Developmental genetics of psychopathology

by

Michel Guillaume Nivard

Reading committee:

Prof.dr. Dale R. Nyholt

Prof.dr. Denny Borsboom

Dr. Marleen H.M. de Moor

Prof.dr. Hein Putter

Dr. Jeffrey C. Glennon

Dr. Ramon J.L. Lindauer

Paranimphs:

Dr. Abdel Abdellaoui

Dr. Sanja Franić

Cover image: Sylvette, by Pablo Picasso (1954)

Printed by: Ridderprint BV

ISBN: 978-94-6299-152-1

VRIJE UNIVERSITEIT

Developmental genetics of psychopathology

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan

de Vrije Universiteit Amsterdam,

op gezag van de rector magnificus

prof.dr. F.A. van der Duyn Schouten,

in het openbaar te verdedigen

ten overstaan van de promotie commissie

Faculteit der Gedrags- en Bewegingswetenschappen

op donderdag 1 oktober 2015 om 11.45 uur

in de aula van de universiteit,

De Boelelaan 1105

door

Michel Guillaume Nivard

geboren te Amsterdam

promotoren:	prof.dr. D.I. Boomsma
	prof.dr. C.V. Dolan
copromotor:	dr. C.M. Middeldorp

Chapter 1: Introduction: Behavior Genetics

The aim of the first chapter of this thesis is to review the current practice of behavior genetics. We define behavior genetics broadly as the study of the role of genetic and environmental variables in phenotypic individual differences. We employ the term phenotype, as this is the accepted term in genetics, for any measured (observed) characteristic. Alternatively we use the term trait (but without the association of stability, as in trait psychology). While behavior genetics often is associated with psychology, we define behavior more broadly to encompass any human phenotype, which is directly or indirectly relevant to behavior, and which is characterized by appreciable individual differences. As such, this definition of behavior genetics encompasses genetic epidemiology, psychiatric genetics, and psychogenetics. We limit this chapter to human behavioral genetics, but acknowledge the common foundation of animal and human genetics in underlying statistical and biological models and methods,[1-4] and the relevance of animal studies to human behavior.[5] For an example of findings from animal behavior genetics informing human behavior genetics, see chapter 6 of this dissertation. Finally, we assume that the phenotypes of interest are complex, i.e., subject to the effects of many genes of small effect and many environmental effects, thereby excluding phenotypes subject to a major effect of one or at most a few genes (e.g., Mendelian or monogenic disease, such as Huntington disease).

Taking a bird's eye view of the recent history of behavior genetics, we note that from the late 1970's to the late 1990's, the field was dominated by family, adoption and twin studies, in which the contributions of unmeasured (or "latent") genetic and environmental variables to phenotypic variation were inferred from the phenotypic resemblance among family members.[6;7] This work established beyond reasonable doubt the importance of genetic and environmental factors in a variety of human phenotypes,[8; 9] and set the stage for addressing follow-up issues, including the identification of the causal genetic variants underlying the phenotypic variation. However, family and twin designs continue to be applied, and chapter 2 in my thesis shows how longitudinal twin data can inform on genetic stability across the lifespan.

From the late 1990's onwards, the scope of behavioral genetics was expanded by the advent of high throughput genotyping technologies, which enabled researchers to measure actual genetic variants at a genome-wide level. Initially, highly polymorphic genetic variants (i.e., comprising many alleles; e.g., microsatellite markers) were used in linkage analyses of complex behavioral phenotypes (see for

example: Wray et al.[10]). These variants were relatively few in number (say, hundreds), and served the relatively modest objective of identifying chromosomal regions of interest (i.e., loci harboring causal genetic variants). From the early 2000's, advances in genotyping technology enabled researchers to measure single nucleotide polymorphisms (SNPs; diallelic genetic variants) in the hundreds of thousands (presently in the millions), and conduct genome-wide association studies (GWAS). An excellent overview of GWAS by Visscher et al. appeared in 2012.[11] In linkage studies, a within family design is employed and biologically related family members are required. In association studies, the measured genetic variant is tested directly by regression of the phenotype on the variant in a sample of individuals, who can be related or unrelated. Association studies largely superseded linkage studies, because they are generally more powerful and do not require samples of related individuals, which may be difficult to collect. In addition, association studies may identify causal variants, or the regions harboring a causal genetic variant, which are appreciably smaller than those identified in linkage analysis.

The availability of vast amounts of measured genetic variants, and their use in GWAS, has also given rise to new statistical techniques. These include techniques to estimate the contribution of the entire set of measured SNPs to the phenotypic variation, without identifying the association of the phenotype with any individual SNP.[12-15] A second technique to test genetic association of a subset of SNPs and a given phenotype is by aggregating a subset of SNPs into a weighted polygenic score [16-18] In chapter 5 of this thesis, polygenic scores based on the most recent findings for schizophrenia[19] are analyzed (see: de Zeeuw et al.[20] and Groen-Blokhuis et al.[21] for other recent applications of this technique).

The setup of the present review follows the historical outline above. We first discuss twin and family methods, as applied to complex quantitative phenotypes. We discuss genetic linkage and genetic association analysis, with the emphasis on the latter. Finally, we discuss the use of polygenic scores, and Genomic Relationship Matrix Restricted Maximum Likelihood (GREML) as implemented in GCTA and its extensions. All throughout this review we consider the different methods in terms of a basic regression model. This allows the reader to gain insight in the communalities and differences between the methods discussed here and used in behavior genetics from the 1970's up until this day.

Family and twin designs

Family and twin designs provide the means to infer, under explicit assumptions, the contributions of genetic and environmental influences to phenotypic individual differences from the phenotypic resemblance among family members. The statistical model employed in family and twin studies invariably involves the regression of the phenotype y on unobserved or latent genetic (G) and latent total environmental (T) variables:

$$y_{ij} = b_0 + g*G_{ij} + t*T_{ij} \quad (1),$$

Where *i* denotes family, *j* denotes family member, and b_0 is an intercept. Equation 1 represent the linear regression, but given a binary or discrete phenotype, y, generalized linear regression is used (notable probit regression see Falconer & MacKay[2] and Neal & Cardon[22]). Note that it is assumed that interaction (GxT) is absent; we return to this assumption below. The regression coefficient *g* quantifies the contribution of G to the phenotype y. Similarly, we assume T comprises effects of many (unknown) environmental factors, and the regression coefficient *t* quantifies their contribution. We assume that the genetic variable G includes the effect of a possibly large, but unknown, number of genes, each of small effect. Given that genetic variants occupy specific chromosomal locations, the term locus is also used in reference to a gene. To contribute to variance in y_{ij} , the genes that comprise G_{ij} must be polymorphic. Many genetic loci are monomorphic. Monomorphic genes may be functional, but they do not vary between individuals and therefore do not contribute to individual differences in the phenotype. Human autosomal chromosomes come in pairs (homologues), so that each locus on the 22 autosomal chromosomes is characterized by two alternate alleles, one on each member of a homologous pair. A well known example of a polymorphic gene is the ABO gene, a locus (on chromosome 9 at 9q34.1-q34.2) with alleles A, B, and O, which give rise to 6 distinct genotypes (AA, AB, BB, AO, BO and OO).

A diallellic genetic variant with alleles A and B gives rise to three genotypes: BB, AB, AA (assuming that BA and AB cannot be distinguished), which we may code 0, 1, 2 (reflecting the presence of the number of A alleles). The effect of alleles can be additive, which means that the effect of the genetic variant on the phenotype is purely additive (or in regression terms: purely linear). For instance, this is the case if the presence of one A allele increases the phenotype value by a given value *a* (i.e., genotypes coded 0, 1, and 2 correspond to effects 0, 1a, and 2a, respectively). Genetic non-additivity implies a deviation from linearity due to intra-locus allelic interaction (e.g., 0, 1, 2 correspond to 0, 0, 2a). Such intra-locus polygenic non-additivity is referred to as genetic dominance. Interaction between alleles at different loci is referred to as epistasis. To accommodate additive and non-additive genetic effect, the G in the regression model is replaced by A and D, representing additive effects and dominance effects. We discard epistatic effects as they are hard to distinguish statistically from dominance effects.[2]
Effects on the phenotype that are not attributable to genetic variation are referred to as environmental effects. Environmental effects are broad in nature as they may include prenatal exposures, experiences and exposures during childhood and throughout life. Relevant environmental variables are often unidentified and therefore not measured. Rather than estimating them directly, their effects are inferred from twin and family correlations. In the context of the twin model two classes of environmental influences are distinguished: common environmental influences that are shared among twin pairs and or family members, and unique environmental influences that are not shared among family members. When studies focus on older participants, whose phenotypes are assessed simultaneously with the phenotypes of cohabitants (spouses or other adults with whom they share a household), shared environment is sometimes referred to as 'household effects'. Both terms emphasize that resemblance among relatives, whether they are biological relatives or not, can arise from sharing an environment, in addition to sharing of genes.

Given the additive (A) and non-additive genetic affects (D) and, shared (C) and unshared (E) environmental effects we arrive at the following regression model:

$$y_{ij} = b_0 + a*A_{ij} + d*D_{ij} + c*C_{ij} + e*E_{ij} \quad (2)$$

Where a, d, c and e are regression parameters. By definition A and D are uncorrelated[2], and C and E are uncorrelated. Under assumptions that there are no interactions among the predictors, and A and/or D is uncorrelated with C and/or E, this regression model implies the following decomposition of phenotypic variance:

$$\sigma_y^2 = a^2\sigma_A^2 + d^2\sigma_D^2 + c^2\sigma_C^2 + e^2\sigma_E^2 \qquad (3)$$

The relative influence of genetic factors on phenotypic variation, common called the "heritability", is defined as the percentage of total phenotypic variance that can be attributed to genetic effects. "Broad-sense" heritability includes all sources of genetic variance (additive and non-additive; $h_b^2 = [a^2s_A^2 + d^2s_D^2]/ s_y^2$), "narrow-sense" heritability is limited to additive genetic variance ($h_b^2 = [a^2s_A^2]/ s_y^2$)). Note that large heritability (narrow or broad) implies that genetic differences contribute substantially to phenotypic variance, but provides no information concerning the number or location of the relevant genes. Measurement error often cannot be distinguished from unshared environmental effect, and we therefore assume that such error variance is included in $e^2s_E^2$. While the variance term $e^2s_E^2$ represent environmental effects (plus error), it may, in principle, include effects, which are not strictly environmental, such as the effects of private genetic mutations, and so-called "developmental noise".[23; 24]

Family and twin designs are used to arrive at estimates of the variance components in equation 3. In such designs, the independent variables (A, D, C, E) are not actually measured, rather their effects are inferred from the phenotypic correlation among individuals, who are in known genetic and environmental relationships.[25-27] As phenotypic resemblance is summarized in one or more phenotypic covariance matrices, the statistical analysis of such family data is essentially covariance structure modeling or structural equation modeling, in which the predictors A, D, C, and E are treated as latent variables.[6; 22] As the majority of behavior genetic studies of complex traits employed the classical twin design, we base our explanation of genetic covariance structure modeling on this design.

Let us reconsider equation 2, but now explicitly for monozygotic (MZ) and dizygotic (DZ) twin pairs (subscript j is now j=1 or j=2). Given that the predictors are latent, we have to impose some scale on them (this is standard in latent variable modeling[28]). We assume that the predictor variables are standardized (i.e., unit variance, zero mean). This implies that the intercept b_0 equals the mean of the phenotype.

$$y_{i1} - b_0 = a*A_{i1} + d*D_{i1} + c*C_{i1} + e*E_{i1} \quad (4a)$$

$$y_{i2} - b_0 = a*A_{i2} + d*D_{i2} + c*C_{i2} + e*E_{i2} \quad (4b)$$

These equations are associated with an expected 2x2 covariance Σ, with diagonal elements $\sigma_y^2 = a^2 + d^2 + c^2 + e^2$ (see equation 3, bearing in mind the variances are equal to 1). Assuming A, D, C, and E are uncorrelated and display no interaction, covariance between twin 1 (y_{i1}) and twin 2 (y_{i2}), σ_{y12}, equals $\sigma_{y12} = a^2*r(A_1,A_2) + d^2*r(D_1,D_2) + c^2*r(C_1C_2) + e^2*r(E_1E_2)$, where r() denotes correlation. The twin design is based on the fact that we know the values of these correlations given various assumptions. In the case of the environmental variables, the fact that C environmental effects are shared, and E effects are unshared implies $r(C_1C_2) = 1$ and $r(E_1E_2) = 0$. Here we assume that shared environmental influences in the MZ twins are the same as in the DZ twins. MZ twins are genetically identical, i.e., they share 100% of their genes (barring de novo mutations[29]). Genetic identicalness implies $r(A_1,A_2)=1$ and $r(D_1,D_2)=1$ in the MZ twins.[2] Assuming the genetic correlation among their parents is zero (random mating), dizygotic (DZ) twins on average 50% of their segregating genes[30; 31] ,and so $r(A_1,A_2) = .5$. The dominance correlation can be shown to equal .25 (see Mather & Jinks[31]) so that $r(D_1,D_2)=.25$ in DZ twins. We arrive at the following expected covariance matrices Σ in MZ and DZ twins (the ADCE model):

$$\Sigma_{MZ} = \begin{matrix} a^2 + d^2 + c^2 + e^2 & \\ a^2 + d^2 + c^2 & a^2 + d^2 + c^2 + e^2 \end{matrix} \quad (5a)$$

$$\Sigma_{DZ} = \begin{matrix} a^2 + d^2 + c^2 + e^2 & \\ \tfrac{1}{2}a^2 + \tfrac{1}{4}d^2 + c^2 & a^2 + d^2 + c^2 + e^2 \end{matrix} \quad (5b)$$

The mean vectors are $\mu_{MZ} = \mu_{DZ} = [b_0\ b_0]$. In this model, we assume no differences between e.g. first and second-born twins or between zygosity in means (all equal b_0) or in variance (σ_y^2). These assumptions are easy to test. This model, including the 4 parameters a, d, c, and e, is not identified. It is common practice to limit the models to effects of A, D, and E (an "ADE model") or A, C, and E (an "ACE model"), depending on the observed correlations (see Keller et al.[32] for a discussion of this issue). When D contributes significantly to a phenotype, we expect the correlation in MZ pairs to be larger than twice the correlation in DZ pairs. If C contributes significantly to a phenotype, we expect the correlation in MZ pairs to be less than twice as large as DZ correlations.

It is instructive to consider an ACE model briefly in terms of the correlations it might generate. Supposing the phenotypic variance equals 1 ($\sigma_y^2 = 1$), any genetic effects are bound to render the MZ phenotypic correlation (rMZ) greater that the DZ correlation (rDZ) while shared environmental effects will increase both MZ and DZ phenotypic correlations. Given $c^2>0$ (say, $c^2=.2$) and $a^2=0$, the correlations are expected to be equal (both .2). As a rule of thumb, if 2*rDZ>rMZ, this is taken to be indicative of an ACE model, and a quick estimate of a^2 is obtained as 2*(rMZ - rDZ). E.g., given rMZ=.6 and rDZ=.4, 2*(.6-.4) = .4, and c^2 is 2*rDZ-rMZ (e.g., 2*.4 - .6 = .2). The unshared environmental component, which includes measurement error, finally, is e^2=1-rMZ (1-.6=.4). If 2*rDZ<rMZ, this is indicative of the effects of non-additive genetic effects (dominance and/or epistasis). The fact that we are limited to ACE or ADE is a clear weakness of the twin design, which can be overcome by adding data from additional family members, such as half-siblings growing up in the same household, or parents of twins.

Genetic covariance structure modeling

Genetic covariance structure modeling (GCSM) is used to fit a given model (say, ACE model; i.e., assuming d=0) to the MZ and DZ twin data simultaneously to

obtain estimates of the parameters a, c, and e.[6] Typically this is done using programs for structural equation modeling with maximum likelihood (ML) estimation,[33] such as LISREL,[34; 35] Mplus,[36] Mx,[37] and OpenMx.[38] Mx and OpenMx were written specifically to facilitate twin and family modeling. ML estimation has the advantages of providing goodness of fit indices to evaluate overall model fit , standard error of parameter estimates, and nested model comparison using the likelihood ratio test.[28]

Covariance structure models, including twin and family models, can be represented graphically in a path diagram. An example of the path model that corresponds to the covariance model in equation 5 is shown in figure 1. The variables in squares are the observed phenotypes in the separate twins. The latent genetic and environmental variables are represented by circles. Their influence on the phenotype is given by path coefficients a, c, d and e.

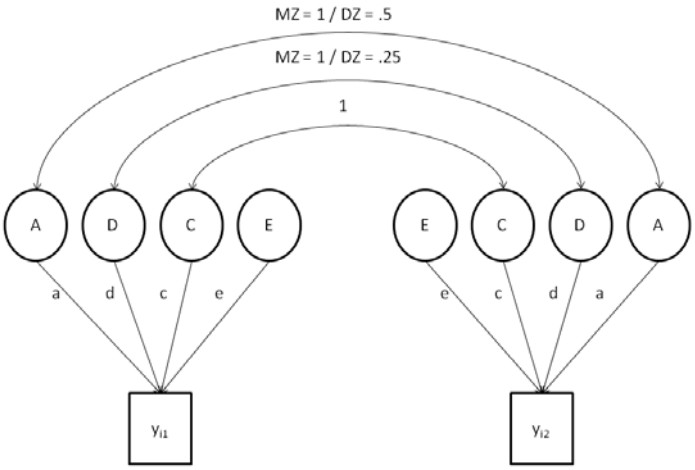


Figure 1: Path diagram corresponding to the ACDE covariance model described in equation 5. As the complete ACDE model is not identified either parameter d is set to 0 resulting in an ACE model or parameter c is set to 0 resulting in an ADE model.

Path diagrams provide highly insightful and intuitive representations, which are mathematically complete.[39; 40] This means that the covariance structure implied from the diagram can be inferred using path tracing rules,[22] Keller et al.[41] discussed the application of the tracing rules in some detail.

So far, we have considered the linear regression model, in which the phenotype is continuously and (conditional on fixed covariates such as sex and age) approximately normally distributed. Non-normality can be handled by transforming the data to approximate normality, or by adopting a robust estimator. Often traits of phenotypes of interest, such as for example the presence and or absence of disease, are measured on a discrete scale. Other phenotypes of interest are measured on an ordinal discrete scale. Fitting models to discrete data can be done by applying a suitable least squares estimator,[28; 42] or by full information maximum likelihood estimations, which is the main method in OpenMx. Regardless of the estimator, the approach is closely related to probit regression analysis[43; 44] where a binary phenotype is regressed on predictors. Models for discrete phenotypes assume a latent normally distributed variable, usually called the liability. The frequencies (or prevalences) of the observed discrete values (say 0,1) determine the position of thresholds (specific points) on the liability. The frequency of a given response (say the frequency of the response 0 is .5) can be derived from the liability by integration of the liability distribution of (say) minus infinity to the threshold value. This approach can be generalized to discrete phenotypes comprising more than two values. In this model (the "liability-threshold model"), the liability is the variable that is subject to genetic covariance structure modeling using twin data, where it is assumed that the a bivariate normal distribution underlies the bivariate discrete twin data.

Multivariate and Longitudinal Genetic Analyses

Above, we considered a single phenotype, measured in twins. As twin pair (rather than individual twin) is the sampling unit, the data are by definition bivariate, as is the covariance structure model. Genetic covariance structure modeling can also be applied to multivariate data. Multivariate data arise when two or more phenotypes are analyzed simultaneously.[6; 26; 34; 45] Here we consider the multivariate genetic model. We again start with the multivariate regression:

$$y_{ij} - b_0 = a*A_{ij} + d*D_{ij} + c*C_{ij} + e*E_{ij} \quad (6)$$

Letting p denote the number of phenotypes, y_{ij} is the p dimensional phenotypic vector in twin *j* of family *i*, with p dimensional mean vector b_0. The p dimensional vector A represents the p additive genetic variables (zero-mean, unit variance), associated with the p phenotypes (D, C, and E are defined analogously). The pxp matrices a, d, c, and e are diagonal matrices containing the regression coefficients (or path coefficients) as they express the change in the phenotype as a function of a change in the predictors (A, D, C and E). The within twin expected covariance structure is:

$$\Sigma_y = aR_Aa^t + dR_Dd^t + cR_Cc^t + eR_Ee^t = \Sigma_A + \Sigma_D + \Sigma_C + \Sigma_E \quad (7)$$

Where R_A is the pxp correlation matrix of the additive genetic effects A, and aR_Aa^t is the pxp genetic covariance matrix, i.e., $\Sigma_A = aR_Aa^t$ ($\Sigma_D = dR_Dd^t$, $\Sigma_C = cR_Cc^t$, and $\Sigma_E = eR_Ee^t$ are defined analogously). Note that the covariance matrices Σ_A , Σ_D, Σ_C and Σ_E are not modeled or constrained in anyway. We assume only that they are positive (semi-) definite covariance matrices, and that Σ_y is positive definite[a]. The covariance matrix equals the 2p x 2p covariance matrices:

$$\Sigma_{MZy} = \begin{matrix} \Sigma_A + \Sigma_D + \Sigma_C + \Sigma_E & \Sigma_A + \Sigma_D + \Sigma_C \\ \Sigma_A + \Sigma_D + \Sigma_C & \Sigma_A + \Sigma_D + \Sigma_C + \Sigma_E \end{matrix} \quad (8a)$$

[a] Positive definite implies that the eigenvalues are positive. This is a characteristic of a valid covariance matrix. Positive (semi-) definite implies that the eigenvalues are zero or greater than zero. Note that the phenotypic covariance matrix is required to be positive definite, but the underlying covariance matrices may be positive semi definite.

$$\Sigma_{DZy} = \begin{matrix} \Sigma_A + \Sigma_D + \Sigma_C + \Sigma_E & \tfrac{1}{2}\Sigma_A + \tfrac{1}{4}\Sigma_D + \Sigma_C \\ \tfrac{1}{2}\Sigma_A + \tfrac{1}{4}\Sigma_D + \Sigma_C & \Sigma_A + \Sigma_D + \Sigma_C + \Sigma_E \end{matrix} \quad (8b)$$

Often this model is estimated by subjecting the covariance matrices to a triangular or Cholesky decomposition (e.g., $\Sigma_A = K_A K_A^t$, where K_A is a lower triangular matrix[22; 46]). This particular parameterization of the covariance matrices is convenient as it guarantees that these matrices are positive (semi-) definite. As in the case of a univariate model, the full model, as presented in Eq 8a-8b, is not identified. Again it is standard practice to consider models including either ACE or ADE (or submodels thereof). Figure 2 shows the path diagram that corresponds to the Choleski decomposition of the covariance between 2 traits.

The multivariate twin model can be used to determine the contributions of genetic and environmental effects to the phenotypic variances and covariances. In addition, each covariance matrix (in the ACE model: Σ_A, Σ_C, and Σ_E) can be subjected to its own covariance structure model, allowing one to determine the covariance structure of the genetic and environmental effects. In this connection, it is interesting to note that any phenotypic covariance structure is the sum of the genetic and environmental structures (eq. 7). One may ask what the relationship is between the phenotypic covariance structure and the latent genetic and environmental structure. For instance, is the phenotypic 5 factor structure of personality[47; 48] a reflection of, say, 5 factor structures of the underlying Σ_A and Σ_E covariance matrices.[49; 50]

Multivariate data arise naturally in longitudinal studies, where the same phenotype(s) is (are) measured repeatedly. The longitudinal twin model again can reveal contributions of genetic and environmental effects to the stability over time. The genetic and environmental covariance structures can be modeled using well established models for repeated measures, such as the simplex model ,[45; 51; 52] which emphasizes genetic and environmental contributions to stability and change, and growth curve models, which emphasize the contributions to growth.[53] Chapter 2 of this dissertation includes an application of the simplex model to a cohort

sequential classical twin design to study the stability of anxious depression from 3 to 63 years.

