G5 and G6, E 1 1 2
eq P 1 2 1 S 1 2 1   ! and G 1 1 2 will be equal as well
end
save equal.mxs

! Test for main effect QTL
get equal.mxs
Fi A 1 1 1 D 1 2 1   ! fix with-family allelic effects to zero
Va 0 A 1 1 1 D 1 2 1
end

! Test for main effect environment
get equal.mxs
Fi H 1 1 2           ! fix with-family environmental effects to zero
Va 0 H 1 1 2
end

! Test for main effect QTL

get equal.mxs
Fi E 1 1 1 S 1 2 1   ! fix with-family interaction effects to zero
Va 0 E 1 1 1 P 1 2 1
end
```

Appendix 2

```
! Mx script for the conduction of sib-pair association
! models including the effects of measured genes (G), measured continuous environment (E)
! and the GxE interaction.
! Note that alleles need to be coded 1,2,… n,

#define n 2          ! number of alleles
#define nvar 1       ! univariate
#define nsib 2       ! sibshipsize
#define ngroups 6    ! G1= precalc group,
                     ! G2=data group sibs
                     ! G3 - G6 constraint groups

G1: calculation group between and within effects
data calc Ngroups=6
Begin Matrices;
  A full 1 n free ;   ! additive genetic effects within
  C full 1 n free ;   ! additive genetic effects between
  D sdiag n n free ;  ! dominance deviations within
  F sdiag n n free ;  ! dominance deviations between

  H full 1 1 free ;   ! environmental effect within
  J full 1 1 free ;   ! environmental effect between

  E full 1 n free ;   ! int effect homozyg within
  G full 1 n free ;   ! int effect homozyg between

  I unit 1 n ;        ! unit vector

  S sdiag n n free ;  ! int effect  heterozyg within
  P sdiag n n free ;  ! int effect heterozyg between
  T full 1 nvar free; ! grand mean

! variance components
  Q lower nvar nvar free ! familial variance
  R lower nvar nvar free ! non-familial variance
End Matrices;
```

```
Begin Algebra;
  W = ((A'@I)+(A@I')) + (D + D');   ! total within QTL effect
  B = ((C'@I)+(C@I')) + (F + F');   ! total between QTL effect
  X = ((E'@I)+(E@I')) + (S + S');   ! total within interaction effect
  Y = ((G'@I)+(G@I')) + (P + P');   ! total between interaction effect
End Algebra;

st .3 all
end
```

!!

```
G2: datagroup  ! sibs
data Nlinput=8
rect file=file.dat
! all refer to allele so all11 is allele 1 sib 1 and all12 is allele 2 sib 1, etc.
! env1 and env2 are the environmental conditions of sib1 and 2 respectively
! and pheno1 and pheno2 are the continuous phenotypic scores of sib1 and 2 respectively
Labels
all11 all12 env1 pheno1
all21 all22 env2 pheno2

select
all11 all12 env1 pheno1
all21 all22 env2 pheno2 ;

definition_variables
all11 all12 env1
all21 all22 env2 ;

begin matrices;
  I unit nsib nsib fix ;
  K full 1 4 fix ;   ! first and second allele sib1, times 2, coded as 1 and 2
  L full 1 4 fix ;   ! first and second allele sib2, times 2, coded as 1 and 2

  U full 1 1 fix ;   ! env condition sib1, continuous
  O full 1 1 fix ;   ! env condition sib2, continuous

  S full 1 1 fix ;   ! contains sibship

  B computed n n = B1 ;   ! spurious and genuine genotypic effects
  W computed n n = W1;    ! genuine genotypic effects
  Y computed n n = Y1 ;   ! spurious and genuine interaction effects
  X computed n n = X1;    ! genuine interaction effects
  J full 1 1 = J1 ;       ! spurious and genuine environmental effects
  H full 1 1 = H1;        ! genuine environmental effects
  Q lower nvar nvar = Q1 ;
  R lower nvar nvar = R1 ;
  T full 1 nvar = T1;

End Matrices ;

matrix S nsib   ! sibship
                ! Note that with varying sibship sizes, one can read in sibship-size
                ! as a variable and use it as a definition variable
matrix K 1 1 1 1   ! provide initial values otherwise matrix with 0's
matrix L 1 1 1 1   ! which is problematic in part-statement
matrix U 1
matrix O 1

Specify K all11 all12 all11 all12   ! genotypes sib1 to be used for \part
Specify L all21 all22 all21 all22   ! genotypes sib2 to be used for \part

Specify U env1   ! env sib 1
Specify O env2   ! env sib 2

Begin Algebra;
! between family effect; genetic part, env part and interaction part
  V = ( \part(B,K) + \part(B,L)
      + U@J + O@J
      + U@(\part(Y,K)) + O@(\part(Y,L))  ) % S ;

! within family effect; genetic part, env part and interaction part
  D = ( \part(W,K) + \part(W,L)
      + U@H + O@H
      + U@(\part(X,K)) + O@(\part(X,L))  ) % S ;

End Algebra;

Means T + V +((\part(W,K) +  U@H + U@(\part(X,K)) )-D) |
      T + V +((\part(W,L) +  O@H + O@(\part(X,L)) )-D)  ;

Covariance
  Q*Q' + R*R' | Q*Q' _
          Q*Q' | Q*Q' + R*R' ;
end
```

!!