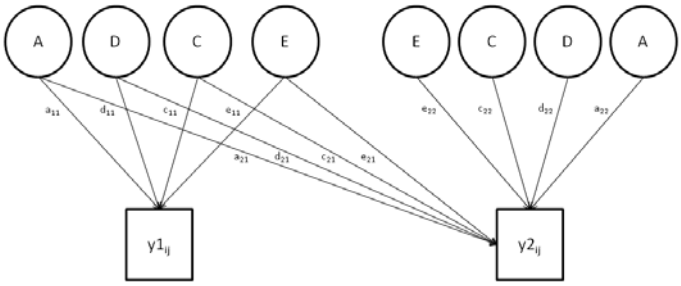


Figure 2: The path model corresponding to a Choleski decomposition of A, C D and E for two phenotypes measured in a single individual. Parameters a_{11}, a_{12} and a_{22} correspond to the values in the lower triangular matrix Ka the genetic covariance matrix is the product of the matrix K_a and its transpose. (i.e. $\Sigma_A = K_A K_A^t$). The 4 covariance matrices in the full ACDE model (K_a, K_d, K_c and K_e) cannot be estimated simultaneous. It is therefore customary to restrict d_{11}, d_{12} and d_{22} to 0 resulting in an ACE model or to restrict c_{11}, c_{12} and c_{22} to 0 resulting in an ADE model.

The classical twin model, as presented, is based on various explicit assumptions.[2] These include random mating, absence of any interaction among the latent variables, the latent variables A and D being uncorrelated with C and E, and the assumption of equal environmental (including prenatal intra-uterine environment) effects in MZ and DZ twins.[2] A lot of work has been devoted to extending the classical twin model to arrive at a design, which is less dependent on these assumptions. For instance, assortative mating (a positive phenotypic correlation among spouses) can result in a spurious c^2 component in the classical twin design. By extending the twin design to include the parents, assortative mating can be included in the model, and its effects accounted for.[41; 54] Below, we discuss in some detail the assumption of additivity of effects, i.e., the absence of any form of gene-by-environment (GxE) interaction, and the assumption that A and D are uncorrelated with C and E.

Interaction: Moderation of genetic and environmental effects

GxE interaction (or in notation consistent with equation 1 GxT) is classically conceptualized in terms of genetic control of sensitivity to the environment, i.e., the effect of an environmental risk factor on an individual depends on the genetic make-up of the individual.[25; 55] We cast this in terms of moderation, i.e., the moderation of any effect estimated in the twin of family design (environmental or genetic) by any measure of latent variable (genetic or environmental). As such, we can consider AxE interaction in the twin model, where the moderation of environmental effects (E) by additive genetic effects (A) is detectable as heteroskedasticity: the variance of E, summarizing the effects of E, varies as a function of the level of A. This is complicated by the fact that A is a latent variable, but various approaches have been developed that can detect such heteroskedasticity.[25; 56-59] Alternatively we can consider interaction in terms of the effects of a measured moderator on both genetic and environmental effects.[60; 61] In this case the moderator can be any variable, and may itself be subject to genetic and environmental effects.

The exact model used to test moderation depends on the nature of the moderator. The simplest case is a moderator that has the same value in both twins, such as age ,[52; 62] religious upbringing,[63] or social economic status. In this case, one can estimate the effects of genotype and environment on the phenotype of interest

conditional on the moderator.[25; 55; 64-66] In the absence of any moderation, the genetic and environmental effects, as expressed in terms of variance components (a^2, c^2 or d^2, and e^2), are the same regardless of the value of the moderator. Note that the absence of moderation does not imply the absence of a main effect. For instance, if the genetic and environmental effects on height are the same in 12 and 18 year olds, this means that the size of the variance components are identical, but it does not rule out a main effect of age on height. Clearly, on average, 12y olds and 18y olds are likely to differ in height.

Sex is another example of a relatively simple moderator, as conditioning on sex and testing whether the variance components are equal is straightforward. Introducing a subscript for sex (f for female; m for male), in the univariate model we test $a_f^2 = a_m^2$, etc. However, the presence of DZ opposite-sex twins (DZOS) offers the unique possibility of further investigating the nature of sex differences in a^2 ($a_f^2 \neq a_m^2$) or c^2 ($c_f^2 \neq c_m^2$). Given $a_f^2 \neq a_m^2$ there are two possibilities concerning the nature of the difference in variance, which can be distinguished by testing the DZOS genetic correlation. Either the same genes are active in males and females, but the effect of the genes is moderated by sex (a quantitative sex difference), or different genes are active in males and females (a qualitative sex difference). The DZOS additive genetic correlation is expected to be .5 in the former case, but less than .5 in the latter.[67] The presence of DZOS twins, in combination with male and female MZ and DZ twins allows one to test this. The qualitative sex differences model can also be applied in the context of an environmental hypothesis: instead of fixing the correlation between C factors at 1 in DZOS twins, it can be estimated as a free parameter. If it is judged to be significantly lower than 1.0, this indicates that the influence of the shared environment differs in the two sexes. However, note that, as there is only one group of opposite-sex twins (there are no MZ twins of opposite sex) this analysis is limited to either the genetic or the common environment correlation, as they cannot be estimated simultaneously.

Purcell [60; 61] developed a general model to investigate moderation by a continuously distributed moderator, which is itself possibly subject to genetic and environmental effects. This model is based on a bivariate model, which includes the simultaneous ACE (ADE) decomposition of the moderator and the phenotype using a Cholesky decomposition. This model accommodates any phenotype covariance between these variables stemming from shared genetic or environmental influences. In addition, the paths specific to the phenotype and the paths accounting for the phenotype-moderators covariance are subject to

moderation. Various simplifications of the Purcell model are possible, but if the full model holds (i.e., full moderation of ACE effect on the phenotype and of the ACE effect shared by the moderator and the phenotype), more simplified models may produce spurious moderation results.[68]

The Purcell model is often presented as a model for GxE interaction in the presence of gene-environment correlation. This interpretation is based on the conceptualization of the moderator as an "environmental" variable, which may be subject to genetic influence. For instance, general parental support (encouragement, taking an active interest, helping with homework, etc.) may moderate genetic and environmental influences on twins' intelligence. But parental support, while contributing to the environment of the twins, is itself likely to be subject to genetic influences. Gene-environmental correlation could arise if the amount of support was a function of parental intelligence. We return to this subject below.

Gene-environment correlation

The possibility that the genetic latent variables (A and D) are correlated with the environment variables (C and E) has been discussed extensively.[25; 55; 69] Various plausible processes are expected to give rise to gene-environment correlation. For instance, the contributions of a parent to the home environment, as experienced by the offspring, may depend on the parent's genotype.[70] This process can be investigated in models including parents and twins, by including the regression of the twin environment on the parental.[41; 70] Similarly, siblings (including twins) that grow up together may contribute (negative or positively) to each other's environment.[71] If the phenotypes involved in such contribution are subject to genetic effects (e.g., rowdiness, aggression), this will give rise to gene-environment correlation. This process can be studied by including the regression of the twins phenotypes on each other in a cross-sectional [71] or longitudinal twin study.[72; 73; 73]

The classical twin design can be used to estimate the genetic and environmental influences to individual differences in a given "environmental" variable, such as marital status. For example, the individual differences in the tendency to marry is subject to genetic influences , [74; 75] as is the tendency to divorce.[76] To determine whether the association between a specific environment and a trait is due to gene-

environment correlation, the bivariate twin design (Figure 2) can be applied, as mentioned above.[60]

This issue can also be addressed using the co-twin control design.[77; 78] In this design MZ and DZ twins, who are discordant for a given condition, are studied along with unrelated individuals. In addition to the condition (e.g., disease status; say long cancer), a risk factor is measured (say, smoking), which is related to the condition. Assuming genetic and environmental influences are uncorrelated, and assuming a direct causal effect of the risk factor on the condition, the strength of the association will be the same in the MZs, DZs, and unrelateds. If the phenotypic association between the risk factor and the condition is due to pleiotropic genetic effects (genes affect both phenotypes, in the absence of any direct relationship), the strength of the association will be greatest in the unrelateds, but smaller in the DZs, and zero in the MZs (as the MZs are completely matched for genetic influences).

An example of this is the study by Kendler et al.[78] of the relationship between smoking and depression in women. They concluded that the relationship between depression and smoking in women is due to familial (probably genetic) factors that predispose to smoking and depression, and not a direct effect of depression on smoking. Groen- Blokhuis et al.[79] investigated the association between low birth weight and attention problems. As in MZ pairs, DZ pairs and unrelated pairs of children, the child with the lowest birth weight scored higher on attention problems at age 3, 7, 10 and 12, the association is not a function of environmental or genetic correlation, but causal: that is, a lower birth weight directly causes increased attention problems.

Linkage analysis

Since the late 1990's, the advent of high throughput genotyping technologies enabled researchers to measure actual genetic variants in unprecedented volumes. These data can be exploited in the hunt for the causal genetic variants contributing to the variance of complex phenotypes. Initially, the focus was on

highly polymorphic genetic markers (i.e., microsatellites, comprising many alleles), which were exploited in linkage analysis to locate chromosomal regions associated with a phenotype. Such regions included multiple genes, and were likely to harbor the gene that represented the variants causally linked to the phenotype. Linkage analysis relies on the fact that genes in close chromosomal proximity are transmitted together (linked). Such linkage is disrupted by recombination. That is, when gametes (sperm and egg cells) are produced during meiosis, the paired homologous chromosomes separate so that each gamete contains only one of the pair of alleles for each trait. During the first division of meiosis, sections near the ends of chromosomes commonly exchange parts of their chromatids with the other chromosome of their homologous pair. The probability of the linkage between two loci being disrupted, i.e., recombination occurring, depends on the distance between the loci. Highly polymorphic markers have been used in parametric and non-parametric linkage analysis. Here we only consider non-parametric linkage.

In non-parametric linkage analysis, variation in the proportion of alleles that family members share identically by descent (IBD, i.e. from the same ancestor) at a given marker locus (see below) is exploited to identify the contribution of the marker to the phenotypic differences. A wide range of software packages is available to perform linkage analysis.[80; 81] We introduce this type of analyses within the context of genetic structural equation modeling (GCSM), as used in the classical twin model. The model employed in linkage can be written in terms of regression of the phenotype on the (latent) genetic effect of a given quantitative trait locus (QTL) (Q) the latent genetic variance (G) ant the total environmental variance (T).

$$y_{ij} = b_0 + q * Q_{ij} + g * G_{ij} + t * T_{ij} \quad ` \quad (9)$$

As applied to full sibs, the model is often limited to the QTL, and additive genetic variable (A) and unshared environmental effects (E), i.e., $y_{ij} = b_0 + q * Q_{ij} + a * A_{ij} + e * E_{ij}$. This regression model implies the following decomposition of phenotypic variance:

$$\sigma_y^2 = q^2\sigma_Q^2 + a^2\sigma_A^2 + e^2\sigma_E^2 = q^2 + a^2 + e^2, \qquad (10)$$

as scaling of the latent variables (Q, A, and E) implies that their variances equal one. We can derive standardized variance components in the model, i.e., the total heritability equals $(q^2 + a^2)/ (q^2 + a^2 + e^2)$, and the variance explained by the QTL is $q^2/ (q^2 + a^2 + e^2)$.

Linkage analysis can be performed in complex pedigrees, but we consider linkage performed in DZ twins or full sib pairs, where again we consider the implied phenotypic covariance matrix. This implied variances are given in eq 10. The covariance between sibs is $\sigma_{y12} = q^2*r(Q_1,Q_2) + a^2*r(A_1A_2) + e^2*r(E_1E_2)$, where, as mentioned above, $r(A_1,A_2)=.5$ and $r(E_1E_2)=0$. The correlation between QTL factors of DZ twins or siblings $r(Q_1,Q_2)$, which is often denoted pi-hat ($\hat{\pi}$), is obtained from measured genotypic (marker) data. IBD status for the marker data determines this correlation. IBD status at a given locus equals 0, 1, or 2 in sibs, depending on the exact configuration of parental alleles that the sibs have inherited. To illustrate this in the simplest case, suppose that the parental genotypes at the locus of interest are A1A2 and A3A4, in mother and father, respectively. If the sibs are have identical genotypes (e.g., both A1A3), they share two alleles IBD (namely A1 and A3). If the sibs have no alleles in common (e.g., A1A3 and A2A4), they share zero alleles IBD. Finally if they share one and the same allele from a given parent (e.g., A1A3 and A1A4), they share one allele IBD (namely A1). Often IBD status cannot be established with certainty (suppose the parental genotypes are A1A1 and A1A2, sibs with genotypes A1A2 and A1A2 may be IBD2 or IBD1 depending on whether they inherited the same A1 allele from the mother). Ii is however always possible to assign IBD probabilities. The value of pi-hat is then obtained by calculating the mean of the sib pair specific IBD distribution, where the IBD is expressed in terms of the proportion of alleles shared IBD (i.e., the proportions 0, .5, and 1 correspond to 0, 1 and 2 alleles IBD). For instance, in the unambiguous case that prob(0)=0, prob(.5)=0 and prob(IBD=1)=1, pi-hat equals 0*0+0*.5+1*1 = 1. If the parents are identical homozygotes (both A1A1), we have prob(0)=.25, prob(.5)=.5 and prob(1)=.25, and pi-hat = 0*.25 + .5*.5 + 1*.25 = .5.[82; 83] So, we can specify the phenotypic covariance between siblings as:

$$\sigma_{y12i} = q^2 * \pi_i + a^2 * .5, \quad (11)$$

where the covariance bears a sib-pair subscript, as the covariance depends on the value of pi-hat.

Note that pi-hat is indicative of genetic resemblance in the region of the marker. Suppose that the marker happens to be the QTL (i.e., causal variant). In that (exceptional) case, the QTL will contribute to the phenotypic resemblance of the sibs as a direct function of pi-hat. E.g., pi-hat equals 1 (or prob(IBD=2) = 1) means that the sibs are genetically identical at the QTL and so the QTL contributes fully to their phenotypic resemblance ($\sigma_{y12i} = q^2 + .5*a^2$). If pi-hat is zero, the sibs are essentially genetically unrelated at the QTL, and so the QTL contributes nothing to their phenotypic resemblance ($\sigma_{y12i} = .5*a^2$). Of course, the marker is unlikely to be the actual QTL, and the further away the marker is from the QTL, the more the IBD relationship will be diluted by recombination, i.e., the less indicative the individual pi-hat value is of genetic resemblance at the QTL.

This specification of the within sib pair covariance allows us to construct two models for the observed covariance for sib ships and DZ twin pairs. Figure 3 shows a path model for DZ twins or siblings that incorporates the effect of a QTL on a measured phenotype.

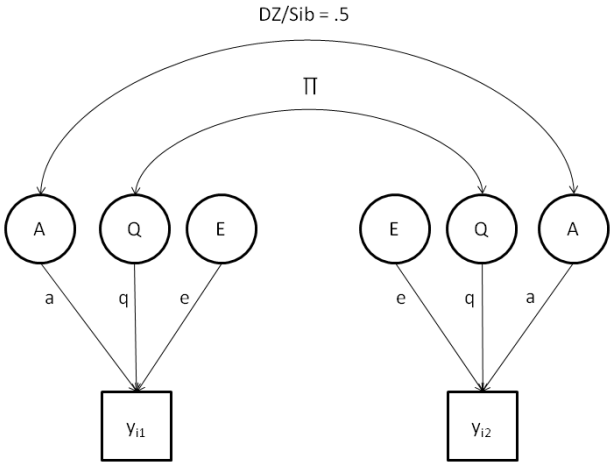


Figure 3: path representation of the linkage covariance model between two siblings or dizogotic twins as described in equation 12a.

The effects of Q, A, and E are not measured directly, they are inferred from family members for which the genetic relation is known (here sibs and DZ twin pairs) and the genetic correlation (π) at locus Q is estimated.

$$\Sigma_{y\text{-}qtl} = \begin{matrix} q^2 + a^2 + e^2 & \\ \pi_i * q^2 + .5 * a^2 & q^2 + a^2 + e^2 \end{matrix} \qquad (12a)$$

$$\Sigma_{y\text{-}null} = \begin{matrix} a^2 + e^2 & \\ .5 * a^2 & a^2 + e^2 \end{matrix} \qquad (12b)$$

The models in equation 12a and 12b are fitted using maximum likelihood and compared based on a likelihood-ratio test which is distributed as χ^2. In a linkage analysis, results are summarized in the form of a LOD score.[84] There is an easy conversion between χ^2 and LOD scores: LOD = χ^2 / 2ln10.[85] Evidence for linkage is present when the maximum LOD-score exceeds a pre-defined threshold, This threshold is generally based on the size of the genome and the number of measured markers.[83] A commonly used threshold is a LOD score of 3, this value corresponds to the data being 1000 times more likely given the model including the QTL than given the null model. A LOD score of -2 is generally taken to indicate strong evidence against linkage, a value of -2 corresponds to the observed data being 100 times more likely given the null model than given the model including the QTL. While the usage of linkage has largely been abandoned in favor of association analysis, recent papers based on linkage techniques have been used to determine the heritability of complex phenotypes without relying on common assumptions associated with the twin model.[30] Linkage can further be combined with association to fine map a specific region of the genome, for application see van Dongen et al.[86].

Association analysis

Candidate gene studies

In contrast to twin and linkage studies, association studies focus on the direct effect of genetic variants on the trait. If the phenotype is continuous and conditionally (on the predictor) normally distributed, association can be cast in terms of a simple linear regression analysis, which can be conducted in a sample of N unrelated individuals:

$$y_j = b_0 + b_s * SNP_j + e_j \qquad (13)$$

where y_j is the phenotype, SNP_j is a measured genotypic marker (usually coded 0,1, 2 with the code conferring the number of reference allele's present in individual *j*), e_j is the residual, and *j* denotes the individual (j=1,..,N). Note that in this model the residual includes all effects not attributable to the SNP. These include environmental effects, genetic effects (attributable to all other genetic variants) and measurement error. Association can tested statistically by testing the null-hypothesis $b_s=0$ vs. $b_s\neq0$. If the phenotype is binary (i.e., disease status in a case-control design), generalized linear regression is used, but the test is the same. Association analysis is statistically more powerful than (sib pair) linkage analysis, because the test $b_s=0$ concerns first-order statistics (conditional mean or prevalence), whereas the linkage test $q^2=0$, see eq 12a-b concerns second-order statistics (covariances). Association analysis was initially targeted at biologically plausible candidate loci. Several strategies can be used in selecting candidate genes. Genes that are part of physiological systems known to influence the trait may be suitable candidates. Genes (or chromosomal regions) that are known to influence the trait in animal models can inform the selection of candidate genes (or regions) in humans (see Chapter 2). Or, genes can be selected in regions of the genome, which were identified as being of interest in linkage analysis.

In association studies the observed association between a given trait and a given genetic variant may be due to population stratification. If a population comprises several subpopulations that differ in allele frequencies and in phenotypic values,

then the observed association may express mainly these differences, not any true relationship between a given genetic variant and the phenotype. Hamer and Sirota[87] illustrated this by means of a (fictional) study in which genetic variants were related to the use of chopsticks in an sample of American students. The students includes both Asian Americans and European americans, who differ in the frequency of using chopsticks, and in allele frequencies at various loci. The association between any of these loci and chopstick use is spurious, as it is driven by subpopulation differences, not by the effect of the loci on chopstick use. If one corrects for this type of population stratification, the association disappears. E.g., one might consider testing the association separately in the Asian American and European American students. To correct for stratification in candidate gene studies one may adopt a within family association design. These are not subject to the effect of stratification, because they base the test of association on the association as observed within families, where the phenotypic differences are not subject to stratification, and the family members are match with respect to many variables.[88] For instance, full sibling pairs do not differ ethnically and racially. The association between a genetic variant and a given phenotype therefore is very unlikely to be spurious. There are several family based association tests, including the Transmission Disequilibrium association Test (TDT) and the Haplotype Relative Risk Test (HRR). An approach based on covariance structure modeling involves decomposing the effects of a given genetic variant into between-family and within-family components. The true assocation is reflected in the within family component, while the between family component may include the effects of stratification.[89; 90]

Genome wide association studies (GWAS)

Candidate gene association studies are hypothesis driven: i.e., they focus on a given gene (possibly encompassing many genetic variants), which is judged to be of special interest on prior ground. GWAS, in contrast, are exploratory in that they test the association between a given phenotype and many hundreds of thousands of genetic variants, usually single nucleotide polymorphisms (SNPs). SNPs, which are diallelic loci, and so give rise to three genotypes, are chosen to broadly cover the whole genome. In this exploratory approach, a significant association may concern a causal genetic variants (with a biological interpretation), but is more likely to be a genetic variant that is in linkage disequilibrium with this causal

variant. Linkage disequilibrium (LD) refers to the fact that genetic variants that are located so closely on a given chromosome, that recombination between them is a rare event. As a consequence, the loci co-segregate in the population over many generations.

Before performing a GWAS, quality control (QC) and data cleaning need to be applied to the genetic data.[91; 92] SNPs that do not meet quality control criteria should be excluded from further analysis. GWAS are identical to candidate association studies with respect to the statistical test (see equation 13). Both GWAS and candidate locus association studies require a correction of the alpha (significance threshold) for multiple testing. In a GWAS, comprising a large number of test (>500K to >7000K), the corrected alpha level is usually set at $5*10^{-8}$; which corresponds to a testing burden of one million independent tests .[93-95]

Both GWAS and candidate gene studies have to control for population stratification. The presence of many measured genetic variants in a GWAS, allows one to control for population stratification by means of a principal component (PC) analysis of the genetic data.[96] Let the (KxN) matrix X contain the standardized K SNPs in N individual. Standardization implies that the genotypic values of each SNP, as observed in the N individuals, are centered by subtracting the SNP mean (μ_k; k=1...K), and scaled by dividing by the SNP standard deviation (σ_k; k=1...K). As the SNPs are diallelic, the mean and standard deviation can be expressed as a function of the minor allele frequency of SNP k, p_k, i.e., $\mu_k = 2*p_k$ and $\sigma_k = \sqrt{p_k * (1 - p_k)}$. Let $\mathbf{\Psi}$ denote the NxN covariance matrix of the individual, i.e., $\mathbf{\Psi} = \mathbf{X}^t\mathbf{X}/K$. Stratification can be controlled using the eigenvectors or principal components (PCs) of this matrix. The first PCs of the matrix $\mathbf{\Psi}$ reflect the possible ancestry on a global scale[96] but also on a local scale.[29] The global scale concerns people of different ethnic background (e.g., Asian vs Europeans). These are identified and usually removed, after which the $\mathbf{\Psi}$ and the PCs are recalculated. The local scale concerns individuals of different ancestry within a continent or nation (e.g., west Europeans). A number of PCs which are judged to reflect local differences in ancestry, are computed and included in the regression model for association:

$$y_j = b_0 + b_1 * PC_{1j} + b_2 * PC_{2j} + \ldots + b_s * SNP_j + e_j \qquad (14)$$

Software is available for the calculation of these principle components.[96]

Note that in equation 13 and 14, we included no family index, i.e., we assumed that the sample consists of unrelated individuals. The standard test of the SNP effect (i.e., the test of the null-hypothesis $b_s = 0$) in (logistic) linear regression assumes that, conditional on the predictors, the residuals (e_j) are independent. However, the sample may include individuals clustered in families (e.g., MZ and DZ twins, their parents, their sibs, etc.). Given a sample including family members, we have the following regression model:

$$y_{ij} = b_0 + b_1 * PC_{1ij} + b_2 * PC_{2ij} + \ldots + b_s * SNP_{ij} + e_{ij} \qquad (15)$$

where i denotes family and j denotes individual. In this model, the residual terms e_{ij} are expected to be correlated between family members, due to all the influences they share, in addition to the predictors in the model. That is, in de case of an ACE model (see above), the residuals would be subject to $e_{ij} = a * A_{ij} + c * C_{ij} + e * E_{ij}$. To arrive at a correct statistical test of $b_s=0$, one can model the background covariance structure (i.e., the covariance structure of e_{ij}) correctly (equation 16).

$$y_{ij} = b_0 + b_1 * PC_{1ij} + b_2 * PC_{2ij} + \ldots + b_s * SNP_{ij} + a * A_{ij} + c * C_{ij} + e * E_{ij} \qquad (16)$$

If the families consist of MZ and DZ twins, this would involve estimating the fixed regression coefficients, and fitting a covariance structure model to the residuals (to estimate the parameter a^2, c^2, and e^2; see eq 5a, 5b). However, this can be computationally too slow, given the large number of SNPs in a GWAS. One solution is to estimate the background covariance matrix once (assuming $b_s=0$) and retain this matrix in the test of all SNPs (avoiding repeated estimation).[97] Alternatively, one can forgo the modeling of the background covariance, treat e_{ij} as independent, and correct the standard error of the estimate of b_s for this misspecification by means of a sandwich correction.[98; 99] The latter is convenient,

as it is fast and has been shown to produce accurate type I error rates. However, the degree of misspecification has a bearing on the power to detect the SNP effect (given $b_s \neq 0$). Minicâ, et al.[100] demonstrated that the robust test based on an independence model (as implemented in Plink[80]), is less powerful that the robust test, based on the exchangeable model. In the exchangeable model (in which the covariances between the family members are constrained to be equal), the power of the latter was almost the same as the power of the test given full correct modeling of the background covariance matrix, while being computationally faster.

Recent methodological advances in association studies have seen the development of a more integrated approach based on the linear mixed model to testing association given population stratification, the presence of known closely related individuals, and cryptic relatedness, i.e. relatedness that is unknown to the researcher.[101; 102] For this model we switch from scalar notation to matrix notation. Let **y** be the Nx1 random vector containing the phenotypic values in the sample of N individuals, who may be related. The phenotype is modeled as follows*:*

$$\mathbf{y} = b_0 * J + b_s * SNP + \ldots. + \mathbf{X}^t\mathbf{u} + \mathbf{e} \qquad (17a)$$

$$\Psi = \mathbf{X}^t\mathbf{X}/K \qquad (17b)$$

$$\Sigma(y) = \Psi \otimes \sigma^2_{\upsilon} + \mathbf{I} \otimes \sigma^2_{e} \qquad (17c)$$

In equation 17a, the phenotype **y** (Nx1) is regressed on the intercept (multiplied by a unit vector *J* of dimensions (Nx1), as y is a multivariate vector) and on the vector (*SNP*) (Nx1) containing the SNP of interest (coded 0,1,2) for each individual (other covariates, such as age and sex may be included). The KxN matrix **X** is the matrix of the standardized genotypic values, as defined above. The Kx1 vector **u** contains the random regression coefficients associated with the K genotypes in **X**. The parameters b_s are estimated and subject to testing the K parameters in u are treated as random values, realizations of the normal distribution with mean 0 and variance σ^2_u, $u_k \sim N(0, \sigma^2_u)$. Equation 17c shows the expected covariance matrix

conditional on the fixed regressors, where $\boldsymbol{\Psi}$ is defined in eq. 17b, and σ^2_u and σ^2_e are the variances of the random regression coefficients **u** and the residual **e**, respectively. Note that in this model any genetic relatedness conditional on the fixed regressors is accommodated in the matrix $\boldsymbol{\Psi}$, which contains the average over the K loci of the allelic correlation estimates for any two individuals. If the individuals are closely related their average correlation tends to the expected additive genetic correlations (e.g., the $r(A_1A_2)$, as discussed above in the context of MZ and DZ twins). It is possible to reintroduce the variance attributable to shared environment, discussed above in the context of twin and family studies, into the model. More distantly related individual will have expected values of the average correlation, which are consistent with their degree of distant genetic relatedness. In essence, all individuals in the sample are treated as related, bearing in mind that the degree of relatedness, as expressed in the off-diagonals of $\boldsymbol{\Psi}$ may vary from ~1 (MZ twins) to ~.5 (1st degree relatives), to values approaching zero (very distant relatedness). Note that the SNP tested with fixed effect b_s, should not be included in the matrix **X**, as this will reduce the statistical power to detect $b_s \neq 0$.[102] One way to prevent this loss of power, is to calculate the matrix $\boldsymbol{\Psi}$ 22 times, where in the calculation of $\boldsymbol{\Psi}$ one chromosome is excluded. So in regressing the phenotype on any SNP on chromosome *c*, the matrix $\boldsymbol{\Psi}$ is based on the SNPs on all chromosomes barring chromosome c. We denote this matrix $\boldsymbol{\Psi}_{\neq c}$, which equals $\mathbf{X}_{\neq c}{}^{t}\mathbf{X}_{\neq c}/K_{\neq c}$. The standardized genotype matrix now has $K_{\neq c}$ rows, and N column, so that $\boldsymbol{\Psi}_{\neq c}$, remains NxN. This model is conveyed in equation 18:

$$\mathbf{y} = b_0 * J + b_1 * SNP + \ldots. + \mathbf{X}_{\neq c}{}^{t}\mathbf{u}_{\neq c} + \mathbf{e} \qquad (18a)$$

$$\boldsymbol{\Psi}_{\neq c} = \mathbf{X}_{\neq c}{}^{t}\mathbf{X}_{\neq c}/K_{\neq c} \qquad (18b)$$

$$\boldsymbol{\Sigma}(\psi) = \boldsymbol{\Psi}_{\neq c} \otimes \boldsymbol{\sigma}^2_{\upsilon \neq c} + \mathbf{I} \otimes \boldsymbol{\sigma}^2_{\varepsilon} \qquad (18c)$$

As discussed by Yang et al.,[102] this approach is becoming the method of choice in conducting a GWAS, as it handles distant (including stratification) and closed relatedness in a single model, and produces the best estimates of b_s in terms of precision (i.e., standard error). The precision of the estimate of b_s is important in

its own right as a test of the SNP effect, but also important in the calculation of genetic risk scores, as explained in more detail below.

Meta and mega analysis of genome wide association studies

Given the potential confounders discussed above, and the large number of tests, GWAS require replication in independent samples. The results of multiple studies can be used in a meta-analysis to arrive at a single test of association based on all available results. However, as different studies often use different genotype arrays, and the arrays measure different sets of SNPs, it is possible that only a small number of SNPs are genotyped in all replication samples. However, SNPs in close proximity on the genome are generally in strong LD (i.e. strongly correlated). This information can be leveraged to impute SNPs to arrive a set of SNPs, common to all studies. In specific reference samples, all SNPs are measured using whole genome sequencing.[103; 104] Given that all SNPs in the reference population are characterized, and the LD between these SNPs is known, one can impute the SNPs in the set that are not directly measured on a given genotyping platform.[105-107] After imputing all SNPs in the reference set for all samples, the association test is performed for all SNPs that are imputed with acceptable accuracy. This yields effect size estimates for a homogeneous set of SNPs across all cohorts. These results can then subsequently be meta-analyzed. Software is available to perform such genome wide meta-analysis (for example: METAL[108]).

The need to harmonize phenotypes and the need to impute SNPs in individual cohorts to a common reference set before performing primary analysis requires close cooperation of many labs and groups in large consortia, involving hundreds of collaborators.[109] Some consortia go beyond meta analysis and combine the raw genotype data to perform mega-analysis.[110] This allows across cohort quality control before imputation and allows centralized analysis of the complete dataset. However, not all cohorts are legally allowed to store genetic data offsite. The use of consortium driven meta- and mega- analysis has enabled identification of multiple risk loci for phenotypes as diverse as schizophrenia,[19] educational attainment,[111] height,[112] and sub-cortical brain volumes.[113]

Estimation of genetic (co)variance based on measured genotypes.

Genomic Relationship Matrix Restricted Maximum Likelihood (GREML)

As discussed above, SNPs are measured to economically cover a substantial portion of genetic variation in the human genome. Yang et al.[12] developed a method to estimate the variance in a given phenotype explained by all measured SNPs. This model is based on SNPs collected in distantly related individuals. The model can be formulated within the linear mixed model. The regression equation and variance decomposition of their model can be expressed as:

$$\mathbf{y} = \mathbf{Ab} + \mathbf{X}^t\mathbf{u} + \mathbf{e} \quad (19a)$$

$$\boldsymbol{\Psi} = \mathbf{X}^t\mathbf{X}/K \quad (19b)$$

$$\boldsymbol{\Sigma}(y) = \boldsymbol{\Psi}\otimes\sigma^2_u + \mathbf{I}\otimes\sigma^2_e \quad (19c)$$

where the matrix (NxP) **A** contains fixed regressors, such as principal components, sex and age, **X** (KxM) contains the standardized measured SNPs, and the off-diagonals of $\boldsymbol{\Psi}$ contains the average over the K loci of the allelic correlation estimates between individual. Given the nature of $\boldsymbol{\Psi}$, Yang et al.[12] name this the "Genetic Relatedness Matrix" or GRM, but we refer to it as a genetic covariance matrix. The regression model in equation 19a is similar to equation 17a, but differs in that the present model excludes any fixed SNP effect. That is, the focus is not on an individual SNP effect, but rather on the test of estimate of the variance attributable to all SNPs, i.e., σ^2_u. Furthermore, to ensure that σ^2_u reflects the effects of the measured SNPs, the sample is selected to include only distantly related individuals, i.e., characterized by off diagonal values in $\boldsymbol{\Psi}$ smaller than .025.[12] The inclusion of closely related individuals would introduce possible sources of bias in the estimate of σ^2_u. First, closely related individuals may share environmental influences contributing to the phenotypic resemblance. Second, closely related individuals will have values in $\boldsymbol{\Psi}$ converging on their expected additive genetic correlations (e.g., .5 in DZs, 1 in MZs). Their presence will bias the estimate of σ^2_u towards the value of σ^2_A, the total additive genetic variance. Yang

et al.[12] used this model to show that genotyped SNPs commonly used in GWAS may explain up to half the additive genetic variation in height. Further application of this model has shown that the SNPs used in GWAS account appreciable proportions of variance in complex phenotypes[15; 114-116] . Lee & Chow[117] present an extensive mathematical account of this model. User friendly software is available to fit this model, namely the GCTA software suite.[12]

Various extensions of this basic GREML model have been developed, and some are included in GCTA. One useful extension is to conduct the analysis for the SNPs on the individual chromosomes. Let $\mathbf{X}_c$ denote the standardized genotype matrix including only the K_c standardized SNPs on chromosome c, and let $\boldsymbol{\Psi}_c = \mathbf{X}_c{}^t\mathbf{X}_c/K_c$ (c=1,...,.22), then the variance decomposition is partitioned by chromosome as follows:

$$\boldsymbol{\Sigma}(y) = \boldsymbol{\Psi}_1 \otimes \boldsymbol{\sigma}^2_{u1} + \boldsymbol{\Psi}_2 \otimes \boldsymbol{\sigma}^2_{u2} + \ldots\ldots + \boldsymbol{\Psi}_{22} \otimes \boldsymbol{\sigma}^2_{u22} + \mathbf{I} \otimes \boldsymbol{\sigma}^2_{e}. \qquad (20)$$

The SNP variance components have been shown to correlate positively with chromosome length, as expected if on average a longer chromosomes harbor more causal SNPs.

The method of partitioning of SNP variance, as shown in Eq 20, has also been used to partition the SNP variance over different functional categories of SNPs.[118] These categories include SNPs in the portion of genome that is expressed (i.e. exonic SNPs), SNPs that regulate genes (e.g.,, DNaseI hyperactivity sites, gene promoter region, un-translated regions) , SNPs that alter gene function (i.e., coding variants), SNPs in intergenic regions (intronic regions). Gusev et al.[118] partitioned the variance in 11 common complex traits in functional categories of SNPs. They showed SNPs in regulatory regions explain a relatively large portion of phenotypic variance

A second extension implemented in GCTA generalizes the GREML model to the bivariate phenotypic case. In this model, the genetic covariance between traits attributable to the SNPs is estimated.[119] While such genetic covariance terms can be estimated readily in multivariate twin and family studies, the present estimate has the virtue of pertaining to the measured SNPs, and of being methodologically

independent of family and twin studies (with their many assumptions). Their method reduces to the following regression and (co) variance decomposition (see also Lee et al.[119]):

$$\mathbf{y_1} = \mathbf{A_1 b_1} + \mathbf{X_1^t u_1} + \mathbf{e_1} \qquad (21a)$$

$$\mathbf{y_2} = \mathbf{A_2 b_2} + \mathbf{X_2^t u_2} + \mathbf{e_2}$$

$$\mathbf{\Psi_{11}} = \mathbf{X}_1^t \mathbf{X_1}/K \qquad \mathbf{\Psi_{12}} = \mathbf{X}_1^t \mathbf{X_2}/K$$

$$\mathbf{\Psi_{21}} = \mathbf{X}_2^t \mathbf{X_1}/K \qquad \mathbf{\Psi_{22}} = \mathbf{X}_2^t \mathbf{X_2}/K \qquad (21b)$$

$$\mathbf{\Sigma}(y_1, y_2) = \boldsymbol{V} \qquad (21c)$$

$$V = \begin{matrix} \mathbf{\Psi_{11}} \otimes \sigma^2_{u1} + \mathbf{I} \otimes \sigma^2_{e1}, & \mathbf{\Psi_{12}} \otimes \sigma^2_{u12} \\ \mathbf{\Psi_{21}} \otimes \sigma^2_{u12}, & \mathbf{\Psi_{22}} \otimes \sigma^2_{u2} + \mathbf{I} \otimes \sigma^2_{e2} \end{matrix} \qquad (21d)$$

The bivariate (co)variance decomposition (Eq 21) is possible if each phenotype is measured in a distinct sample, but the precision of the covariance estimate is improves if the phenotypes are measured in the same sample. Application of this method to data collected by the psychiatric genetics consortium demonstrated a SNPs based genetic covariance between schizophrenia, bipolar disorder, and depression. A possible multivariate extension to this model, which allows for partitioning of the genetic variance of multiple matrices $\mathbf{\Psi}_c$ as in equation 20b is discussed in chapter 8 of this dissertation.

As the primary goal of GREML is to obtain an estimate of the variance explained by measured SNPs, closely related individuals are generally excluded, as explained above. A recently proposed model allows for the estimation of the variance attributable to SNPs in the presence of related individuals.[120] It specifically allows

for the estimation of both the variance attributable to SNPs and the variance attributable to the additive genetic influences. To fit this model we require individual who are closely and distantly related. Given data from such as sample, we can fit the following model:

$$\boldsymbol{\Sigma}(y) = \boldsymbol{\Psi}_s \otimes \sigma^2_u + \boldsymbol{\Psi}_{res} \otimes [\sigma^2_A - \sigma^2_u] + \mathbf{I} \otimes \sigma^2_e \qquad (22)$$

where $\boldsymbol{\Psi}_s = \mathbf{X}^t\mathbf{X}/K$ is the NxN matrix, as observed in the sample, and $\boldsymbol{\Psi}_{res}$ equals the NxN matrix $\boldsymbol{\Psi}_s$, with all off diagonals <.05 set to equal zero. As above σ^2_u is the additive genetic variance attributable to the K measured SNPs , and $[\sigma^2_A - \sigma^2_u]$ is the residual variance (i.e., the total additive genetic variance minus the SNP variance). Given we can estimate σ^2_u and $[\sigma^2_A - \sigma^2_u]$, we can calculate σ^2_A. Zaitlen et al.[120] applied this model to a large number of phenotypes. Chapter 7 of this dissertation includes results obtained with this model pertaining to body mass index, height, anxious depression, and attention problems.

The GCTA software suit also allows for the assessment of gene-environment interaction, given a binary environmental variable, coded 0/1 .[12]

$$\boldsymbol{\Sigma}(y) = \boldsymbol{\Psi} \otimes \sigma^2_u + \boldsymbol{\Psi}_{ge} \otimes \sigma^2_{ge} + \mathbf{I} \otimes \sigma^2_e \qquad (23)$$

Where $\boldsymbol{\Psi}$ is the NxN genetic covariance matrix as describe above, and $\boldsymbol{\Psi}_{ge}$ equal the NxN matrix $\boldsymbol{\Psi}$,where all cells which correspond to the genetic covariance between a pair of individuals discordant for the environmental exposure are set to zero. The gene -environment model described above is extended in Chapter 7 of this dissertation.

Polygenic risk scores.

GCTA can be applied to demonstrate that all measured SNP, or a subset thereof, explain an appreciable portion of variance in a trait. Bivariate GCTA can be applied to determine whether a shared set of SNPs is associated with two distinct traits. GCTA can be applied to find evidence of a genetic effect even if no individual fixed effect is significant. However bivariate GCTA requires access to both the measured genotypic data and both the phenotypes. An alternative method to demonstrate the presence of signal in genetic markers, or pleiotropy between two traits is by means of polygenic risk scores. This involves the selection of a subset of SNPs which satisfy a given alpha level (not necessarily the genome-wide alpha of $5*10^{-8}$), and the calculation of the weighted linear combination of the SNPs in the set, where the weights are set to equal the regression coefficients associated with the individual SNPs (i.e., the parameter b_s in eq. 14, 15, 17a or 18a above). This linear function of the SNPs is called the polygenic risk score.

There regression coefficients used are derived from a genome wide meta analysis of a phenotype of interest (i.e the discovery sample). The polygenic scores are then calculated for individuals that are not included in the discovery sample(i.e. the target sample). To determine the presence of signal in the genetic markers the phenotype, measured in the target sample, is regressed on the polygenic risk score derived from the discovery sample. To determine genetic overlap between two traits, one can obtain weights form a meta analysis of schizophrenia GWAS and use the polygenic scores based on these weights to predict bipolar disorder.[121] The regression of the phenotype on the risk score is expected to be significant (i.e., explained variance, $R^2 > 0$) if the set of SNPs are associated with the phenotype of interest. The alpha used for inclusion of SNPs in the polygenic score, may be set at varying values to assess the effect on the explained variance of progressively less stringent alpha (e.g, $\alpha = 0.001$, $\alpha = 0.01$, $\alpha = 0.1$, and $\alpha = .2$). The discovery and target sample need to be independent, dependency may result in overestimation and false positives.

While polygenic risk scores are often found to be predictive, the predictive power, even if based on the best available discovery samples, has been found to be too low to be clinically relevant.[17] For a full discussion of the common pitfalls associated with calculation and usage of polygenic risk scores see Wray et al.[122] For application of polygenic risk score in psychology and psychiatry see: [16; 20; 21; 123]

and chapter 5 of this dissertation. Software packages are available to handle data management and computation of polygenic risk scores such as Plink[80] or PRSice[124].

This thesis

In the present chapter I reviewed historical and current practice in behavior genetics and discussed the different techniques used in terms of regression models. This regression approach allows the reader to compare the different methods used. While procedures and methods used to fit the regression models in twin studies, linkage and association can differ, the models are compatible. The biggest change in the field of behavior genetics has been to move from latent inferred genotypic variables (twin and family studies) via latent genetic variables inferred from measured loci (linkage), to directly observed genetic variables (association). In the following chapters of this dissertation many of the techniques discussed here are applied to population based measured of psychopathology in children and in adults. Chapter 2 involves a longitudinal twin study of anxious depression, assessing heritability and stability between ages 3 and 60+. Chapter 3 aims to describe the development of broad measures of internalizing and externalizing psychopathology between childhood and adolescence, using growth models. Chapter 4 is a genome wide meta-analysis of preschool internalizing problems. In Chapter 5 the overlap between polygenic risk scores based on the schizophrenia GWA meta-analysis and childhood psychopathology is tested. Chapter 6 also studies anxiety, aiming to replicate a gene associated with anxiety in rats in humans. In Chapter 7 methods are developed to extend the GREML model and GCTA framework to allow for gene -environment interaction. In Chapter 8 a method to rapidly estimate genetic covariance in the context of GREML/GCTA is developed and tested. Chapters 9 and 10 offer a summary of the work and discuss it broader implications.

References

1. Fisher RA (1918) The correlation of relatives on the assumption of Mendelian inheritance.

2. Falconer D.S., Mackay T.F. (1996) Introduction to quantitative genetics. Harlow. UK: Longman

3. Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. 1 ed. Sinauer Associates., Sunderland, MA:

4. Plomin Robert, DeFries John C, Knopik Valerie S, Neiderhiser Jenae M (2013) Behavioral Genetics. McMillan Education, New York City

5. van der Weyden L., White J.K., Adams D.J., and Logan D.W. (2011) The mouse genetics toolkit: revealing function and mechanism. Genome Biol *12*, 224

6. Martin N.G., Eaves L.J. (1977) The genetical analysis of covariance structure. Heredity *38*, 79-95

7. Posthuma D., Beem A.L., De Geus E.J., Van Baal G.C., von Hjelmborg J.B., Iachine I., and Boomsma D.I. (2003) Theory and practice in quantitative genetics. Twin research *6*, 361-376

8. McGue M., Bouchard Jr T.J. (1998) Genetic and environmental influences on human behavioral differences. Annual review of neuroscience *21*, 1-24

9. Bouchard T.J., McGue M. (2003) Genetic and environmental influences on human psychological differences. Journal of neurobiology *54*, 4-45

10. Wray N.R., Middeldorp C.M., Birley A.J., Gordon S.D., Sullivan P.F., Visscher P.M., Nyholt D.R. et al. (2008) Genome-wide linkage analysis of multiple measures of neuroticism of 2 large cohorts from Australia and the Netherlands. Archives of General Psychiatry *65*, 649-658

11. Visscher P.M., Brown M.A., McCarthy M.I., and Yang J. (2012) Five years of GWAS discovery. The American Journal of Human Genetics *90*, 7-24

12. Yang J., Lee S.H., Goddard M.E., and Visscher P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics *88*, 76-82

13. Visscher P.M., Yang J., and Goddard M.E. (2010) A commentary on "common SNPs explain a large proportion of the heritability for human height" by Yang et al.(2010). Twin Research and Human Genetics *13*, 517-524

14. So H.C., Li M., and Sham P.C. (2011) Uncovering the total heritability explained by all true susceptibility variants in a genome–wide association study. Genetic epidemiology *35*, 447-456

15. Lubke G.H., Hottenga J.J., Walters R., Laurin C., De Geus E.J., Willemsen G., Smit J.H. et al. (2012) Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. Biological psychiatry *72*, 707-709

16. Purcell S.M., Wray N.R., Stone J.L., Visscher P.M., O'Donovan M.C., Sullivan P.F., Sklar P. et al. (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature *460*, 748-752

17. Dudbridge F. (2013) Power and predictive accuracy of polygenic risk scores. PLoS genetics *9*, e1003348

18. Wray N.R., Lee S.H., Mehta D., Vinkhuyzen A.A., Dudbridge F., and Middeldorp C.M. (2014) Research Review: Polygenic methods and their application to psychiatric traits. Journal of Child Psychology and Psychiatry *55*, 1068-1087

19. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014) Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421-427

20. de Zeeuw E.L., van Beijsterveldt C.E., Glasner T.J., Bartels M., Ehli E.A., Davies G.E., Hudziak J.J. et al. (2014) Polygenic scores associated with educational attainment in adults predict educational achievement and ADHD symptoms in children. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics *165*, 510-520

21. Groen-Blokhuis M.M., Middeldorp C.M., Kan K.J., Abdellaoui A., van Beijsterveldt C.E., Ehli E.A., Davies G.E. et al. (2014) Attention-deficit/hyperactivity disorder polygenic risk scores predict attention problems in a population-based sample of children. Journal of the American Academy of Child & Adolescent Psychiatry *53*, 1123-1129

22. Neale M, Cardon L (1992) Methodology for genetic studies of twins and families. Vol. 67. Springer Science & Business Media,

23. Molenaar P.C., Boomsma D.I., and Dolan C.V. (1993) A third source of developmental differences. Behavior Genetics *23*, 519-524