```
G3: Constraint sum allelic effects = 0
constraint
Begin Matrices ;
  A full 1 n = A1 ;
  O zero 1 1 ;
End Matrices ;
Begin Algebra ;
  B = \sum(A) ;
End ALgebra ;
constraint B=O ;
option NO-output
End


G4: Constraint sum allelic effects = 0
constraint
Begin Matrices ;
  C full 1 n = C1 ;
  O zero 1 1 ;
End Matrices ;
Begin Algebra ;
  B = \sum(C) ;
End ALgebra ;
constraint B=O;
option NO-output
End


G5: Constraint sum interaction effects homozyg = 0
constraint
Begin Matrices ;
  E full 1 n = E1 ;
  O zero 1 1 ;
End Matrices ;
Begin Algebra ;
  B = \sum(E) ;
End ALgebra ;
constraint B=O;
option NO-output
End


G6: Constraint sum interaction effects homozyg = 0
constraint
Begin Matrices ;
  G full 1 n = G1 ;
  O zero 1 1 ;
End Matrices ;
Begin Algebra ;
  B = \sum(G) ;
End ALgebra ;
constraint B=O;
option NO-output
option multiple issat
End

!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
save cont.mxs

! Test for spurious allelic association
get cont.mxs
eq A 1 1 1 C 1 1 1   ! because of constraint groups G3 and G4, A 1 1 2
eq D 1 2 1 F 1 2 1   ! and C 1 1 2 will be equal as well
end

! Test for spurious environmental association
get cont.mxs
eq H 1 1 1 J 1 1 1
end

! Test for spurious interaction association
get cont.mxs
eq E 1 1 1 G 1 1 1   ! because of constraint groups G5 and G6, E 1 1 2
eq P 1 2 1 S 1 2 1   ! and G 1 1 2 will be equal as well
end
```

```
! If the equality constraints tested above are tenable, then the significance of the allelic effects,
! the environmental effects, and the interaction effects can be tested in the context of the
! model in which these equality constraints are implemented.  One can then test the
! significance of allelic effects, environmental effects and interaction effects across families,
! i.e., both between and within effects are fixed to zero, and the deterioration of the fit of the
! model is indicative of the significance of the effects.
! If the equality constraints are not tenable, then only the within family effects (i.e., the genuine
! effects) are fixed to zero.

get cont.mxs
eq A 1 1 1 C 1 1 1   ! because of constraint groups G3 and G4, A 1 1 2
eq D 1 2 1 F 1 2 1   ! and C 1 1 2 will be equal as well
eq H 1 1 1 J 1 1 1
eq E 1 1 1 G 1 1 1   ! because of constraint groups G5 and G6, E 1 1 2
eq P 1 2 1 S 1 2 1   ! and G 1 1 2 will be equal as well
end

save equal.mxs

! Test for main effect QTL
get equal.mxs
Fi A 1 1 1   D 1 2 1
Va 0 A 1 1 1 D 1 2 1
end