24. Freund J., Brandmaier A.M., Lewejohann L., Kirste I., Kritzler M., Kr++ger A., Sachser N. et al. (2013) Emergence of individuality in genetically identical mice. Science *340*, 756-759

25. Eaves L.J. (1987) Including the environment in models for genetic segregation. Journal of psychiatric research *21*, 639-647

26. Boomsma D., Busjahn A., and Peltonen L. (2002) Classical twin studies and beyond. Nature Reviews Genetics *3*, 872-882

27. Martin N.G., Boomsma D.I., and Machin G. (1997) A twin-pronged attack on complex traits. Nat. Genet *17*, 387-392

28. Bollen KA (1989) Structural equations with latent variables. John Wiley & Sons,

29. Abdellaoui A., Hottenga J.J., de Knijff P., Nivard M.G., Xiao X., Scheet P., Brooks A. et al. (2013) Population structure, migration, and diversifying selection in the Netherlands. European journal of human genetics *21*, 1277-1285

30. Visscher P.M., Medland S.E., Ferreira M.A., Morley K.I., Zhu G., Cornes B.K., Montgomery G.W. et al. (2006) Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. PLoS genetics *2*, e41

31. Mather K, Jinks JL (1977) Introduction to biometrical genetics. Cambridge Univ Press,

32. Keller M.C., Coventry W.L. (2005) Quantifying and addressing parameter indeterminacy in the classical twin design. Twin Research and Human Genetics *8*, 201-213

33. Azzalini A (1996) Statistical inference based on the likelihood. 68 ed. CRC Press,

34. Boomsma D.I., Molenaar P.C. (1986) Using LISREL to analyze genetic and environmental covariance structure. Behavior Genetics *16*, 237-250

35. Fulker D.W., Baker L.A., and Bock R.D. (1983) Estimating components of covariance using LISREL. Data Analyst *1*, 5-8

36. Prescott C.A. (2004) Using the Mplus computer program to estimate models for continuous and categorical data from twins. Behavior Genetics *34*, 17-40

37. Maes H.H., Neale M.C., Medland S.E., Keller M.C., Martin N.G., Heath A.C., and Eaves L.J. (2009) Flexible Mx specification of various extended twin kinship designs. Twin Research and Human Genetics *12*, 26-34

38. Boker S., Neale M., Maes H., Wilde M., Spiegel M., Brick T., Spies J. et al. (2011) OpenMx: an open source extended structural equation modeling framework. Psychometrika *76*, 306-317

39. Wright S. (1922) Coefficients of inbreeding and relationship. American Naturalist, 330-338

40. Wright S. (1960) Path coefficients and path regressions: alternative or complementary concepts? Biometrics *16*, 189-202

41. Keller M.C., Medland S.E., Duncan L.E., Hatemi P.K., Neale M.C., Maes H.H., and Eaves L.J. (2009) Modeling extended twin family data I: description of the Cascade model. Twin Research and Human Genetics *12*, 8-18

42. Browne M.W. (1973) Generalized least squares estimators in the analysis of covariance structures. ETS Research Bulletin Series *1973*, i-36

43. Bliss C.I. (1934) The method of probits. Science *79*, 38-39

44. Fisher R.A. (1935) The Case of Zero Survivors in Probit Assays. Annals of Applied Biology *22*, 164-165

45. Boomsma D.I., Molenaar P.C. (1987) The genetic analysis of repeated measures. I. Simplex models. Behavior Genetics *17*, 111-123

46. Loehlin J.C. (1996) The Cholesky approach: A cautionary note. Behavior Genetics *26*, 65-69

47. McCrae R.R., Costa P.T. (1987) Validation of the five-factor model of personality across instruments and observers. Journal of personality and social psychology *52*, 81

48. McCrae R.R., Costa Jr P.T. (1999) A five-factor theory of personality. Handbook of personality: Theory and research *2*, 139-153

49. Franic S., Borsboom D., Dolan C.V., and Boomsma D.I. (2014) The big five personality traits: psychological entities or statistical constructs? Behavior Genetics *44*, 591-604

50. Heath A.C., Cloninger C.R., and Martin N.G. (1994) Testing a model for the genetic structure of personality: a comparison of the personality systems of Cloninger and Eysenck. Journal of personality and social psychology *66*, 762

51. Hewitt J.K., Eaves L.J., Neale M.C., and Meyer J.M. (1988) Resolving causes of developmental continuity or "tracking" I. Longitudinal twin studies during growth. Behavior Genetics *18*, 133-151

52. Kan K.J., Dolan C.V., Nivard M.G., Middeldorp C.M., van Beijsterveldt C.E., Willemsen G., and Boomsma D.I. (2013) Genetic and environmental stability in attention problems across the lifespan: evidence from the Netherlands twin register. Journal of the American Academy of Child & Adolescent Psychiatry *52*, 12-25

53. McArdle J.J. (2006) Latent curve analyses of longitudinal twin data using a mixed-effects biometric approach. Twin Research and Human Genetics *9*, 343-359

54. Cloninger C.R., Rice J., and Reich T. (1979) Multifactorial inheritance with cultural transmission and assortative mating. II. a general model of

combined polygenic and cultural inheritance. American journal of human genetics *31*, 176

55. Kendler K.S., Eaves L.J. (1986) Models for the joint effect of genotype and environment on liability to psychiatric illness. The American journal of psychiatry

56. Jinks J.L., Fulker D.W. (1970) Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of the human behavior. Psychological bulletin *73*, 311

57. van der Sluis S., Dolan C.V., Neale M.C., Boomsma D.I., and 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 Research and Human Genetics *9*, 377-392

58. Molenaar D., van der Sluis S., Boomsma D.I., Haworth C.M., Hewitt J.K., Martin N.G., Plomin R. et al. (2013) Genotype by environment interactions in cognitive ability: a survey of 14 studies from four countries covering four age groups. Behavior Genetics *43*, 208-219

59. Molenaar D., Dolan C.V. (2014) Testing systematic genotype by environment interactions using item level data. Behavior Genetics *44*, 212-231

60. Purcell S. (2002) Variance components models for gene–environment interaction in twin analysis. Twin research *5*, 554-571

61. Medland S.E., Neale M.C., Eaves L.J., and Neale B.M. (2009) A note on the parameterization of Purcell´s GxE model for ordinal and binary data. Behavior Genetics *39*, 220-229

62. Haworth C.M.A., Wright M.J., Luciano M., Martin N.G., De Geus E.J.C., Van Beijsterveldt C.E.M., Bartels M. et al. (2010) The heritability of general cognitive ability increases linearly from childhood to young adulthood. Molecular psychiatry *15*, 1112-1120

63. Boomsma D.I., De Geus E.J., Van Baal G.C., and Koopmans J.R. (1999) A religious upbringing reduces the influence of genetic factors on disinhibition: evidence for interaction between genotype and environment on personality. Twin research *2*, 115-125

64. Boomsma D.I., Martin N.G. (2002) Gene –environment interactions. Biological psychiatry, 181-187

65. Heath A.C., Jardine R., and Martin N.G. (1989) Interactive effects of genotype and social environment on alcohol consumption in female twins. Journal of Studies on Alcohol and Drugs *50*, 38

66. Heath A.C., Eaves L.J., and Martin N.G. (1998) Interaction of marital status and genetic risk for symptoms of depression. Twin research *1*, 119-122

67. Vink J.M., Bartels M., Van Beijsterveldt T.C., Van Dongen J., Van Beek J.H., Distel M.A., De Moor M.H. et al. (2012) Sex differences in genetic architecture of complex phenotypes? PloS one *7*, e47371

68. van der Sluis S., Posthuma D., and Dolan C.V. (2012) A note on false positives and power in Gx E modelling of twin data. Behavior Genetics *42*, 170-186

69. Scarr S., McCartney K. (1983) How people make their own environments: A theory of genotype > environment effects. Child development, 424-435

70. Eaves L.J. (1976) The effect of cultural transmission on continuous variation. Heredity *37*, 41-57

71. Eaves L.J. (1976) A model for sibling effects in man. Heredity *36*, 205-214

72. Carey G. (1986) Sibling imitation and contrast effects. Behavior Genetics *16*, 319-341

73. Dolan C.V., de Kort J.M., Van Beijsterveldt T.C., Bartels M., and Boomsma D.I. (2014) GE Covariance through phenotype to environment transmission: An assessment in longitudinal twin data and application to childhood anxiety. Behavior Genetics *44*, 240-253

74. Johnson W., McGue M., Krueger R.F., and Bouchard Jr T.J. (2004) Marriage and personality: a genetic analysis. Journal of personality and social psychology *86*, 285

75. Middeldorp C.M., Cath D.C., Vink J.M., and Boomsma D.I. (2005) Twin and genetic effects on life events. Twin Research and Human Genetics *8*, 224-231

76. McGue M., Lykken D.T. (1992) Genetic influence on risk of divorce. Psychological Science *3*, 368-373

77. Cederlof R., Friberg L., and Lundman T. (1977) The interactions of smoking, environment and heredity and their implications for disease etiology. A report of epidemiological studies on the Swedish twin registries. Acta Medica Scandinavica. Supplementum *612*, 1

78. Kendler K.S., Neale M.C., MacLean C.J., Heath A.C., Eaves L.J., and Kessler R.C. (1993) Smoking and major depression: a causal analysis. Archives of General Psychiatry *50*, 36-43

79. Groen-Blokhuis M.M., Middeldorp C.M., van Beijsterveldt C.E., and Boomsma D.I. (2011) Evidence for a causal association of low birth weight and attention problems. Journal of the American Academy of Child & Adolescent Psychiatry *50*, 1247-1254

80. Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., Bender D., Maller J. et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics *81*, 559-575

81. Dudbridge F. (2003) A survey of current software for linkage analysis. Human genomics *1*, 63-65

82. Haseman J.K., Elston R.C. (1972) The investigation of linkage between a quantitative trait and a marker locus. Behavior Genetics *2*, 3-19

83. Kruglyak L., Lander E.S. (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. American journal of human genetics *57*, 439

84. Morton N.E. (1955) Sequential tests for the detection of linkage. American journal of human genetics *7*, 277

85. Sham P.C., Zhao J.H., and Curtis D. (1997) Optimal weighting scheme for affected sib-pair analysis of sibship data. Annals of human genetics *61*, 59-67

86. Van Dongen J., Jansen R., Smit D., Hottenga J.J., Mbarek H., Willemsen G., Kluft C. et al. (2014) The Contribution of the Functional IL6R

Polymorphism rs2228145, eQTLs and Other Genome-Wide SNPs to the Heritability of Plasma sIL-6R Levels. Behavior Genetics *44*, 368-382

87. Hamer D.H. (2000) Beware the chopsticks gene. Molecular psychiatry *5*, 11-13

88. Spielman R.S., McGinnis R.E., and Ewens W.J. (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). American journal of human genetics *52*, 506

89. Fulker D.W., Cherny S.S., Sham P.C., and Hewitt J.K. (1999) Combined linkage and association sib-pair analysis for quantitative traits. The American Journal of Human Genetics *64*, 259-267

90. Posthuma D., De Geus E.J.C., Boomsma D.I., and Neale M.C. (2004) Combined linkage and association tests in Mx. Behavior Genetics *34*, 179-196

91. Neale B.M., Purcell S. (2008) The positives, protocols, and perils of genome–wide association. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics *147*, 1288-1294

92. Laurie C.C., Doheny K.F., Mirel D.B., Pugh E.W., Bierut L.J., Bhangale T., Boehm F. et al. (2010) Quality control and quality assurance in genotypic data for genome–wide association studies. Genetic epidemiology *34*, 591-602

93. Hoggart C.J., Clark T.G., De Iorio M., Whittaker J.C., and Balding D.J. (2008) Genome–wide significance for dense SNP and resequencing data. Genetic epidemiology *32*, 179-185

94. Daly M.J. (2007) Estimation of the Multiple Testing Burden for Genomewide Association Studies of Common Variants.

95. Pe'er I., Yelensky R., Altshuler D., and Daly M.J. (2008) Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genetic epidemiology *32*, 381-385

96. Price A.L., Patterson N.J., Plenge R.M., Weinblatt M.E., Shadick N.A., and Reich D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nature genetics *38*, 904-909

97. Li X., Basu S., Miller M.B., Iacono W.G., and McGue M. (2011) A rapid generalized least squares model for a genome-wide quantitative trait association analysis in families. Human heredity *71*, 67

98. Rogers W. (1994) Regression standard errors in clustered samples. Stata technical bulletin *3*

99. Williams R.L. (2000) A note on robust variance estimation for cluster–correlated data. Biometrics *56*, 645-646

100. Minicâ C.C., Dolan C.V., Kampert M.M., Boomsma D.I., and Vink J.M. (2014) Sandwich corrected standard errors in family-based genome-wide association studies. European journal of human genetics

101. Lippert C., Listgarten J., Liu Y., Kadie C.M., Davidson R.I., and Heckerman D. (2011) FaST linear mixed models for genome-wide association studies. Nature Methods *8*, 833-835

102. Yang J., Zaitlen N.A., Goddard M.E., Visscher P.M., and Price A.L. (2014) Advantages and pitfalls in the application of mixed-model association methods. Nature genetics *46*, 100-106

103. International HapMap Consortium (2005) A haplotype map of the human genome. Nature *437*, 1299-1320

104. Genomes Project Consortium (2012) An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56-65

105. Marchini J., Howie B., Myers S., McVean G., and Donnelly P. (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. Nature genetics *39*, 906-913

106. De Bakker P.I., Ferreira M.A., Jia X., Neale B.M., Raychaudhuri S., and Voight B.F. (2008) Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Human molecular genetics *17*, R122-R128

107. Howie B., Fuchsberger C., Stephens M., Marchini J., and Abecasis G.R. (2012) Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nature genetics *44*, 955-959

108. Willer C.J., Li Y., and Abecasis G.R. (2010) METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190-2191

109. Manolio T.A., Rodriguez L.L., Brooks L., Abecasis G., Ballinger D., Daly M., Donnelly P. et al. (2007) New models of collaboration in genome-wide association studies: the Genetic Association Information Network. Nature genetics *39*, 1045-1051

110. Sullivan P.F. (2010) The psychiatric GWAS consortium: big science comes to psychiatry. Neuron *68*, 182-186

111. Rietveld C.A., Medland S.E., Derringer J., Yang J., Esko T.+., Martin N.W., Westra H.J. et al. (2013) GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. Science *340*, 1467-1471

112. Wood A.R., Esko T., Yang J., Vedantam S., Pers T.H., Gustafsson S., Chu A.Y. et al. (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. Nature genetics *46*, 1173-1186

113. Hibar D.P., Stein J.L., Renteria M.E., rias-Vasquez A., Desrivi+¿res S., Jahanshad N., Toro R. et al. (2015) Common genetic variants influence human subcortical brain structures. Nature

114. Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nature genetics *45*, 984-994

115. Yang J., Benyamin B., McEvoy B.P., Gordon S., Henders A.K., Nyholt D.R., Madden P.A. et al. (2010) Common SNPs explain a large proportion of the heritability for human height. Nature genetics *42*, 565-569

116. Yang J., Manolio T.A., Pasquale L.R., Boerwinkle E., Caporaso N., Cunningham J.M., de Andrade M. et al. (2011) Genome partitioning of genetic variation for complex traits using common SNPs. Nature genetics *43*, 519-525

117. Lee J.J., Chow C.C. (2014) Conditions for the validity of SNP-based heritability estimation. Human genetics *133*, 1011-1022

118. Gusev A., Lee S.H., Neale B.M., Trynka G., Vilhjalmsson B.J., Finucane H., Xu H. et al. (2014) Regulatory variants explain much more heritability than coding variants across 11 common diseases. bioRxiv, 004309

119. Lee S.H., Yang J., Goddard M.E., Visscher P.M., and Wray N.R. (2012) Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics *28*, 2540-2542

120. Zaitlen N., Kraft P., Patterson N., Pasaniuc B., Bhatia G., Pollack S., and Price A.L. (2013) Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. PLoS genetics *9*, e1003520

121. Purcell S.M., Wray N.R., Stone J.L., Visscher P.M., O'Donovan M.C., Sullivan P.F., Sklar P. et al. (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature *460*, 748-752

122. Wray N.R., Yang J., Hayes B.J., Price A.L., Goddard M.E., and Visscher P.M. (2013) Pitfalls of predicting complex traits from SNPs. Nature Reviews Genetics *14*, 507-515

123. Wray N.R., Lee S.H., Mehta D., Vinkhuyzen A.A., Dudbridge F., and Middeldorp C.M. (2014) Research Review: Polygenic methods and their application to psychiatric traits. Journal of Child Psychology and Psychiatry *55*, 1068-1087

124. Euesden J., Lewis C.M., and O'Reilly P.F. (2014) PRSice: Polygenic Risk Score software. Bioinformatics, btu848

Chapter 2: Stability in symptoms of anxiety and depression as a function of genotype and environment: A longitudinal twin study from age 3 to 63 years

This chapter is based on the publication:

Nivard, M. G., Dolan, C. V., Kendler, K. S., Kan, K. J., Willemsen, G., van Beijsterveldt, C. E., Lindauer, R.J., van Beek, J.H.D.A., Geels, L.M., Bartels, M., Middeldorp, C.M. & Boomsma, D. I. (2014). Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. *Psychological medicine*.

Abstract

Background: The influence of genetic factors on major depressive disorder is lower than on other psychiatric disorders. Heritability estimates mainly derive from cross-sectional studies, and knowledge on the longitudinal etiology of symptoms of anxiety and depression (SxAnxDep) across the lifespan is limited. We aim to assess phenotypic, genetic and environmental stability in SxAnxDep between ages 3 and 63 years.

Method: Cohort-sequential design combining data from 49,524 twins followed from birth to age 20+, and from adolescence into adulthood. SxAnxDep were assessed repeatedly with a maximum of 8 assessments over a 25 year period. Data were ordered in 30 age groups and analyzed with longitudinal genetic models.

Results: Over age, there was a significant increase during adolescence in mean scores with sex differences (women > men) emerging. Heritability was high in childhood and decreased to 30-40% during adulthood. This decrease in heritability was due to an increase in environmental variance. Phenotypic stability was moderate in children (correlations across ages ~ .5) and high in adolescents (r=.6) , young adults (r=.7), and adults (r=.8). Longitudinal stability was mostly attributable to genetic factors. During childhood and adolescence there was also significant genetic innovation, which was absent in adults. Environmental effects contributed to short term stability.

Conclusions: The substantial stability in symptoms of anxiety and depression is mainly due to genetic effects. The importance of environmental effects increases with age and explains the relatively low heritability of depression in adults. The environmental effects are transient, but the contribution to stability becomes larger with age.

Introduction

Insight into the course and the etiology of variation in trajectories of psychopathology from childhood into adolescence and adulthood is required to address questions concerning origins and prognosis of psychopathology. In this paper we aim to unravel the causes of variation in trajectories of symptoms of anxiety and depression (SxAnxDep) between ages 3 and 63 years, and specifically to assess the extent to which such variation is caused by genetic factors. A large prospective cohort study would be the optimal design for identifying the importance of genetic etiological factors, but would require following the same subjects for 50 years or more. Here we made use of the long-term data collection in nearly 50,000 twins from the Netherlands Twin Register (NTR) over the past 25 years, and analyzed data on SxAnxDep reordered according to a cohort-sequential design.

Stability in symptoms and diagnoses of anxiety and depression is evident throughout the lifespan; stability is lowest between childhood and adolescence, and increases from adolescence into adulthood.[1-3] During adolescence, there is a rise in the prevalence of anxious and depressive symptoms and diagnoses, especially in women (with the exception of separation anxiety disorder and specific phobia), and there is a marked continuity of symptoms into adulthood.[1;4] When considering stability in longitudinal studies, outcomes depend in part on how anxiety and depression are measured. Greater stability of depression has been observed in continuous measures (e.g., symptom counts) than in binary measures indicating that individuals, who no longer satisfy the criteria for a diagnosis, may still have residual symptoms.[5]

A case has been made that studies of the genetic and environmental contributions to individual differences in health and disease are important to the understanding of illness.[6] In addition, such studies inform molecular genetics studies.[7] Knowledge regarding the contributions of genetic and environmental factors to the long-term course of symptoms of anxiety and depression is still lacking, as most studies using genetically informative subjects have been limited to cross-sectional or short-term follow-up analyses. Moreover, in these short-term follow-up analyses, the age range of the participants at the start of a longitudinal study is often greater than the duration of the follow-up. As a consequence, the age ranges of the participants at the different assessments overlap and age-specific explanations of variation over time might be missed.

The empirical evidence suggests that between ages 3 and 12 years genetic effects are important, and relatively stable, while adolescence is characterized by genetic innovations, i.e., the emergence of novel genetic effects. In an overview of cross-sectional genetic studies of young and adolescent twins, heritability estimates for anxiety, depression and internalizing symptoms range from 0% to 74%, but the majority is over 30%.[8] The great variability over studies is attributed to rater and age effects; in most studies heritability is higher for parental reports than for self reports, and heritability is higher at adolescence than in.[9] Longitudinal studies of children in the Netherlands and in the UK showed that stability in SxAnxDep between ages 3 to 12 is mainly attributable to stable genetic,[10;11] although genetic innovation was also evident in this age range. The period of transition from childhood to adolescence was investigated in Swedish twins with SxAnxDep, assessed repeatedly between ages 8 and 20.[12-14] Stability was partly explained by genetic factors. However, the influence of genetic innovation was large, and the contribution of genetic factors expressed during childhood declined in adolescence. Thus, genetic effects on SxAnxDep were developmentally dynamic from middle childhood to young adulthood.

In adulthood, heritability of anxiety disorders and major depression was estimated around 40.[15-17] For depression in middle age (50 to 70 years) heritability estimates ranged from 20% to 50%.[18] In one longitudinal twin study, participants aged 20 to 70 rated themselves on SxAnxDep twice or thrice with an interval of 10 or 20 years. Genetic effects on symptom scores showed a large degree of stability, with some evidence in females for new genetic effects on anxiety and depression in mid-life and later-life, respectively.[19]

The role of environmental effects on stability has been found to be smaller than the role of genetic factors. Environmental influences shared by members of the same family, often referred to as the common environment, contributed to stability in SxAnxDep during, [10;11] but effects were small and waned from around 15% explained variance in childhood to zero in adolescence and.[20] Environmental factors that are not shared by family members, referred to as unique environment, mainly seemed to have short-term effects on SxAnxDep. The impact of life events, for example, on the risk for major depression has been shown to last for one to three months.[21] However, two twin studies found enduring unique environmental effects from adolescence into adulthood and.[12;19] This was confirmed in a meta-analysis of longitudinally assessed SxAnxDep in eight samples of monozygotic twins, spanning an age range of 10 to 66 years.[22] Within-pair

differences between MZ twins in SxAnxDep increased from childhood into late adulthood. By middle adulthood environmental factors contributed substantially to stable and predictable individual differences in SxAnxDep.
The aim of the present study was to gain insight into the genetic architecture of symptoms of anxiety and depression (SxAnxDep) across the lifespan by analyzing SxAnxDep assessed with a standardized instrument in twins aged 3 to 63 years. Data collection spanned a period of 25 years, and entry into the study was at different ages, running from birth to old age. The data were reorganized according to a cohort sequential longitudinal design, covering the entire age range from age 3 to age 63 years with a maximum follow-up time of 25 years.

Methods

Subjects: Longitudinal survey data were collected in twins registered with the Netherlands Twin Register (NTR), which includes the Young NTR (YNTR)[23] and the Adult NTR (ANTR).[24] In the YNTR, twins have been registered at birth by their parents since 1987.[25] Maternal ratings at ages 3, 7, 10, and 12, and self-ratings at age 12, 14, 16 and 18 were included in the analysis. When young twins reach age 18, they are enrolled in the ANTR. The ANTR includes adolescent and adult twins, who were recruited through city councils and other means.[24] The twins completed the SxAnxDep subscale in 1991, 1995, 1997, 2000, 2002, and 2009. All twins between the ages of 12 and 63 were included in the current study. The total dataset comprised 49,524 twins, including 7,863 monozygotic (MZ) and 15,815 same-sex and opposite-sex dizygotic (DZ) complete twin pairs. The majority (60%) participated in more than one survey; with 10% taking part 5 times or more (see appendix I eTable 1). To analyze the data as a function of age, data were reordered into age bins spanning two years. For instance, a 21 year old twin in the 1991 and a 21 year old twin in the 2003 survey were both included in the 20-21 year group. Up to age 30, all age bins included more than 1000 observations. Across the entire dataset, no 2 year age bin included fewer than 130 observations.
Phenotypes: SxAnxDep scores were obtained from the “anxious-depressed“ subscale from the age appropriate questionnaires from the Achenbach System of Empirically Based Assessment (ASEBA): the Child Behavior Checklist CBCL/1.5-5 [26] and CBCL/4-18,[27] the Youth Self Report/YSR,[28] and the Adult Self Report/ASR.[29] The instruments were designed to measure comparable constructs over the ages, and are similar in item content. Mothers were asked to rate on a 3-point scale *("not true", "somewhat or sometimes true", "very or often true")* the extent to

which a statement described their child. An example item of the mother rating SxAnxDep scale is “Unhappy, Sad or Depressed”. In the self-rating scales, used from age 12 onward, the item is phrased as “I am Unhappy, Sad or Depressed”, and the response format was the same. Similar strong parallels exist for the other items for mother rated and self rated SxAnxDep. Full sample questionnaires can be found at http://www.aseba.com. At all ages, except age 3, the number of items was the same (9 items at age 3, 16 items at all other ages). The instrument has been found to be measurement invariant in adolescence across age and sex.[30] While the CBCL, YSR and ASR “anxious depressed’ scale score is a good predictor of anxiety disorders and depression diagnosis, a high score is not equivalent to a diagnosis.[31] Using composite international diagnostic interviews (CIDI), DSM-IV anxiety and depressive diagnoses were assessed in 1331 (345 cases, 986 controls) adults from the present sample in 1997 and 2007. We compared the SxAnxDep scores of these subjects with their CIDI diagnoses using a ROC curve analysis. The area under the curve for all anxiety and mood disorders was fair at .76. For major depressive disorder, the area under curve was .75, and for generalized anxiety disorder .78. The use of a continuous measure is consistent with the dimensional view of psychopathology, and with diagnoses based on continuous measures.[32]
Statistical model: Structural equation modeling was employed to analyze the mean trend across age, sex differences in the mean trend, and the covariance structure of SxAnxDep across age. Estimates of the mean trend for men and women are further analyzed using weighted least squares (WLS) (see Supplementary information). A genetic simplex model[33;34] was chosen to analyze the longitudinal data and to estimate heritability at each age as well as the phenotypic, genetic, and environmental stability across the lifespan. Figure 1 provides a graphical description of the model, which is further detailed in the supplemental material online. In brief, the model allows partitioning the variance in the observed data into variance due to additive genetic factors (A), unique environment (E, not shared between twins) and a common environment (C, shared between twins within a family). The model further allows estimation of the stability of the effects of genetic and environmental factors over age, and to establish the extent to which new effects, called innovations, come into play.[19] From the longitudinal model we derived heritability estimates at each age and calculated genetic and environmental correlations between the ages. For model details, and identification, please see the online supplementary information.

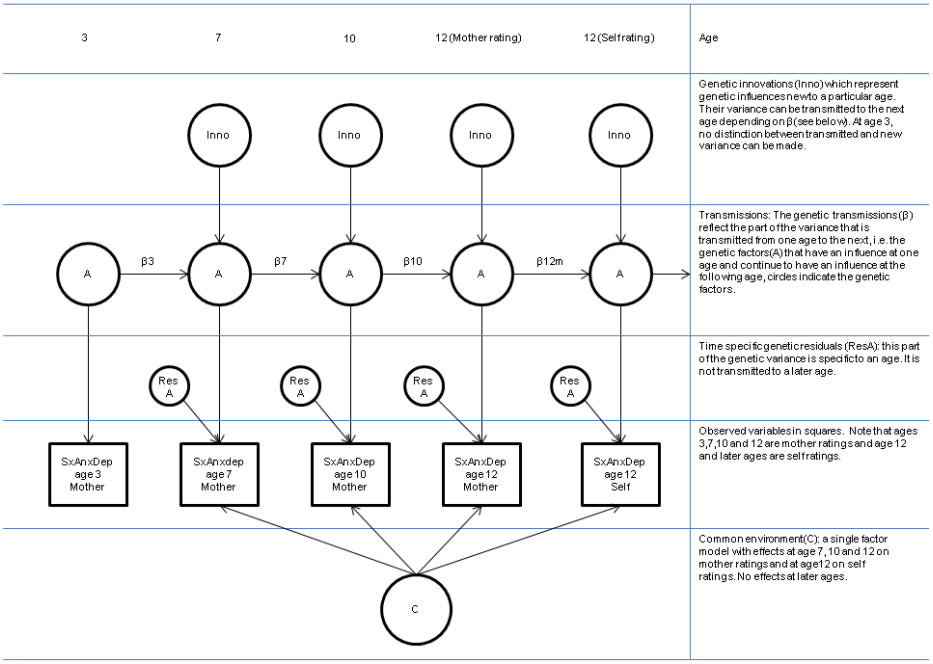


Figure 1**:** Path model representing the genetic simplex model. The simplex model describes an autoregressive process, in which latent genetic values (A) at age t are regressed (β) on previous latent values (t-1). In addition, at each age, novel genetic influences, called innovations (Inno) may come into play. Genotypic values thus consist of a part that is transmitted and a part that is innovation (except at the first age at which data are observed, where such a distinction cannot be made). The residual variance in the phenotype may also be influenced by additive genetic (rA) factors. The unique environmental (E) process is not depicted to avoid clutter but is structurally identical to the genetic process. The common environment shared between twins is modeled as a single factor, loading on observed SxAnxDep scores at ages 7,10,12 (mother) and age 12 (self) ratings.

Results

Figure 2 shows means for men and women and the post-hoc fitted mean trends. The mean in males for SxAnxDep in childhood was 2.79. Between ages 12 and 28 the mean SxAnxDep scores significantly increased (b=.19, t=5.568, df=1, p<0.001). After age 28, the means in men no longer significantly changed (b=-.06, t=-1.922, df=1, p=0.062). In females, the mean SxAnxDep scores in childhood was 2.97. Between the ages of 12 and 28 the mean SxAnxDep scores significantly increased (b=.68, t=14.35, df=1, p<0.001). Ater age 28 the SxAnxDep scores in women significantly declined (b=-.18, t=-5.477, df=1, p<0.001). The sex difference in SxAnxDep score in childhood was not significant (b=0.17, df=1, t=0.911, p =0.36), while the increase between age 12 and 28 was significantly steeper for females than for males (b=.43, t=6.665, p <0.001). In this trends analysis means were weighted for sample size (see Appendix I).

Table 1: Parameter estimates and variance components from the best longitudinal model. A: genetic factors, E unique environment, C common environment, h^2: proportion of variance explained by genetic factors = heritability, c^2 proportion of variance explained by common environmental factors, e^2 proportion of variance explained by unique environmental factors. See Figure 1 for an explanation of the terms transmission, innovation and residual variance.

t	Age	Transmission (regression) (β)		Innovation (variance) (ζ)		Residual (variance) (r)		C-Factor Loadings	Variance decomposition:		
		A	E	A	E	A	E	C*	h^2	c^2	e^2
1	3	0.35	0.03	6.51	2.87	0	0		0.69		0.31
2	7	1	0.59	3.15	2.05	0.97	1.01	0.38	0.61	0.02	0.38
3	10	0.87	0.57	1.26	2.48	0.97	1.01	0.48	0.58	0.02	0.4
4	12(m)	0.48	0.46	1.22	1.53	0.97	1.01	0.58	0.61	0.03	0.36
5	12(s)	1.22	0.63	4.17	3.09	0.34	4.28	1.28	0.37	0.11	0.52
6	14	0.9	0.73	0.25	2.56	0.34	4.28		0.51		0.49
7	16	0.93	0.86	2.58	3.37	0.34	4.28		0.5		0.5
8	18	1	0.61	2.7	3.58	1.81	3.43		0.53		0.47
9	20	1	0.85	0	6.68	1.81	3.43		0.49		0.51
10	22	1	0.83	0	4.71	1.81	3.43		0.46		0.54
11	24	1	0.76	0	5.11	0.67	4.72		0.39		0.61
12	26	1	0.78	0	4.87	0.67	4.72		0.4		0.6
13	28	1	0.97	0	2.93	0.67	4.72		0.43		0.57
14	30	1	0.87	0	1.76	0	6.55		0.37		0.63
15	32	1	0.72	0	3.54	0	6.55		0.36		0.64
16	34	1	0.93	0	4.69	0	6.55		0.38		0.62
17	36	1	0.94	0	0.07	1.55	3.88		0.48		0.52
18	38	1	0.81	0	0	1.55	3.88		0.5		0.5
19	40	1	0.76	0	4.01	1.55	3.88		0.48		0.52
20	42	1	0.75	0	7.63	0.65	3.8		0.4		0.6
21	44	1	0.74	0	3.23	0.65	3.8		0.44		0.56
22	46	1	0.87	0	7.32	0.65	3.8		0.4		0.6
23	48	1	0.56	0	6.93	0.4	0		0.4		0.6
24	50	1	1.4	0	0.53	0.4	5.22		0.5		0.5
25	52	1	0.81	0	0	0.4	5.22		0.4		0.6
26	54	1	1.26	0	2.98	1.93	3.25		0.48		0.52
27	56	1	0.57	0	4.36	1.93	3.25		0.35		0.65
28	58	1	0.52	0	0.86	1.93	3.25		0.54		0.46
29	60	1	1.38	0	1.2	1.93	3.25		0.66		0.34
30	62			0	0.44	0	0		0.63		0.37

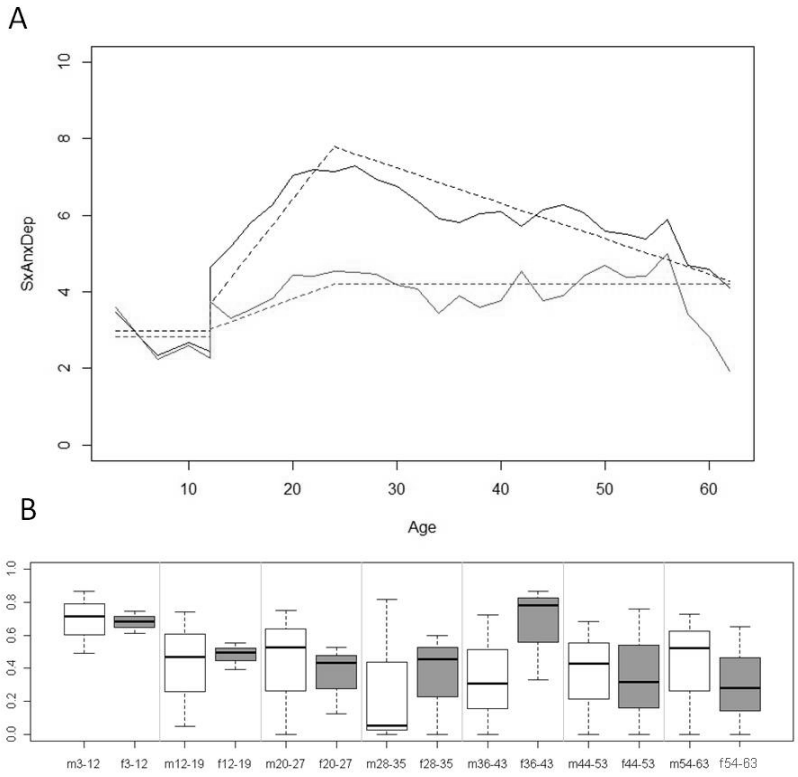


Figure 2: A: Mean trends in SxAnxDep across the lifespan for females (black) and males (grey). Solid lines are raw means. Dashed lines are fitted trend lines. **B:** Univariate heritability estimates across 7 broad ages (maternal ratings 3-12, selfreport ratings 12-19, 20-27, 28-35, 36-43, 44-53, 54-63) separately by sex (m= males, f=females).Figure 2B depicts the heritability estimates by sex and age for 7 broad age intervals, with their 99% confidence intervals. The proportion of variance attributable to genetic factors (i.e., heritability) in males and females differed little, except in the 36 to 43 year old groups.

The observed MZ and DZ twin correlations (presented in appendix I eTable 2) and the cross-twin cross-age correlations (eFigure 1) suggested a genetic simplex model as longitudinal correlations decreased over time and MZ correlations were higher than DZ correlations. Based on previous analyses in this sample[10] we expected the presence of common environmental effects in childhood.

Having fitted the genetic simplex model to the data (model fits Appendix I, table 3), we obtained heritability estimates for SxAnxDep at each 2 year age bin between 3 and 63 years. The estimates are given in Figure 3A, which shows that heritability (h^2, Figure 3A) declined from childhood (around 60 to 70%) to adulthood (around 40 to 50%). This decrease in heritability was due to an increase in environmental variance (V_E, figure 3B), and not to a decrease in genetic variance (V_A, Figure 3B), as can be seen in the plot of the absolute variance components estimates in Figure 3B. Part of the increase in the unique environmental variance is associated with the switch from mother ratings to self ratings (see the "jump" in Figure 3B at age 12). However, Figure 3B shows that the unique environmental variance continues to rise after the switch from mother to self ratings.

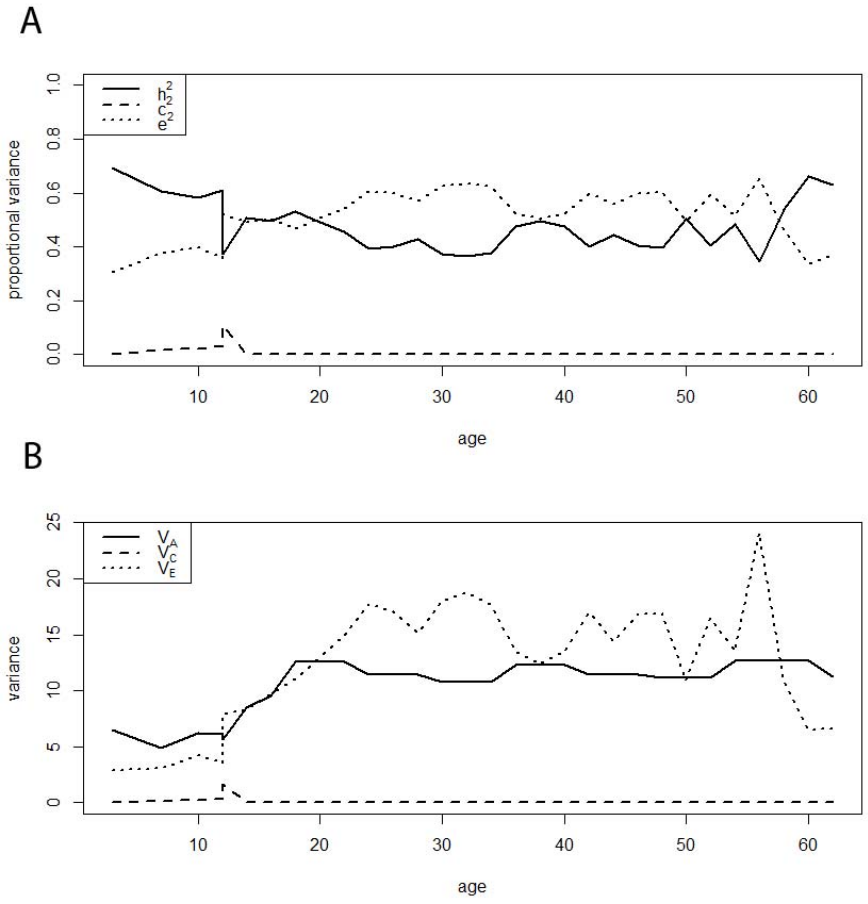


Figure 3: A: Proportions of variance explained by genetic factors (h^2; solid), common environment (c^2; dashed), and unique environment (e; dots) at each age as

erived from the model. **B:** variance components V_A, V_C and V_E at each age as derived from the model.

At age 12, we looked at the correlations between self ratings and maternal ratings and found that the agreement was moderate (phenotypic correlation = .35). The genetic and environmental contributions to this correlation were 56% (genetic), 27% (unique environmental), and 17% (shared environmental).

Based on the longitudinal model, we derived the phenotypic, genetic, and environmental correlations between SxAnxDep across different ages and visualized these in heat maps (Figure 4 and Appendix I figure 1). The phenotypic correlations (i.e., observed stability) between two successive ages ranged from .29 to .63 during childhood (up to age 12), from .48 to .70 during adolescence (age 12 to 18), from .64 to .77 between age 18 and age 32, and from .45 and .86 from age 32 onwards. Correlations between successive ages did not differ between males and females (t=- 0.57, df=814, p=.57). The genetic correlations between subsequent ages during childhood (mean r=.71, ranging from .4 to .94) and adulthood (mean r=0.92, ranging from .85 to 1) were large (Figure 4). Environmental correlations between subsequent ages were clearly lower than genetic correlations in childhood (mean r= .31, ranging from .03 to .47), and adulthood (mean r= .60, ranging from .47 to .73), although an increase was observed with age. The 10-year lag genetic correlations (Appendix I figure 2) were large (mean .92, ranging from .89 to .97), while 10-year lag environmental correlations were moderate (mean .32, ranging from .18 to .66, eFigure2) due to the relatively large proportion of environmental variance attributable to innovation All together, stability was largely attributable to genetic effects, while the increase in stability from childhood to adulthood was due to an increase in environmental correlations with age.

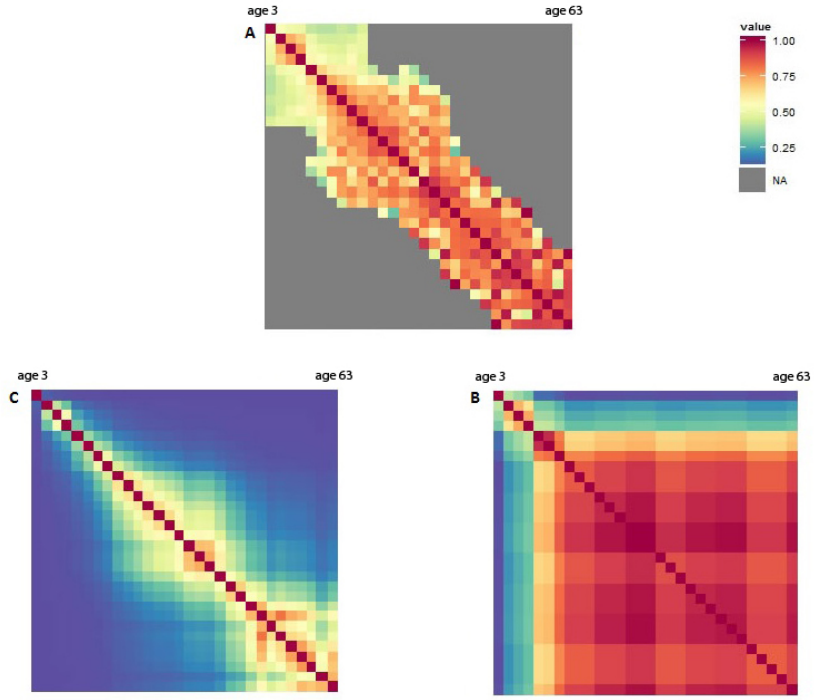


Figure 4: Correlation heatmaps to represent **A)** the observed correlations between ages 3 and 63, **B)** the model implied genetic correlations and **C)** the model implied unique environmental correlations

Discussion

The aim of this study was to gain insight into the etiology of the variation in symptoms of anxiety and depression (SxAnxDep) across the lifespan. Overall, genetic effects explained around 40% of the phenotypic variance at each age beyond age 7, and contributed greatly to the stability in SxAnxDep across age. We observed important differences between childhood and adulthood. Specifically, genetic innovation was observed during childhood and adolescence, but not after age 18. So after age 18, genetic effects are highly stable. This differs markedly from the results obtained for the unique environmental influences, which showed innovations at each age, and less stability. Unique environmental factors contributed primarily to short-time stability, but with increasing age, the contribution to stability of these environmental factors increased. This resulted in increasing stability in SxAnxDep with age corroborating pervious findings by Kendler *et al.* (2011).[22] This comprehensive picture of the etiology of the course of SxAnxDep is in line with previous studies of SxAnxDep in either childhood/adolescence or in adulthood.[10-13;19]

When analyzing data from children (maternal ratings) and adolescents and adults (self ratings), the question arises if the results are influenced by the different raters. The availability of self- and maternal reports of SxAnxDep at age 12 allowed us to address this question. We found a moderate correlation between mother and child ratings (.35). Moderate correlations among different raters have been observed before.[35] In the current dataset, we could establish that the correlation was largely attributable to genetic effects (56%), indicating that different raters to a large extent seem to agree on the genetically influenced SxAnxDep phenotype .

An important result from the current study is the clear explanation of the decline in heritability, which coincides with the switch from maternal to self-ratings. Heritability was highest in childhood (70%-50%), and dropped in adulthood (50%-35%), and remained stable to the age of 63. These heritability estimates are largely in line with those found in earlier studies.[8,15,17,18,20,36,37] The longitudinal data indicate that a decrease in heritability is not due to a decrease in genetic variance, but that there is an increase in environmental variance (see Figure 3), which leads to a *relatively* lower influence of genetic factors. Common environmental effects were only present at ages 7 to 12 and explained a relatively small proportion of the variance (~.2-.11%). Mean sex differences in SxAnxDep

emerged after age 12, around the same time when the increase in environmental variance is seen, peaked at age 28, and decreased thereafter. Figure 2B shows that there were no large sex differences in the estimates of genetic and environmental influences.

For clinical practice, it is important to note that environmental effects contribute to change, but also to short-term stability. This suggests that an improvement in SxAnxDep can be accomplished by positive environmental experiences, such as beneficial therapy or positive life events, and that increases of SxAnxDep can be caused by negative experiences, such as adverse life events. Importantly, these effects can endure for several years. An earlier study in part of this sample showed that SxAnxDep increase after negative life events, but that higher scores on SxAnxDep also precede negative life events.[38] Thus, individuals already suffering from anxiety or depression are at increased risk of experiencing negative events that can exacerbate their symptoms over an extended period of time. These results further underline the importance of addressing the environment in therapy (e.g., increasing social support or involving significant others). They can also suggest a possible interdependency between the individual and the environment, which may give rise to genotype-environment covariance[39] but it was shown that the association between life events and SxAnxDep was not explained by a shared genetic background. [38]

The results may have several implications for future research. The stability of genetic effects from childhood into adulthood is important as it indicates that genetic vulnerability is present from early onwards and remains a risk factor throughout life. These results suggest that gene-finding studies may include adults between 18 and 63, as we have observed little age-related heterogeneity in genetic effects. However, we recognize that high stability in polygenic effects does not rule out age effects at the level of a single causal genetic variant.

Rapee *et al.*[20] have already pointed out that with respect to childhood anxiety disorders "current knowledge of the role and mechanisms of environmental factors is especially poor…" (Rapee *et al.*[20] page 331). The increase in environmental stability over time is an intriguing finding that warrants further investigation. The question is whether new environmental effects on SxAnxDep, decrease with age, or whether the impact of events lasts longer as people age. Our results suggest the latter given that the environmental innovation parameters do not decrease with age, while the transmission parameters increase between age 3 and age 16 and then remain at a similar level. It is well known that

childhood and especially adolescence are characterized by large developmental changes in the brain.[40] Possibly, the maturation of the brain is accompanied by a more enduring effect of the environment. Kendler *et al.*[22] also found that the environmental effects on long term stability of SxAnxDep reach a plateau after adolescence.
We have shown that even during childhood and adolescence, part of the genetic and unique environmental innovations that appear at each age are transmitted to other ages. This signifies that there is a group of children and adolescents with a risk of enduring symptoms. Given the high disease burden for these children and society it is important to identify risk and protective factors that lead to stability over ages.