! Test for main effect environment
get equal.mxs
Fi H 1 1 1
Va 0 H 1 1 1
end

! Test for main effect QTL
get equal.mxs
Fi E 1 1 1   S 1 2 1
Va 0 E 1 1 1   P 1 2 1
end
```

References

Abecasis GR, Cardon LR, Cookson WOC (2000a) A general test of association for quantitative traits in nuclear families. Am J Human Genet 66:279–292

Abecasis GR, Cookson WOC, Cardon LR (2000b) Pedigree tests of transmission disequilibrium. Eur J Human Genet 8:545–551

Azzelini A (1996) Statistical inference based on the likelihood. Chapman and Hall, London

Beem AL, Boomsma DI (2006) Implementation of a combined association-linkage model for quantitative traits in Linear Mixed Model procedures of statistical packages. Twin Res Human Genet 9(3):325–333

Berg K, Kondo I, Drayna D, Lawn R (1989) 'Variability gene' effect of cholesteryl ester transfer protein (CEPT) genes. Clin Genet 35:437–445

Bollen KA, Stine R (1993) Bootstrapping goodness of fit measures in structural equation models. In: Bollen KA, Long JS (eds) Testing structural equation models. Sage, Newbury Park, CA, pp 111–135

Boomsma DI, Martin NG (2002) Gene-environment interactions. In: D'haenen H, den Boer JA, Willner P (eds) Biological psychiatry. John Wiley & Sons, London, pp 181–187

Cardon LR, Bell JI (2001) Association study designs for complex diseases. Nat Genet 2:91–99

Caspi A, McClay J, Moffitt TE, Mill J, Martin J, Craig IW, Taylor A, Poulton R (2002) Role of genotype[e in the cycle of violence in maltreated children. Science 297(5582):851–854

Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H et al (2003) Influence of life stress on depression: Moderation by a polymorphism in the 5-htt gene. Science 301(5631):386–389

Caspi A, Moffitt TE (2006) Gene-environment interactions in psychiatry: joining forces with neuroscience. Nat Neurosci 7:583–590

Dolan CV, van der Sluis S, Grasman R (2005) A note on normal theory power calculation in SEM with data missing completely at random. Struct Equat Model 12(2):245–262

Eaves LJ (1984) The resolution of genotype-environment interaction in segregation analysis of nuclear families. Genet Epidemiol 1:215–228

Eaves LJ (2006) Genotype × environment interaction in psychopathology: fact or artifact? Twin Res Human Genet 9(1):1–8

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th edn. Pearson Education Ltd., Essex, England

Fisher RA (1918) The correlation between relatives on the supposition of Mendelian inheritance. Trans R Soc Edinb 52:399–433

Foley DL, Eaves LJ, Wormley B, Silberg JL, Maes HH, Kuhn J, Riley B (2004) Childhood adversity, monoamine oxidase A genotype, and risk for conduct disorder. Arch Genet Psychiatry 61:738–744

Fulker DW, Cherny SS, Sham PC, Hewitt JK (1999) Combined linkage and association sib-pair analysis for quantitative traits. Am J Human Genet 64:259–267

Huizinga D, Haberstick BC, Smolen A, Menard S, Young SE, Corley RP, Stallings MC, Grotpeter J, Hewitt JK (2006) Childhood maltreatment, subsequent antisocial behavior, and the role of monoamine oxidase A genotype. Biol Psychiatry 60:677–683

Kendler K, Eaves L (1986) Models for the joint effect of genotype and environment on liability to psychiatric illness. Am J Psychiatry 143:279–289

Khoury MJ, Adams MJ, Flanders WD (1988) AN epidemiologic approach to ecogenetics. Am J Human Genet 42:89–95

Khoury MJ, James LM (1993) Population and family relative risks of disease associated with environmental factors in the presence of gene-environment interaction. Am J Epidemiol 137:1241–1250

Lasky-Su J, Faraone SV, Lange C, Tsuang MT, Doyle AE, Smoller JW, Laird NM, Biedermand J (2007) A study of how socioeconomic status moderates the relationship between SNPs encompassing BDNF and ADHD symptom count in ADHD families. Behav Genet 37:487–497

Mather K, Jinks JL (1977) Introduction to biometrical genetics. Chapman and Hall, London

Martin N (1999) Gene-environment interaction and twin studies. In: Spector TD, Snieder H, MacGregor AJ (eds) Advances in twin and sib-pairanalysis. Greenwich Medical Media, London

Moffitt TE, Caspi A, Rutter M (2005) Strategy for investigating interactions between measured genes and measured environments. Arch Genet Psychiatry 62:473–481

Moffitt TE, Caspi A, Rutter M (2006) Measured gene-environment interactions in psychopathology: concepts, research strategies, and implications for research, intervention, and public understanding of genetics. Perspect Psychol Sci 1(1):5–27

Neale MC, Cherny SS, Sham PC, Whitfield JB, Heath AC, Birley AJ, Martin NG (1999) Distinguishing population stratification from genuine allelic effects with Mx: association of ADH2 with alcohol consumption. Behav Genet 29(4):233–243

Neale MC, Boker SM, Xie G, Maes HH (2003) Mx: statistical modeling, 6th edn. Department of Psychiatry, Richmond, VA

Plomin R, DeFries JC, Loehlin JC (1977) Genotype-environment interaction and correlation in the analysis of human behavior. Psychol Bull 84:309–322

Posthuma D, de Geus EJC, Boomsma DI, Neale MC (2004) Combined linkage and association tests in Mx. Behav Genet 34(2):179–196

Purcell S (2002) Variance components models for gene-environment interaction in twin analysis. Twin Res 5:554–571

Rutter M, Silberg JL (2002) Gene-environment interplay in relation to emotional and behavioral disturbance. Ann Rev Psychol 53:463–490

Saris WE, Satora A (1993) Power evaluations in structural equation models. In: Bollen KA, Long JS (eds) Testing structural equation models. Sage, Newbury Park, CA, pp 181–204

Turkheimer E, D'Onofrio BM, Meas HH, Eaves LJ (2005) Analysis and interpretation of twin studies including measures of the shared environment. Child Dev 76(6):1217–1233

Van den Berg SM, Glas CAW, Boomsma DI (2007) Variance decomposition using an IRT measurement model. Behav Genet 37:604–616

Van den Oord EJCG (1999) Method to detect genotype-environment interactions for quantitative trait loci in association studies. Am J Epidemiol 150(11):1179–1187

Van der Sluis S, Dolan CV, Neale MC, Posthuma D (2006) Detecting genotype-environment interaction in monozygotic twin data: comparing the Jinks and Fulker test and a new test based on marginal maximum likelihood estimation. Twin Res Human Genet 9(3):377–392

Van der Sluis S, Dolan CV, Neale MC, Posthuma D (2008) Power calculations using exact data simulation: a useful tool for genetic study designs. Behav Genet 38:202–211

Yaffe K, Haan M, Byers A, Tangen C, Kuller L (2000) Estrogen, apoe, and cognitive decline: evidence for gene-environment interaction. Neurology 54:1949–1954