Finally, the current results for SxAnxDep are partly different from the results in a similar analysis of the etiology of stability in attention problems.[41] For attention problems, the genetic variance decreases from childhood to adulthood, while for SxAnxDep the genetic variance increases. Moreover, the increase in stability due to environmental effects that is observed for SxAnxDep after age 18 is not present in attention problems till the age of 30. Although it is of note that, for both phenotypes, genetic factors are most important in explaining stability, these differences indicate that the pattern of genetic and unique environmental factors throughout life do not need to be similar across psychiatric phenotypes. Our results and those for attention problems[41] suggest similar longitudinal genetic analysis are warranted for other psychiatric phenotypes.

Longitudinal results are essential to increase our understanding of the development and genetic architecture of psychiatric phenotypes. We made use of 25 years of data collection in subjects who entered the study at different ages, and tried to inform both clinical and molecular genetics research in psychiatry, using developmental and etiological informative models.

References

1. Beesdo-Baum K., Knappe S. (2012). Developmental epidemiology of anxiety disorders. Child and Adolescent Psychiatric Clinics of North America *21*, 457-478.

2. Hofstra M.B., van der Ende J., Verhulst F.C. (2000). Continuity and change of psychopathology from childhood into adulthood: a 14-year follow-up study. Journal of the American Academy of Child and Adolescent Psychiatry *39*, 850-858.

3. Richards D. (2011). Prevalence and clinical course of depression: a review. Clinical Psychology Review *31*, 1117-1125.

4. Costello E.J., Mustillo S., Erkanli A., Keeler G., Angold A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. Archives of General Psychiatry *60*, 837-844.

5. Ormel J., Oldehinkel T., Brilman E., van den Brink W. (1993). Outcome of depression and anxiety in primary care. A three-wave 3 1/2-year study of psychopathology and disability. Archives of General Psychiatry *50*, 759-766.

6. Collins F.S. (2004). The case for a US prospective cohort study of genes and environment. Nature *429*, 475-477.

7. van Dongen J., Slagboom E.P., Draisma H.H.M., Martin N.G., Boomsma D.I. (2012). The continuing value of twin studies in the omics era. Nature Reviews Genetics *13*, 640-653.

8. Rice F. (2009). The genetics of depression in childhood and adolescence. Current Psychiatry Reports *11*, 167-173.

9. Rice F., Thapar A. (2009). Depression and anxiety in childhood and adolescence: developmental pathways, genes and environment. In Handbook of Behavior Genetics, (ed. Y. K. Kim), pp. 377-394. Springer: New York.

10. Boomsma D.I., van Beijsterveldt C.E.M., Bartels M., Hudziak J.J. (2007). Genetic and environmental influences on anxious/depression: A longitudinal study in 3 to 12 year old children. In Genetic and environmental influences on developmental psychopathology and

wellness, (ed. J. J. Hudziak), American Psychiatric Association: Washington.

11. Trzaskowski M., Zavos H.M., Haworth C.M., Plomin R., Eley T.C. (2012). Stable Genetic Influence on Anxiety-Related Behaviours Across Middle Childhood. Journal of Abnormal Child Psychology *40*, 85-94.

12. Kendler K.S., Gardner C.O., Lichtenstein P. (2008). A developmental twin study of symptoms of anxiety and depression: evidence for genetic innovation and attenuation. Psychological Medicine *38*, 1567-1575.

13. Kendler K.S., Gardner C.O., Annas P., Neale M.C., Eaves L.J., Lichtenstein P. (2008). A longitudinal twin study of fears from middle childhood to early adulthood: evidence for a developmentally dynamic genome. Archives of General Psychiatry *65*, 421-429.

14. Kendler K.S., Gardner C.O., Annas P., Lichtenstein P. (2008). The development of fears from early adolesence to young adulthood: a multivariate study. Psychological Medicine *38*, 1759-1769.

15. Kendler K.S., Gatz M., Gardner C.O., Pedersen N.L. (2006). A Swedish national twin study of lifetime major depression. American Journal of Psychiatry *163*, 109-114.

16. Middeldorp C.M., Birley A.J., Cath D.C., Gillespie N.A., Willemsen G., Statham D.J., de Geus E.J., Andrews J.G., Van Dyck R., Beem A.L., Sullivan P.F., Martin N.G., Boomsma D.I. (2005). Familial clustering of major depression and anxiety disorders in Australian and Dutch twins and siblings. Twin Research and Human. Genetics *8*, 609-615.

17. Sullivan P.F., Neale M.C., Kendler K.S. (2000). Genetic epidemiology of major depression: review and meta-analysis. American Journal of Psychiatry *157*, 1552-1562.

18. Franz C.E., Lyons M.J., O'Brien R., Panizzon M.S., Kim K., Bhat R., Grant M.D., Toomey R., Eisen S., Xian H., Kremen W.S. (2011). A 35-year longitudinal assessment of cognition and midlife depression symptoms: the Vietnam Era Twin Study of Aging. The American Journal of Geriatric Psychiatry *19*, 559-570.

19. Gillespie N.A., Kirk K.M., Evans D.M., Heath A.C., Hickie I.B., Martin N.G. (2004). Do the genetic or environmental determinants of anxiety and

depression change with age? A longitudinal study of Australian twins. Twin Research *7*, 39-53.

20. Rapee R.M., Schniering C.A., Hudson J.L. (2009). Anxiety disorders during childhood and adolescence: origins and treatment. Annual Review of Clinical Psychology *5*, 311-341.

21. Kendler K.S., Karkowski L.M., Prescott C.A. (1998). Stressful life events and major depression: Risk period, long-term contextual threat, and diagnostic specificity. Journal of Nervous and Mental Diseases *186*, 661-669.

22. Kendler K.S., Eaves L.J., Loken E.K., Pedersen N.L., Middeldorp C.M., Reynolds C., Boomsma D.I., Lichtenstein P., Silberg J., Gardner C.O. (2011). The impact of environmental experiences on symptoms of anxiety and depression across the life span. Psychological Science *22*, 1343-1352.

23. van Beijsterveldt C.E.M., Groen-Blokhuis M., Hottenga J.J., Franić S., Hudziak J.J., Lamb D. et al. (2013). The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. Twin Research and Human Genetics *16*, 252-267.

24. Willemsen G., Vink J.M., Abdellaoui A., den Braber A., van Beek J.H.D.A., Draisma H.H.M. et al. (2013). The Adult Netherlands Twin Register: Twenty-Five Years of Survey and Biological Data Collection. Twin Research and Human Genetics *16*, 271-281.

25. Bartels M., van Beijsterveldt C.E.M., Derks E.M., Stroet T.M., Polderman T.J., Hudziak J.J., Boomsma D.I. (2007). Young Netherlands Twin Register (Y-NTR): a longitudinal multiple informant study of problem behavior. Twin Research and Human Genetics 10, 3-11.

26. Achenbach T.M., Rescorla L.A. (2000). Manual for the ASEBA Preschool - Forms & Profiles. University of Vermont, Research center for children, youth & families: Burlington, VT.

27. Achenbach T.M., Rescorla L.A. (2001). Manual for the ASEBA school age- Forms & Profiles. University of Vermont, Research center for children, youth & families: Burlington, VT.

28. Achenbach T.M. (1990). The Young Adult Self Report. University of Vermont, Dept of Psychiatry: Burlington, VT.

29. Achenbach T.M., Rescorla L.A. (2003). Manual for the ASEBA Adult- Forms & Profiles. University of Vermont, Research center for children, youth & families: Burlington, VT.

30. Fonseca-Pedrero E., Sierra-Baigrie S., Lemos-Giráldez S., Paino M., Muñiz J. (2012). Dimensional structure and measurement invariance of the Youth Self-Report across gender and age. Journal of Adolescent Health *50*, 148-153.

31. Bird H.R., Gould M.S., Rubio-Stiptec M.A.R.I., Staghezza B.M., Canino G.L.O.R. (1991). Screening for Childhood Psychopathology in the Community Using the Child Behavior Checklist. Journal of the American Academy of Child and Adolescent Psychiatry *30*, 116-123.

32. Markon K.E., Chmielewski M., Miller C.J. (2011). The reliability and validity of discrete and continuous measures of psychopathology: a quantitative review. Psychological Bulletin *137*, 856-879.

33. Boomsma D.I., Molenaar P.C.M. (1987). The genetic analysis of repeated measures. 1. Simplex models. Behavior Genetics *17*, 111-123.

34. Neale M.C., Boker S.M., Xie G., Maes H.H. (2006). Mx: Statistical modeling. Department of Psychiatry, Medical College of Virginia: Richmond (VA).

35. Ferdinand R.F., van der Ende J., Verhulst F.C. (2004). Parent-adolescent disagreement regarding psychopathology in adolescents from the general population as a risk factor for adverse outcome. Journal of Abnormal Psychology *113*, 198-206

36. Sakolsky D.J., McCracken J.T., Nurmi E.L. (2012). Genetics of pediatric anxiety disorders. Child and Adolescent Psychiatric Clinics of North America *21*, 479-500.

37. Johnson W., McGue M., Gaist D., Vaupel J.W., Christensen K. (2002). Frequency and heritability of depression symptomatology in the second half of life: evidence from Danish twins over 45. Psychological Medicine *32*, 1175-1185

38. Middeldorp C.M., Cath D.C., Beem A.L., Willemsen G., Boomsma D.I. (2008). Life events, anxious depression and personality: a prospective and genetic study. Psychological Medicine *38*, 1557-1565.

39. Dolan C.V., de Kort J.M., van Beijsterveldt C.E.M., Bartels M., Boomsma D.I. (2014). GE Covariance through phenotype to environment transmission: an assessment in longitudinal twin data. Behavior Genetics *44*, 240-253

40. Blakemore S.J. (2008). The social brain in adolescence. Nature Reviews Neuroscience *9*, 267-277.

41. Kan K.J., Dolan C.V., Nivard M.G., Middeldorp C.M., van Beijsterveldt C.E.M., Willemsen G., Boomsma D.I. (2013). Genetic and Environmental Stability in Attention Problems Across the Lifespan: Evidence From the Netherlands Twin Register. Journal of the American Academy of Child and Adolescent Psychiatry *52*, 12-25.

Chapter 3: Joint developmental trajectories of internalizing and externalizing disorders between childhood and adolescence.

Abstract

This study aimed to identify trajectories of DSM-IV based internalizing (INT) and externalizing (EXT) problem scores across childhood and adolescence, and to provide more insight into the co-morbidity by modeling the co-occurrence of internalizing and externalizing trajectories. INT and EXT were measured repeatedly between age 7 and age 15 years in over 7000 children and analyzed using growth mixture models. For both INT and EXT, five trajectories were identified including a very low, low, decreasing and increasing trajectory. In addition, an adolescent onset trajectory was identified for INT, and a stable high trajectory was identified for EXT. Multinomial regression showed that similar EXT and INT trajectories were associated. However, the adolescent onset INT trajectory was independent of high EXT trajectories, and persisting EXT was mainly associated with decreasing INT. Sex and early life environmental risk factors predicted EXT and, to a lesser extent, INT trajectories. The association between trajectories indicates the need to consider co-morbidity when a child presents with internalizing or externalizing disorders particularly when symptoms start early. This is less necessary when internalizing symptoms start at adolescence. Future studies should investigate the etiology of co-occurring INT and EXT, and the specific treatment needs of these severely affected children.

Introduction

Longitudinal epidemiological studies have shown that the prevalence of joint internalizing disorders (anxiety and depressive disorders) increases from childhood to adolescence into adulthood, whereas the prevalence of externalizing disorders (Attention Deficit Hyperactivity Disorder (ADHD), conduct disorder (CD) and oppositional defiant disorder(ODD)) decreases (Costello et al., 2011). [1] Research into the stability observed that symptoms and disorders partly persist from childhood into adulthood,[2;3] and that internalizing disorder in childhood predict externalizing disorder in adulthood and vice versa.[2-4] However, these studies did not specifically look at individual differences in trends of development. Therefore, a next step is to investigate the existence of subgroups (classes) with distinct developmental trajectories in internalizing and externalizing problems, using growth mixture modeling (GMM). Few studies using GMM have analyzed data from population based cohorts and covered the period from childhood to.[5-10] As pointed out by,[11] research in population based cohorts is necessary to get unbiased estimates of comorbidity and its risk factors. And since the largest changes in prevalence rates are observed in the transition from childhood to adolescence, it is critical to investigate this period.

In the current study we used GMM to model the development of DSM-IV based internalizing and externalizing problem scores (INT and EXT) measured at four occasions between age 7 and 15 years in a birth cohort of over 7,000 children. INT comprised anxiety disorders and depression, and EXT comprised ADHD, ODD and conduct disorder. The use of internalizing and externalizing summary scores is consistent with the results of several factor analytic studies of these disorders .[11-13] Following an initial separate analysis of internalizing and externalizing trajectories, we focused on the co-occurrence of these trajectories in a combined model. We further added risk factors such as sex, birth weight and social class to the model as predictors of class membership. The results provide insight into the trajectories of clinically relevant internalizing and externalizing problems across childhood and adolescence as well as into the association between the internalizing and externalizing trajectories over this period.

Based on the results of previous trajectory analyses of internalizing and externalizing psychopathology measured during childhood and adolescence in

population based samples[5-10] we expected for both INT and EXT a class of unaffected individuals. In addition, we expected at least a class with increasing symptoms for INT and a class with stable high and a class with decreasing symptoms for EXT. Although results of previous studies using GMM on internalizing symptoms are mixed regarding a class with persisting symptoms over time, we expected such a class given that other longitudinal studies suggest continuity over age.[2-3]

To our best knowledge, this study is the first to investigate the co-occurrence of internalizing and externalizing problems across childhood and adolescence. Two studies during childhood (till age 12) showed that similar trajectories were associated, e.g., children with internalizing problems were more often in the classes with moderate or high scores for externalizing problems.[14-15] We expected that this association continues into adulthood.

Methods

Subjects

The Avon Longitudinal Study of Parents and Children (ALSPAC, also known as "Children of the 90s", http://www.bristol.ac.uk/alspac/) is a long-term health research project project.[16] More than 14,000 mothers from the Avon County in the UK were enrolled during pregnancy in 1991 and 1992, and returned at least one questionnaire. When the oldest children were approximately 7 years of age, an attempt was made to bolster the initial sample with eligible cases who had failed to join the study originally.

The (psychological) health and development of these children has been followed in great detail. At ages 7, 10, 13, and 15 years, DSM-IV psychiatric disorders were assessed as part of the regular assessments. In total, 7202 children were assessed at least once for psychiatric disorders and had data available on risk factors (see table 1). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Please note that the study website contains details of all the data available through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary).

Instruments

The Development And Wellbeing Assessment (DAWBA) was used to establish DSM-IV psychiatric diagnoses.[17] We analyzed externalizing disorders, including ODD, CD and ADHD, and internalizing disorders, including major depression, generalized anxiety disorder, specific phobia, social phobia (at age 7, 10, 13 and 15), separation anxiety disorder (at age 7, 10 and 13) and panic disorder and agoraphobia (at age 15). All disorders were assessed by the child's mother except for self reports of internalizing disorders at age 15.

The analyses were performed on "DAWBA bands", with each category indicating the probability to suffer from a psychiatric disorder as derived from the DAWBA psychiatric interview. Scores of 0, 1, 2, 3, 4, and 5 correspond to probabilities of <0.01%, 0.5%, 3%, 15%, 50%, and >70%, respectively, of satisfying diagnostic criteria. The DAWBA bands have shown a positive association with clinician-rated diagnosis (chance corrected kappa 0.4 to 0.7, sensitivity 0.4 to 0.8, and specificity 0.98 to 0.99), and a strong relation with indicators of mental health.[18] The internalizing DAWBA band scores (INT) and externalizing DAWBA band scores (EXT) reflect the probability of satisfying the diagnostic criteria of any internalizing or any externalizing disorder.[18] Since category 0 did not occur in all assessments, 0 and 1 scores were collapsed into a single category (i.e., <.5%). We used the DAWBA band scores as they provide more information than the dichotomous affected/unaffected variable.

We included maternal smoking during pregnancy (no/yes), maternal highest education (6 categories), maternal and paternal social class (6 categories), maternal age at delivery, birth weight, and sex as predictors of class membership. Maternal smoking was assessed at week 18, social class and educational attainment at week 32. Maternal age at delivery and birth weight were part of the pregnancy and child baseline data.

Statistical methods

First, a latent growth curve model was fitted to repeated measures of INT and EXT. This model included three factors, an intercept (I), linear slope (S), and quadratic slope (Q) factor. The quadratic slope factor allows for curvilinear

development. As we estimated the means (fixed effects) and the variances (random effects) of the I, S, and Q factors, this is a random effects model, implying that each child was characterized by his or her own unique growth curve.

Growth mixture models (GMM) extend the standard growth model with a latent class variable, featuring a distinct growth model within each latent class. Subjects with similar trajectories are grouped into classes in a data-driven fashion, since class membership is not known beforehand. Fixing the variances of the I, S, and Q factors to zero within each class results in a restrictive GMM (also known as latent class growth models, LCGMs), in which only average within class trajectories are estimated (i.e., means of I, S, and Q), and all variability within classes is considered to be occasion specific .[19] We fitted LCGM's as well as models allowing for within class individual differences in the intercepts, i.e., random intercepts.[20;21] Models with an increasing number of classes were considered. Mixture models with random slope or random quadratic terms often failed to converge or converged prohibitively slow and were therefore not considered.

Based on the best fitting separate models, a combined model of INT and EXT trajectories between age 7 and 15 was tested, in which the EXT latent categorical class variable (C_E,) was regressed on the INT latent categorical class variable (C_I) (see Figure 1). This multinomial regression analysis provided an omnibus test of the null hypothesis that INT and EXT classes are unrelated. Note that the direction of this regression is arbitrary, and has no effect on the interpretation of the results. Reversing the direction of the regression to EXT on INT would result in exactly the same model fit and parameter estimates. INT and EXT class variables, denoted C_I and $C_{E,}$ were also regressed on maternal social class, paternal social class, maternal educational level, maternal age at delivery, maternal smoking during pregnancy, birth weight of the child, and sex of the child to test whether these variables predict trajectories.

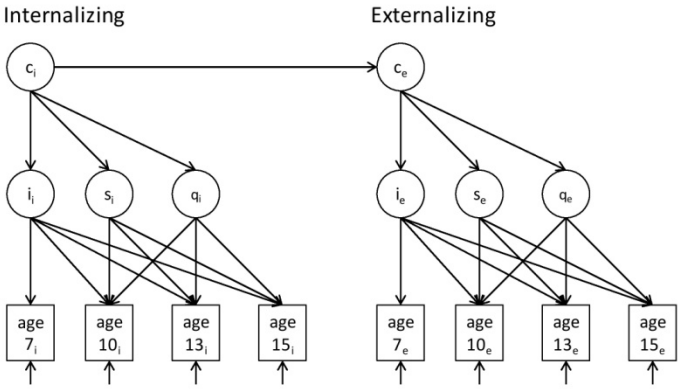


Figure 1. Path model of the final growth mixture model. C_i and C_e are the class variables indicating the distinct growth trajectories between the ages of 7 and 15 years for the internalizing and externalizing problem scores respectively. For each growth trajectory, an intercept (i), a linear slope (s) and a quadratic slope (q) are estimated. The factor loadings of the intercept (i) are 1. The factor loadings of the linear slope (s) are 1, 2 and 3, and of the quadratic slope (q) 1, 4 and 9. The association between the classes is modeled as a categorical regression of the externalizing trajectories on the internalizing trajectories.
Subscript i: internalizing disorders, subscript e: externalizing disorders

Models were fitted with M-plus 6.12[22] using robust full information maximum likelihood. If initial settings did not result in replicated minima, the number of starts was increased from 500 to 2000 and the number of final optimization from 50 to 200. If the best likelihood was not replicated with 2000 starts, the model was considered to have failed. The choice of the best fitting model was based on the sample size adjusted BIC. In case of small differences in fit, the preciseness of individual assignment to a specific trajectory (measured using model entropy) and the interpretability of the model were also considered.

Previous analyses showed that missingness in the ALSPAC data is not random, but that it only marginally affects parameter estimates in statistical analyses.[23] We regressed the covariates on the number of missing DAWBA assessments per individual. This attrition analysis showed that sex, smoking during pregnancy, maternal and paternal social class, maternal highest education and maternal age at delivery predicted missingness. As these variables were included in the model, data were assumed to be missing at random (MAR). We also reran the models on listwise complete data to evaluate the models based on individuals with complete data.

Results

Descriptives

Table 1 provides an overview of the prevalences of the observed DAWBA band scores and the polychoric correlations between INT and EXT at age 7, 10, 13 and 15. Polychoric correlations quantify the association between ordinal variables (Ekstrom, 2011).[24] The estimated prevalence of externalizing disorders in our sample was around 5% between ages 7 and 15. The prevalence of internalizing disorders was around 4% in childhood, and grew to 5% at age 15. As expected, male average EXT scores were greater than female average EXT scores at all ages, and female average INT scores were greater than male average INT at ages 13 and 15. Correlations between EXT and INT were around .20. Longitudinal correlations for INT between age 7 and 15 were .15 to .48, whereas correlations for EXT between the ages of 7 and 15 were higher (.35 to .61).

Table 1: For each age, the number of individuals for each DAWBA band score for externalizing (EXT) and internalizing (INT) disorders, prevalences for EXT and INT and total N (upper part), and the polychoric correlations between EXT and INT (lower part).

Descriptives

	DAWBA band score					Prevalence	N
	1	2	3	4	5		
EXT 7	1509	3948	612	160	96	0.058	6325
EXT 10	1668	3388	520	124	90	0.054	5790
EXT 13	1848	2682	608	115	79	0.055	5332
EXT 15	1714	1848	335	88	65	0.05	4050
INT 7	2749	3063	412	76	45	0.037	6345
INT 10	2092	3101	470	89	58	0.045	5810
INT 13	2181	2686	382	68	33	0.038	5350
INT 15	1506	1885	521	106	15	0.051	4033

Correlations

	EXT 7	EXT 10	EXT 13	EXT 15	INT 7	INT 10	INT 15
EXT 7	1						
EXT 10	0.63	1					
EXT 13	0.497	0.582	1				
EXT 15	0.387	0.451	0.569	1			
INT 7	0.23	0.161	0.102	0.083	1		
INT 10	0.188	0.257	0.151	0.133	0.485	1	
INT 13	0.162	0.177	0.211	0.173	0.406	0.483	
INT 15	0.081	0.08	0.074	0.11	0.142	0.19	1

Internalizing and externalizing trajectories

First, GMMs were fitted for INT and EXT separately. The single class model with a random intercept, slope and quadratic term showed a worse fit than models including a latent class variable, which indicates the existence of subgroups with different trajectories. Models with 2-6 classes were tested with 1) a *fixed* intercept, slope, and quadratic term, and 2) a *random* intercept, and a *fixed* slope and quadratic term. Table 2 provides the model fit and entropy. We retained the quadratic term as models with the quadratic term generally outperformed models without a quadratic term (results available on request from the first author). INT data were best described by a model with 5 classes with a fixed intercept, slope and quadratic term. However, the 4 class fixed model did not differ much in fit and had slightly better entropy then the 5 class model. Visual inspection of the trajectories in the 4 and 5 class model showed that the 5 class model contained an additional class of individuals who showed low internalizing problems up to age 13 and then increase at age 15. Given that a group with increasing internalizing scores in adolescence is consistent with the literature, we preferred the 5 class model to the 4 class model. For EXT, the best fitting model is the 3 class random intercept model. However, this model has a very low entropy compared to the fixed intercept models. Among the fixed intercept models, the best fitting model is the 6 class model, but the 5 class model has a substantially better entropy and only a slightly worse fit. Visual inspection of the trajectories showed that the 6 class model adds a third unaffected class to the very low and low classes, which starts out low and progresses to very low EXT scores. As this extra class is not very informative, the 5 class fixed intercept model was preferred. Results of the analyses of the listwise complete data were similar and also resulted in the selection of a 5 class model for both INT and EXT. Models fitted on listwise complete data had a higher entropy reflecting that individuals with more data available are easier to categorize.

Table 2: Fit indices for the internalizing (INT) and externalizing (EXT) growth mixture models containing 1 to 6 classes. On the left, the models with fixed effects for intercept (I), linear (S) and quadratic (Q) slopes i.e., with the variances of the intercept and slopes fixed to zero in each class. In the middle models with a random I (i.e., with the variance of the intercept estimated in each class) and fixed effects for S and Q. And on the right the reference model with random I, S and Q.

Models INT

N classes	**Fixed ISQ**			**Random I, Fixed SQ,**			**Random ISQ**	
	Entropy	AIC	Adj-BIC	Entropy	AIC	Adj-BIC	AIC	Adj-BIC
1							41748.603	41870.892
2	0.469	42326.785	42389.782	0.323	41946.080	42016.489	-	
3	0.432	41959.951	42063.711	0.455	41674.774	41789.651	-	
4	0.513	41650.452	41794.975	0.392	41536.956	41696.303	-	
5	0.488	41536.394	41721.680		*		-	
6	0.518	41505.755	41731.804		*		-	

Models EXT

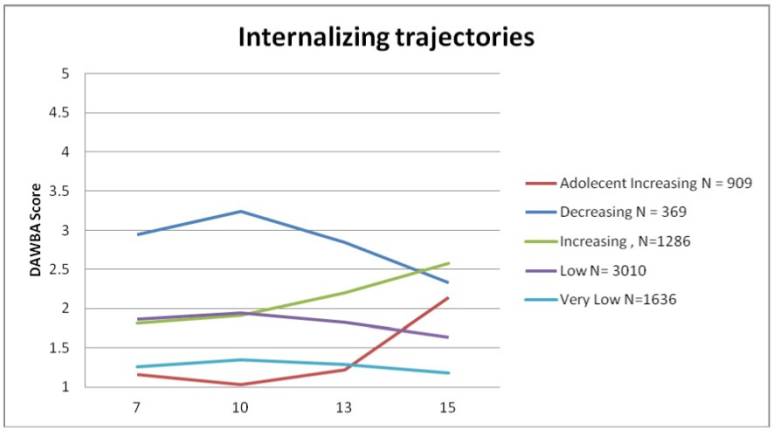
N classes	**Fixed ISQ**			**Random I, Fixed SQ,**			**Random ISQ**	
	Entropy	AIC	Adj-BIC	Entropy	AIC	Adj-BIC	AIC	Adj-BIC
1							41273.385	41395.628
2	0.544	42770.286	42833.259	0.383	41463.420	41533.803	-	
3	0.647	41630.458	41734.179	0.471	41235.535	41350.369	-	
4	0.647	41355.296	41499.765		*		-	
5	0.656	41208.013	41393.230		*		-	
6	0.569	41136.539	41362.503		*		-	

* best log-likelihood not replicated at 2000 starts and 200 final itterations

Combined internalizing/externalizing model.

In the combined model, the association between INT and EXT was analyzed using the multinomial logistic regression of the 5 class EXT trajectories on the 5 class INT trajectories (Figure 1). The model including the multinomial regression fitted the data better than a model that dropped this regression (Likelihood ratio: 477.894, df=16, $p < 0.0001$). We first describe the INT and EXT trajectories and then discuss the association between the EXT and INT trajectory class variables.

A)



B)

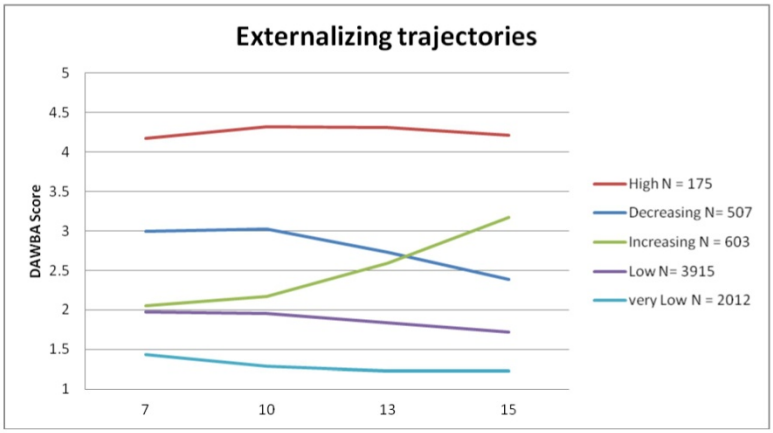


Figure 2. A) Internalizing trajectories. B) Externalizing trajectories. The y axis indicates the expected DAWBA band score for a given class at a given age.

For INT, there were two classes with low scores, called the very-low-INT class (22.7% of the sample based on most likely class membership) and the low-INT class (41.8%) (Figure 2A). A third class contained individuals with decreasing INT scores (5.1%). The remaining two classes contained individuals with increasing scores. The increasing-INT class (17.8%) showed a steady raise in score from childhood on, while in the adolescent-risk-INT class (12.6%) the scores are low until age 13 but sharply increase at age 15 years. Girls were significantly more likely than boys to be a member of the decreasing (OR= 2.011, $p<0.001$), increasing (OR= 7.800, $p<0.001$), or adolescent increasing INT classes (OR=3.128, $p<0.001$). Maternal smoking during pregnancy is a significant risk factor for being a member of the decreasing INT class (OR= 1.991, $p<0.001$). None of the other risk factors showed a significant effect.

Four of the five EXT trajectories show similar patterns as the INT trajectories (Figure 2B), i.e., the very low-EXT class (28%), the low-EXT class (54%), the decreasing-EXT class (7%) and the increasing-EXT class (8.3%). The final high-EXT class (2.4%) was different, as it contained individuals with persisting high scores from childhood to adolescence. Girls were significantly less likely than boys to be a member of the high (OR= 0.074, $p< 0.001$) , increasing (OR=0.475, $p<0.001$), and decreasing (OR= 0.178, $p<0.001$) EXT class. Maternal smoking during pregnancy predicted high (OR= 2.237, $p<0.001$), increasing (OR= 2.053, $p<0.001$), and decreasing (OR= 2.765, $p<0.001$) EXT class membership. A higher social class of the father resulted in a significant reduction of the chance of belonging to the high (OR= 0.787, $p<0.001$) or increasing EXT class (OR= 0.818, $p<0.001$). Higher maternal education is a protective factor for belonging to the increasing EXT class (OR=0.790, $p<0.001$).

Figure 3 displays the conditional probabilities of belonging to the EXT (INT) classes given membership of a given INT (EXT) class. These conditional probabilities show that similar internalizing and externalizing classes were associated. Focusing on the “affected” trajectories reveals that individuals in the decreasing INT class had a high probability of belonging to the decreasing EXT class (38%) and children in the increasing INT class had a substantial probability (22%) of being member of the increasing EXT class (Figure 3A). Vice versa, 27% of the children in the decreasing EXT class belonged to the decreasing INT class and children in the increasing EXT class had a substantial chance (46%) of belonging to the increasing

INT class (Figure 3B). It further becomes apparent that the high EXT class was particularly associated with the decreasing INT class and less with the increasing INT class, whereas the adolescent onset INT class was independent from EXT trajectories.

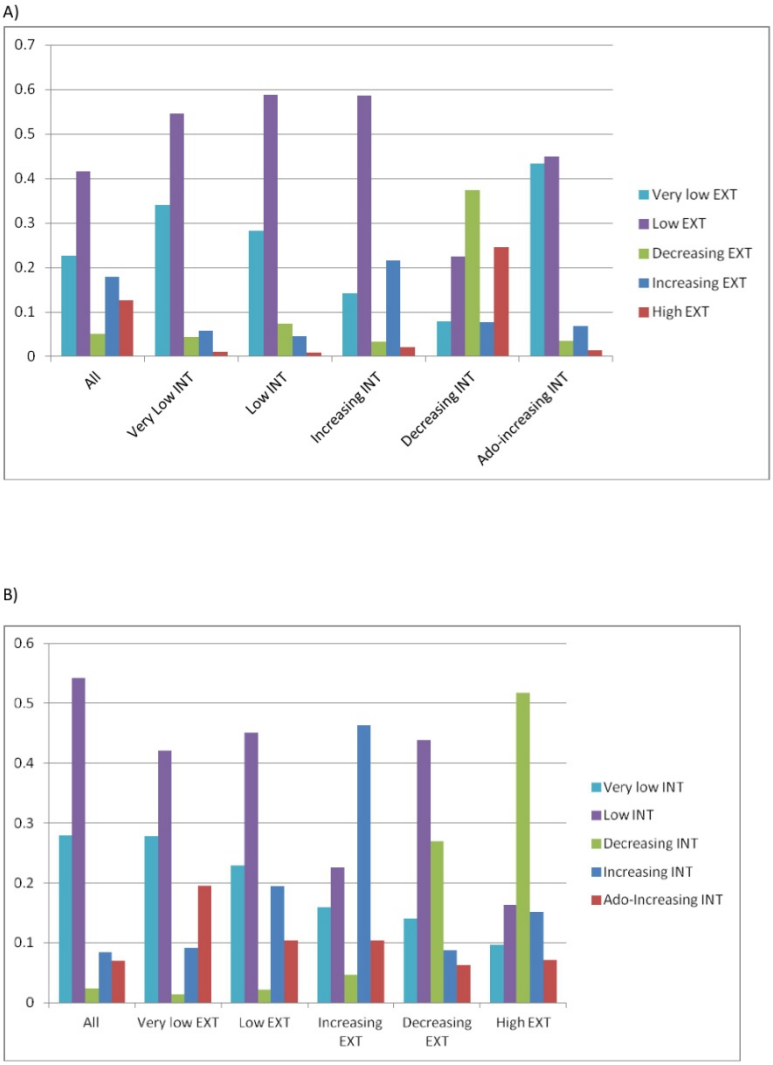


Figure 3: A) Probabilities for EXT class membership and conditional probabilities for EXT class membership given INT class membership. B) Probabilities for INT class membership and conditional probabilities for INT class membership given EXT class membership

Discussion

Using DSM-IV based internalizing and externalizing problem scores across childhood and adolescence, we identified five developmental trajectories for both INT and EXT, including trajectories with very low, low, increasing and decreasing scores. For INT, the last trajectory was characterized by increased scores from adolescence onwards, while for EXT, the last trajectory was characterized by individuals with stable high scores. The combined model showed that similar INT and EXT trajectories were associated. Only the adolescent onset INT group showed no association with affected EXT classes. Further, the high EXT group was most associated with the decreasing INT group signifying that in the group of children that have both externalizing and internalizing symptoms during childhood, in some of the children the externalizing symptoms will persist, whereas the internalizing symptoms attenuate.

Our findings for the INT and EXT trajectories are largely in line with *a priori* expectations. The most apparent differences are the absence of a class characterized by stable high INT and the presence of an increasing EXT class. The absence of a stable high INT group may be due to the relatively low prevalence of these disorders, especially during childhood. In GMM, it requires very large samples to reliably identify classes that include a small proportion of the sample. Our finding therefore cannot rule out a small group of children with persisting symptoms, as suggested by other studies.[2;3] A class with increasing externalizing symptoms was also identified by Van Lier et al.[8] for conduct disorder, and by Larsson et al.[10] for the inattentive subtype of ADHD. The increasing externalizing class probably comprises these groups.

Finally, the only two other studies[14,15] that looked at the combination of internalizing and externalizing trajectories up to age 12 also showed that increasing and decreasing internalizing and externalizing trajectories are mutually dependent. It will be interesting to see whether future studies modeling the trajectories into adolescence will replicate our finding that internalizing disorders starting at adolescence are independent of externalizing disorders and that persisting high EXT is mainly associated with decreasing INT. Our results indicate that the previously reported longitudinal associations between internalizing and externalizing disorders are due to the co-morbidity that exists between these

disorders at childhood. Thus, the recent finding in ALSPAC that adolescent depression is predicted by conduct problems in childhood[25] is probably due to the persistence of childhood internalizing symptoms.

Since mixture modeling as employed here is exploratory, results require validation.[26] However, the agreement between the found trajectories and our expectations underlines their importance. Moreover, the association with sex and the prenatal risk factors were all in the expected directions. Boys were found to be more at risk for externalizing problems and girls more for internalizing problems and adverse prenatal risk factors were associated with externalizing, and to a lesser extent, with internalizing problems.[27]

The use of two broad internalizing and externalizing problem scores could be considered a limitation. Studies investigating specific internalizing or externalizing symptom domains have detected differences in trajectories between the separate disorders.[8;10;28;29] Given the low prevalence rates of the separate disorders, such analyses were not feasible for the clinically based measure used in the current study. Moreover, previous studies have also shown that the analyzed disorders load on common factors interpretable as our INT and EXT (see[11-13]). This indicates that studies focusing on measures of a general tendency to display internalizing or externalizing disorders can also provide important information. Sample size also precluded separate analyses in males and females. Therefore, gender was included as a covariate predicting class membership. Finally, the correlations between internalizing and externalizing disorders (around 0.20) were lower than previously reported by Cosgrove et al.[13] (0.20 to 0.30). However, these differences are relatively small, and might be due to differences in the instrument used.

The strengths of this study were the use of a large population based sample with repeated measures from childhood into adolescence and, importantly, a DSM-IV based psychiatric interview instrument. This enables the translation of these findings to clinical relevance. The results suggest that if internalizing disorders first occur in adolescence, a brief screening for externalizing disorders will suffice. However, when confronted with childhood internalizing or externalizing problems, the association in course indicates the need to focus on co-morbidity from the start of the treatment. Future studies should address the specific

treatment needs of children with co-occurring internalizing and externalizing disorders, especially since the co-occurrence is related to negative outcomes .[15] An interesting question is whether successful treatment of an externalizing disorder also leads to a remission of the internalizing disorder, or vice versa, or whether treatment of both disorders is necessary. Moreover, it is important to identify the factors associated with the combination of the trajectories of decreasing internalizing and externalizing symptoms versus the combination of persisting externalizing symptoms and decreasing internalizing symptoms.

The current study does not address the etiology of comorbidity between internalizing and externalizing disorders. Different hypotheses currently exist about the causes of comorbidity. It has been suggested that depressive symptoms in ADHD are due to demoralization,[30] but in line with our finding that a combination of trajectories of externalizing symptoms in childhood and later increasing internalizing symptoms did not exist,[31] concluded that this does not explain all comorbidity. The opposite, i.e. internalizing symptoms causing externalizing symptoms, has also been hypothesized. Granic[32] ,for example, proposes three mechanisms explaining how anxiety can cause aggression and suggests how future research could investigate whether these mechanisms play a role. Another explanation for comorbidity is that multiple disorders are caused by the same underlying mechanism, which agrees with the observed co-occurrence of similar trajectories. This is supported by cross-sectional twin studies that indicated that co-morbidity between internalizing and externalizing disorders is partly explained by shared genetic risk factors (e.g. Cosgrove et al.[13]). It has already been shown that ADHD trajectories are influenced by genetic factors (Larsson et al.[10]). This could also be the case for co-occurring trajectories, which would be interesting for gene-finding studies. Including a genetic variant-by-course (i.e., decreasing or stable high) interaction term enables the identification of variants associated with a favorable or unfavorable outcome and reveal hints about biological differences in etiology between developmental courses.

To summarize, we showed that both internalizing and externalizing disorders can have a favorable or unfavorable course in time from childhood into adolescence and that trajectories are associated with each other. Future research should focus on unraveling the etiology of the co-occurrence, and focus on the development of treatment designs for the most seriously affected children

References

1. Costello, E.J., Copeland, W., & Angold, A. (2011). Trends in psychopathology across the adolescent years: What changes when children become adolescents, and when adolescents become adults? Journal of Child Psychology and Psychiatry, *52,* 1015-1025.

2. Hofstra, M.B., Van der Ende, J., & Verhulst, F.C. (2002). Child and Adolescent Problems Predict *DSM-IV* Disorders in Adulthood: A 14-Year Follow-up of a Dutch Epidemiological Sample. Journal of the American Academy of Child & Adolescent Psychiatry *41,* 182-189.

3. Hofstra, M B., Van der Ende, J., & Verhulst, F.C. (2000). Continuity and change of psychopathology from childhood into adulthood: a 14-year follow-up study. Journal of the American Academy of Child & Adolescent Psychiatry *39,* 850-858.

4. Copeland, W.E., Shanahan, L., Costello, E.J., & Angold, A. (2009). Childhood and adolescent psychiatric disorders as predictors of young adult disorders. Archives of General Psychiatry *66,* 764-772.

5. Dekker, M.C., Ferdinand, R.F., Van Lang, N.D., Bongers, I.L., van der Ende, J., & Verhulst, F.C. (2007). Developmental trajectories of depressive symptoms from early childhood to late adolescence: gender differences and adult outcome. Journal of Child Psychology and Psychiatry *48,* 657-666.

6. Letcher, P., Smart, D., Sanson, A., & Toumbourou, J.W. (2009). Psychosocial Precursors and Correlates of Differing Internalizing Trajectories from 3 to 15 Years. Social Development *18,* 618-646.

7. Toumbourou, J.W., Williams, I., Letcher, P., Sanson, A., & Smart, D. (2011). Developmental trajectories of internalising behaviour in the prediction of adolescent depressive symptoms. Australian journal of psychology *63,* 214-223.

8. Van Lier, P.A., Der Ende, J.V., Koot, H.M., & Verhulst, F.C. (2007). Which better predicts conduct problems? The relationship of trajectories of conduct problems with ODD and ADHD symptoms from childhood into adolescence. Journal of Child Psychology and Psychiatry *48*, 601-608.

9. Haltigan, J.D., Roisman, G.I., Susman, E.J., Barnett-Walker, K., & Monahan, K.C. (2011). Elevated trajectories of externalizing problems are associated with lower awakening cortisol levels in midadolescence. Developmental psychology *47,* 472-478.

10. Larsson, H., Dilshad, R., Lichtenstein, P., & Barker, E.D. (2011). Developmental trajectories of DSM IV symptoms of attention deficit/hyperactivity disorder: genetic effects, family risk and associated psychopathology. Journal of Child Psychology and Psychiatry *52,* 954-963.

11. Angold, A., Costello, E.J., & Erkanli, A. (1999). Comorbidity. Journal of Child Psychology and Psychiatry *40,* 57-87.

12. Beauchaine, T.P. & McNulty, T. (2013). Comorbidities and continuities as ontogenic processes: Toward a developmental spectrum model of externalizing psychopathology. Development and psychopatholog*, 25,* 1505-1528.

13. Cosgrove, V.E., Rhee, S.H., Gelhorn, H.L., Boeldt, D., Corley, R.C., Ehringer, M.A., Young, S.E., Hewitt, J.K.(2011). Structure and etiology of co-occurring internalizing and externalizing disorders in adolescents. Journal of abnormal child psychology 39*,* 109-123.

14. Brezo, J., Barker, E.D., Paris, J., Hébert, M., Vitaro, F., Tremblay, R.E., Turecki, G. (2008). Childhood trajectories of anxiousness and disruptiveness as predictors of suicide attempts. Archives of pediatrics & adolescent medicine *162,* 1015-1021.

15. Fanti, K.A. & Henrich, C.C. (2010). Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: findings from the National Institute of Child Health and Human Development Study of Early Child Care. Developmental psychology *46,* 1159-1175.

16. Golding, J., Pembrey, M., & Jones, R. (2001). ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology. Paediatric and Perinatal Epidemiology *15,* 74-87.

17. Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development and Well Being Assessment: Description and Initial

Validation of an Integrated Assessment of Child and Adolescent Psychopathology. Journal of Child Psychology and Psychiatry *41,* 645-655.

18. Goodman, A., Heiervang, E., Collishaw, S., & Goodman, R. (2011). The DAWBA bands as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. Social psychiatry and psychiatric epidemiology *46,* 521-532.

19. Nagin, D.S. (1999). Analyzing developmental trajectories: a semiparametric, group-based approach. Psychological methods *4,* 139-157.
20. Jung, T. & Wickrama, K.A.S. (2008). An introduction to latent class growth analysis and growth mixture modeling. Social and Personality Psychology Compass *2,* 302-317.

21. Muthen, B. & Muthen, L.K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. Alcoholism: Clinical and experimental research *24,* 882-891.
22. Muthén, L.K. & Muthén, B.O. (2007). Mplus User's Guide. Fifth Edition. Los Angeles, CA: Muthén & Muthén.

23. Wolke D., Waylen A., Samara M., Steer C., Goodman R., Ford T. & Lamberts, K. (2009). Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. The British Journal of Psychiatry *195*, 249-256.

24. Ekstrom, J. (2011). A generalized definition of the polychoric correlation coefficient. Department of Statistics, UCLA.

25. Stringaris, A., Lewis, G., & Maughan, B. (2014). Developmental pathways from childhood conduct problems to early adult depression: findings from the ALSPAC cohort. The British Journal of Psychiatry 205, 17-23.

26. Lubke, G. (2012). Old Issues in a New Jacket: Power and Validation in the Context of Mixture Modeling. Measurement: Interdisciplinary Research and Perspectives *10,* 212-216.

27. Ormel, J., Raven, D., van Oort, F., Hartman, C. A., Reijneveld, S. A., Veenstra, R. Vollebergh, W.A., Buitelaar, J., Verhulst, F.C., Oldehinkel,

A.J. (2014). Mental health in Dutch adolescents: a TRAILS report on prevalence, severity, age of onset, continuity and co-morbidity of DSM disorders. Psychological medicine, epub ahead of print.

28. Barker, E.D., Séguin, J R., White, H.R., Bates, M.E., Lacourse, E., Carbonneau, R. Tremblay, R.E. (2007). Developmental trajectories of male physical violence and theft: relations to neurocognitive performance. Archives of General Psychiatry *64,* 592-599.

29. Broeren, S., Muris, P., Diamantopoulou, S., & Baker, J.R. (2013). The course of childhood anxiety symptoms: developmental trajectories and child-related factors in normal children. Journal of Abnormal Child Psychology *41,* 81-95.

30. Brown, R.T., Borden, K.A., Clingerman, S.R., & Jenkins, P. (1988). Depression in attention deficit-disordered and normal children and their parents. Child psychiatry and human development *18,* 119-132.

31. Biederman, J., Mick, E., & Faraone, S.V. (1998). Depression in attention deficit hyperactivity disorder (ADHD) children:"True" depression or demoralization? Journal of Affective Disorders *47,* 113-122.

32. Granic, I. (2014). The role of anxiety in the development, maintenance, and treatment of childhood aggression. Development and psychopathology *26*, 1515-1530.

Chapter 4: A Genome-wide Association Meta-analysis of Preschool Internalizing Problems

This chapter is based on the publication:

Benke, K. S., **Nivard, M.G.,** Velders, F.P., Walters, R.K., Pappa, I., Scheet, P.A., Xiao X., Ehli E.A., Palmer, L.J. , Whitehouse, A.J.O., Verhulst, F.C. ,Jaddoe, V.W., Rivadeneira, F. , Groen-Blokhuis, M.M., van Beijsterveldt, C.E.M. , Davies, G.E. , Hudziak, J.J., Lubke, G.H., Boomsma, D.I. , Pennell, D. E., Tiemeier, H. , Middeldorp, C.M. ,for the Early Genetics and Lifecourse Epidemiology Consortium. (2014). A genome-wide association meta-analysis of preschool internalizing problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, *53*(6), 667-676.

Abstract

Objective: Preschool internalizing problems (INT) are highly heritable and moderately genetically stable from childhood into adulthood. Gene-finding studies are scarce. In this study, the influence of genome-wide measured Single Nucleotide Polymorphisms (SNPs) was investigated in three cohorts (total N = 4,596 children) in which INT was assessed with the same instrument (Child Behavior Checklist).
Method: First, genome-wide association (GWA) results were used for density estimation and GCTA analyses to calculate the variance explained by all SNPs. Next, a fixed effect inverse variance meta-analysis of the three GWA analyses was carried out. Finally, the overlap in results with prior GWA studies of childhood and adulthood psychiatric disorders and treatment responses was tested by examining whether SNPs associated with these traits jointly showed a significant signal for INT.
Results: Genome-wide SNPs explained 13% to 43% of the total variance. This indicates that the genetic architecture of INT mirrors the polygenic model underlying adult psychiatric traits. The meta-analysis did not yield a genome-wide significant signal, but was suggestive for the PCSK2 gene located on chromosome 20p12.1. SNPs associated with other psychiatric disorders appeared to be enriched for signals with INT (lambda=1.26, $p< 0.03$).
Conclusions: Our study provides evidence that INT is influenced by many common genetic variants, each with a very small effect, and that, even as early as age 3, genetic variants influencing INT overlap with variants that play a role in childhood and adulthood psychiatric disorders.

Introduction

Preschool internalizing problems (INT) are relatively prevalent, often not self-limited, and associated with significant morbidity. A recent study investigating prevalence rates of DSM-IV disorders in a sample of 2,475 Norwegian 4-year olds found, for example, that 1.5% of the children fulfilled the criteria for any anxiety disorder and 2% for a depressive disorder.[1] Preschool INT are persistent into childhood, as shown by several longitudinal studies.[2-5]

Twin studies have shown a substantial influence of genetic factors on preschool INT. Heritability estimates of INT, assessed across a range of instruments, are mostly around 40% and 50% (range: 36% - 75%) with study samples varying from 822 to 6,783 twin pairs.[5-9] These heritability estimates are similar to or even higher than the estimates found for anxiety and depressive symptoms and disorders in adults.[10, 11] Moreover, genetic factors influencing INT at age 3 continue to have an effect later in life, even into adulthood[5,12]

There are numerous gene-finding studies for anxiety and depression in adults, but gene finding studies on childhood INT are scarce. There has been only one genome-wide association (GWA) study that analyzed anxiety-related behaviors in 2,810 7 year olds.[12] None of the effects of the top ten Single Nucleotide Polymorphisms (SNPs) (p –values between $8.7x10^{-7}$ and $1.2X10^{-4}$) were replicated in an independent sample of 4,804 children. In addition, a Genome-wide Complex Trait Analysis (GCTA) was performed in the discovery sample.[14] Such an analysis does not focus on the effect of each SNP separately, but calculates the variance explained by all genome-wide SNPs. For anxiety-related behaviors, the GCTA yielded estimates were between 0.01 (SE=0.11) to 0.19 (SE=0.12). The authors concluded that common SNPs do not explain as much of the genetic influence on anxiety at age 7 as on other psychiatric phenotypes.

We present a genome-wide approach to investigate the etiology of preschool INT. Genome-wide SNP data were analyzed from three cohorts with a total N of 4,596 children in which INT was measured with the same instrument. Each cohort carried out a GWAS. These results were, first, used to estimate the variance attributable to all SNPs in each cohort. A meta-analysis of the results of the three GWAS was performed next, aiming to identify genetic variants influencing INT. Finally, overlap between our meta-analysis results and results from prior GWA (meta-)analysis studies was investigated. We analyzed whether SNPs associated with a range of psychiatric disorders jointly yielded a significant signal in the meta-

analysis results of INT at age 3. We have not restricted these analyses to SNPs associated with internalizing disorders (depression), but have also analyzed SNPs associated with disorders usually diagnosed in childhood or psychotic disorders . There is frequent co-morbidity between childhood internalizing disorders and other disorders at childhood and internalizing symptoms predict a range of disorders in adulthood, including disruptive disorders and schizophrenia.[1, 15-17] This could be well due to overlapping genetic risk factors. Further, it has been suggested that treatment resistant depression is influenced by specific risk factors including early age of onset,[17] which may signify that disorders resistant to various treatments bear a unique genetic signature. Although literature does not provide a direct link to internalizing problems in children and treatment response in adults, we wished to explore whether treatment resistant SNPs were enriched in preschool internalizing children. Therefore, SNPs were also selected from GWAS of treatment response in adults.

Methods

Subjects

Participants were recruited from 3 large population-based studies (see Table 1).

Generation R (www.generationr.nl). The Generation R study is a prospective population-based cohort of 9745 children born in Rotterdam, the Netherlands, whose due dates were between April 2002 and January 2006.[19,20] Data from a total of 7893 children were available and eligible for follow-up. DNA was extracted from cord blood taken at birth. Children of Northern European descent, as determined by principal component analyses of GWA data, were selected.[21] Of 5908 children with DNA available, 2,841 children of Northern European descent were identified of whom 2037 children (50**%** girls) had SNP data and the INT score available. The Medical Ethical Committee of the Erasmus Medical Centre, Rotterdam (MEC 217.595/2002/20) approved the study protocol, and participants gave informed consent in writing.
NTR (www.tweelinGenerationRegister.org). The Netherlands Twin Register (NTR) is a prospective study involving twin families. The NTR was established at the VU University Amsterdam in 1987 and includes twins born from 1986 and onwards.[22] Data collection is ongoing. Parental ratings of problem behavior are available for

ages 2, 3, 5, 7, 10 and 12 years. Subsamples of young twins were invited to participate in experimental and laboratory studies and to provide a DNA sample, either by whole blood or by buccal swabs.[23] Dizygotic twin pairs were included in the analyses while correcting for the dependence between their measures. From monozygotic twin pairs, one twin was randomly selected. This yielded a sample of 1475 children (50% girls), belonging to 1031 families, with genotype and phenotype data. The study was approved by the Central Ethics Committee on Research involving human subjects of the VU University Medical Centre, Amsterdam, and an institutional review board certified by the U.S. Office of Human Research Protections (IRB number IRB-2991 under Federal-wide Assurance-3703; IRB/institute codes, NTR 03-180).

Raine (www.rainestudy.org.au). The Western Australian Pregnancy (Raine) Cohort Study is a prospective cohort representative of those presenting to an antenatal tertiary referral center in Western Australia.[24, 25] There were 2900 pregnant women recruited between 1989 and 1991 as part of a randomized control trial to investigate the association of repeated ultrasound measurements during pregnancy on birth outcomes. DNA was collected from blood at the fourteen year follow-up. There are 1084 children (49% girls) for whom DNA and INT data at 2 to 3 years of age are available. Participant recruitment and all follow-ups of their families were approved by the Human Ethics Committee at King Edward Memorial Hospital and/or Princess Margaret Hospital for Children in Perth.

Table 1: Description of cohorts, Internalizing Problem scores and measure (upper part) and the estimates of the variance explained by all SNPs obtained with the density estimation method (DE) and with Genetic Complex Trait Analysis (GCTA) (lower part).

	Generation R	**NTR**	**Raine**
N (%girls)	2037 (49)	1475 (50)	1084 (49)
Mean age (sd)	3.0 (0.10)	3.31 (0.26)	2.2 (0.15)
Mean INT score (sd)	4.0 (3.4)	7.8 (6.0)	7.2 (5.1)
INT score range	0,23	0,33	0,37
CBCL[a] version	CBCL 1 ½ - 5	CBCL/2-3	CBCL/2-3
Website	www.generationr.nl	www.tweelingenregister.org/en/	www.rainestudy.org.au
Explained variance in %			
DE (p-value)	41 (0.04)	31 (0.41)	43 (0.58)
GCTA (p-value)	26 (0.07)	18 (0.30)	13 (0.33)

[a]CBCL: Child Behavior Checklist

Genotyping, Quality control and Imputation Procedures

DNA was extracted from blood in Raine and Generation R and from buccal swabs (n=1087) and whole blood (n=388) in NTR. Excellent concordance between SNP genotyped in buccal and blood samples has been shown in 331 individuals for whom both kind of samples were available.[6] For Raine and Generation R, genotyping was performed on Illumina platforms. For NTR, the affymetrix 6.0 platform was used. Table S1 provides details of the genotyping centers, platforms, calling software, analytic and imputation software, and details of the pre-imputation and post-imputation filtering criteria per study. Basic quality checks for each SNP included call rates and Hardy-Weinberg statistics. For Generation R and Raine, each sample was checked for excess heterozygosity, gender accuracy, relatedness (identity by descent) and missingness. NTR samples were also checked for incorrect Mendelian inheritance patterns, as parents of twins were genotyped in 25% of families. After prefiltering, phased genotype data were imputed to build 36 (release 22) of the original HapMap CEU reference panel for NTR and Raine studies, and the consensus panel for the Generation R study resulting in 2.5 to 3 million SNPs for GWA analysis.[27]

Measurement of Internalizing Problems

In all cohorts, INT was assessed by the internalizing problem scale of the Child Behavior Checklist (CBCL). In the most recent version of the CBCL 1½ - 5, the INT scale consists of 36 items. Thirty-four of these items were measured in all three cohorts and were, therefore, used to obtain the INT sum score for the analyses. Example items are “Acts too young for age”, “Worries a lot” and “Clings to adults or too dependent”. For each item, the rater must select a score of 0 (not true), 1 (somewhat or sometimes true) or 2 (very true or often true) resulting in a potential score range of 0 to 68. Table S2 provides a list of the 34 items and their corresponding subscales. In Generation R and Raine, the primary caregiver, usually the mother, filled out the questionnaire. These studies thus included paternal and other caregiver raters, although they were rare (5% in Generation R and data not available in Raine). In NTR, only maternal ratings were analyzed.

Statistical Analysis

GWA-based analysis within cohorts. Each cohort modeled the outcome as the square root of the sum score, which was chosen based on simulation and regression diagnostics (details are provided in the Supplementary Material). The significance of the SNPs was tested in the following regression equation:

$$\sqrt{INT} = \alpha_0 + \beta_1 SNP + \beta_2 PC + \beta_3 Sex$$

SNP was coded as 0, 1 and 2, reflecting the number of alternate alleles for a given individual. To correct for population stratification, ten principal components (PC) were included as covariates in the NTR GWA, and two were included in Raine and Generation R. The Generation R and NTR studies did not adjust for age, given the age restriction at time of data collection to 3 years; the Raine study accepted children of 2 to 3 years of age, and adjusted for age. In Generation R and Raine, the analyses were performed in mach2qtl software.[28,29] In NTR, the analyses were performed in PLINK[30] using the option "—family" to account for the dependence between INT measured in DZ twins from the NTR.

Variance explained by all SNPs. The joint effect of all SNPs in explaining the variance of INT was calculated in each cohort using the density estimation method (DE) proposed by So et al [31] and using GCTA. In GCTA, a genetic relationship matrix is calculated, based on all SNPs, reflecting the genetic similarity in *unrelated* individuals. Next, the variance explained by these SNPs is estimated in a linear mixed model in which the measures of the genetic similarity is included as a random effect to predict the phenotype.[14] The analyses were corrected for the covariates included in the GWA analyses. Moreover, related individuals were excluded (threshold = 0.025).

In contrast to GCTA, DE does not use measured SNPs, but uses the *z*-transformed t-statistics of the regression coefficients as obtained in a GWA analysis to estimate the explained variance. As it is more common to provide GWA results to a consortium than genotype data, this method is more suitable when using data from several cohorts. The basic idea is to compare the distribution of *z*-transformed t-statistics of the regression coefficients of genome-wide SNPs to the theoretical null distribution of *z* statistics representing no effects. Deviation of the observed statistics from the theoretical null distribution indicates that SNPs explain part of the variance. Specifically, the observed *z* statistics, which contain

error due to sampling fluctuation, are first corrected to obtain *z* statistics representing "true" effect sizes. The *z* statistics can then be combined by summing the contributions of all SNPs using sums of squares as in analysis of variance. These sums of squares are computed based on the estimated effect sizes and the study sample size, as well as the number of included covariates and their joint effect size. The resulting sum is an estimate of the total proportion of phenotypic variance explained by the SNPs in the analysis. Before applying DE, we carried out linkage disequilibrium (LD) pruning as suggested by So et al.[31] using HapMap CEU genotypes as a reference set which had been used for imputation in all three cohorts.
P-values for the variance explained estimates are calculated using Monte Carlo *p*-values. This is necessary because the sampling distribution for the estimates is skewed and biased, and standard errors are potentially misleading. Briefly, in each of 4000 Monte Carlo replications, variance explained is estimated using simulated *t* statistics for each SNP under the null hypothesis (i.e. zero variance explained). Then the estimated *p*-value for the observed data is the proportion of Monte Carlo replications with estimated variance explained greater than or equal to the estimate for the observed data. Simulation results showing that this method is appropriate for the variance explained estimates are available on request from the fourth author.

Meta-analysis. An inverse variance meta-analysis was performed in METAL.[32] Comparing a fixed effect analysis with a random effect analysis did not show different results. Therefore, we report the findings from the fixed effect model. Because the genomic control lambdas (the median X^2 association statistic divided by the median expected under the null) within each cohort were close to 1.0 and thus indicated no evidence for inflation, we applied a genomic control solely at the meta-analysis stage. We only considered SNPs whose minor allele frequency was greater than 0.01. We also applied quality filtering, requiring that the imputed SNPs had a quality score above 0.30 for the Generation R and Raine studies, and a PLINK info score between 0.80-1.1 for the NTR study. We considered any SNPs with a p-value below $5x10^{-8}$ to achieve genome-wide statistical significance.

Analysis of SNPs associated with psychiatric disorders and treatment response. We conducted a search in www.genome.gov/gwastudies in January 2013 for SNPs

with a p-value below $1x10^{-5}$ in GWAS of psychiatric disorders and treatment response performed in children or adults. We first analyzed SNPs associated with internalizing disorders (major depression). Next, in three steps, SNPS were added that were associated with disorders usually diagnosed in childhood (ADHD, conduct disorder and autism) and psychotic disorders (bipolar disorder and schizophrenia), that were associated with treatment response for antidepressants, lithium and antipsychotics and that were located in candidate genes for major depression based on hypotheses regarding the etiology, derived from Table 1 in Bosker et al.[33]

To determine whether these candidate SNPs were associated beyond expectation under the null in our meta-analysis, we calculated the genomic control lambda. Next, a null distribution of lambdas is created by sampling 10,000 sets of p-values equal in size to the set of SNPs being tested. The observed lambda in each of the 4 sets is significant when it exceeds the lambdas in the null distribution.

Results

Variance explained by all SNPs

In the Raine and NTR studies, the GWA analysis was performed on 2,543,887 autosomal SNPs imputed from the HapMap original panel that passed initial quality control measures (see Table S1). In Generation R, 3,021,329 autosomal SNPs imputed from the HapMap consensus panel passed initial quality control measures and were analyzed in the GWA. QQ and Manhattan plots for each cohort are provided in Figures S1 and S2. The genomic control lambda was very close to 1 (λ_{GC}=1.02), indicating there was no evidence for inflated test statistics. Next, DE analyses and GCTA were performed (see lower part of Table 1). After pruning for LD, 29,588, 29,612 and 42,800 SNPs were used in the DE analyses of NTR, Raine, and Generation R, respectively. The variance explained by these pruned SNPs varied between 31% and 43%. The variance explained by covariates in each analysis was negligible. In GCTA, the variance explained varied between 13% and 26%. In Generation R, the estimate based on DE was significant (p=0.04) and the GCTA estimate approached significance (p=0.07). The findings were not significant for NTR and Raine. Given a heritability estimate of 59% for INT at age 3,[6] this signifies that the SNPs capture between 22% and 72% (13%/59% and 43%/59% respectively) of the genetic variance.

Genome-wide association analyses, quality control and meta-analysis

After applying post-GWA control measures, there were 2,403,520 SNPs present in both the consensus and the original HapMap panels included in the meta-analysis. At a threshold of $P \leq 5x10^{-8}$, no genome-wide significant findings were detected for the meta-analyzed results, as shown in the Manhattan plot (Figure 1). Table 2 lists the top 30 SNP whose p-values were smaller than $1x10^{-5}$, which we considered to be potentially suggestive. Interesting signals that contained multiple suggestive SNPs in an independent region from chromosome 9q33.1 and chromosome 20p12.1 were apparent, and regional plots are shown in supplementary figures S3A and S3B. The 9q33.1 region is an intergenic region and is not in LD with the nearest gene upstream (DBC1 implicated in bladder cancer) or downstream (CDK5 genes implicated in rheumatoid arthritis). The current ENCODE annotations do list two interesting nearby non-coding RNAs (lincRNAs) in this region: RP11-360A18.2 and RP11-360A18.1 (Ensembl version ENSG00000261432.1 and ENSG00000225960.1,respectively), but the SNPs located near these coordinates are also not in strong LD with our region. The 20p12.1 signal, however, appeared to be in the PCSK2 gene, as indicated in the regional plots provided in supplementary figures S3A and S3B

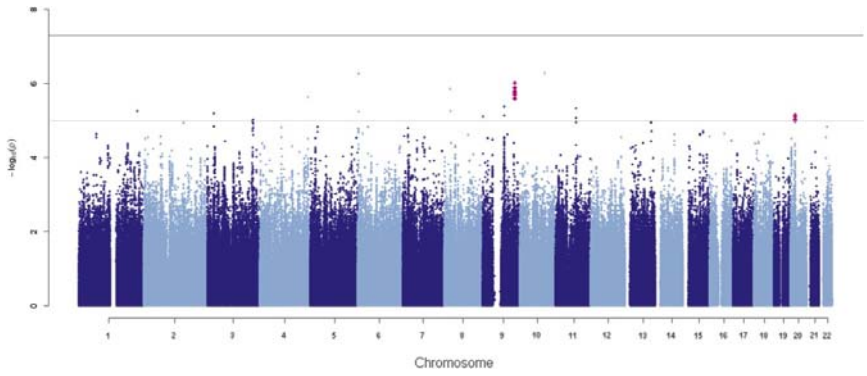


Figure 1. Manhattan Plot for the genome wide association meta-analysis of preschool Internalizing Problems across three cohorts**.** Chromosome is displayed on the x-axis and the association statistic, expressed as –log10(p-value), is given on the y-axis. Points that fall within the black (-log10($5x10^{-8}$) and gray lines (-log10($1x10^{-5}$) are signals suggestive for association. Pink diamonds denote the SNPs of interest on chromosome 9 and 20.

Table 2. The 30 most strongly associated genetic variants with preschool Internalizing Problems in the Genome-Wide Association meta-analysis. A direction is provided for each study in the following order: NTR, GenR, and Raine. A plus indicates that the beta for the association between the SNP and square root of the Internalizing Problem score is positive. A minus indicates a negative association. A question mark indicates that the SNP did not survive the post-imputation QC process for that study.

Chr Region	SNP	Position	MAF	Ref allele	Direction[a]	Beta (95%CI)	*P*-value
1q42.12	rs360059	224143755	0.3564	a	---	-0.11 (-0.15, -0.061)	5.5810^{-6}
1q42.12	rs1223245	224149048	0.3578	a	+++	0.11 (0.061, 0.15)	5.4410^{-6}
3q26.31	rs4894527	173424920	0.2921	t	+++	0.11 (0.060, 0.16)	9.8210^{-6}
3q26.31	rs1878007	174281494	0.1702	a	---	-0.12 (-0.18, -0.069)	9.3410^{-6}
4q35.1	rs10013166	183714323	0.4416	a	+++	0.11 (0.062, 0.15)	2.3610^{-6}
5q33.1	rs1062177	151164894	0.2475	t	-?-	-0.16 (-0.23, -0.091)	8.2610^{-6}
6p25.2	rs2272990	3022140	0.0823	t	?--	-0.39 (-0.55,-0.24)	5.3910^{-7}
6p25.2	rs9405191	3031952	0.153	c	?--	-0.29 (-0.42, -0.17)	5.7510^{-6}
6p25.2	rs9391981	3032005	0.1526	c	?--	-0.29 (-0.42, -0.17)	5.7510^{-6}
8p21.3	rs6557600	22819193	0.2273	a	---	-0.13 (-0.18, -0.076)	1.3710^{-6}
8p21.3	rs310272	23699434	0.3718	c	+++	0.10 (0.060, 0.15)	5.5210^{-6}
9p24.3	rs12000567	779499	0.0165	a	?++	0.63 (0.35, 0.91)	7.7510^{-6}
9q21.31	rs17266958	82455069	0.0756	t	---	-0.20 (-0.28, -0.11)	4.2010^{-6}
9q21.31	rs17083743	82481285	0.071	a	+++	0.20 (0.11, 0.29)	7.2110^{-6}
9q33.1	rs10818415	121811389	0.1769	t	---	-0.14 (-0.12, -0.081)	1.7810^{-6}
9q33.1	rs10984795	121814167	0.1853	a	+++	0.14 (0.082, 0.20)	1.6110^{-6}
9q33.1	rs16909317	121814340	0.1849	t	---	-0.14 (-0.19, -0.081)	2.0010^{-6}
9q33.1	rs2416740	121817157	0.1842	a	---	-0.14 (-0.19, -0.079)	2.5610^{-6}
9q33.1	rs2416741	121817280	0.1835	a	---	-0.14 (-0.20,-0.081)	1.7510^{-6}
9q33.1	rs10984803	121820192	0.1749	a	---	-0.14 (-0.20, -0.084)	1.2410^{-6}
9q33.1	rs10818418	121820770	0.1841	a	+++	0.14 (0.083, 0.20)	1.3410^{-6}
9q33.1	rs2416745	121821075	0.1766	t	+++	0.14 (0.084, 0.20)	9.5110^{-7}
9q33.1	rs10984807	121822291	0.188	t	+++	0.14 (0.082, 0.20)	2.5310^{-6}
10q23.33	rs640090	95382575	0.0715	c	?++	0.55 (0.33, 0.76)	5.2410^{-7}
11q14.1	rs12270115	79162785	0.1334	a	+++	0.15(0.088, 0.22)	4.6610^{-6}
11q14.1	rs12287037	79205014	0.1918	t	+++	0.12 (0.068, 0.17)	8.4610^{-6}
20p12.1	rs2281204	17364812	0.3416	a	+++	0.10 (0.056, 0.14)	9.7710^{-6}

20p12.1	rs890609	17365013	0.3411	a	---	-0.10 (-0.14, -0.056)	9.9410^{-6}
20p12.1	rs2021786	17369978	0.3356	t	---	-0.10 (-0.15, -0.058)	6.9510^{-6}
20p12.1	rs2021785	17370063	0.3356	t	---	-0.10 (-0.15, -0.058)	6.9810^{-6}
20p12.1	rs13039651	17371040	0.3359	t	---	-0.10 (-0.15, -0.057)	7.6710^{-6}
20p12.1	rs2269020	17375229	0.3442	c	---	-0.10 (-0.15, -0.056)	9.1610^{-6}

Analysis of SNPs associated with psychiatric disorders and treatment response

Table 3 shows the results of the analyses of the joint effect of the four sets of SNPs with a p-value below $1x10^{-5}$ in GWAS of psychiatric disorders and treatment response as found in www.genome.gov/gwastudies. The GWAS of internalizing disorders (depression), psychotic disorders (schizophrenia and bipolar disorder) and treatment response were performed in adults. For the GWAS of the disorders usually diagnosed in childhood both adult and children's samples were used. The SNPs associated with only internalizing disorders did not show a lambda significantly greater than 1.0. Adding SNPs associated with disorders usually diagnosed at childhood and with psychotic disorders yielded a significant lambda of 1.26. The addition of SNPs associated with treatment response for antidepressants, lithium and antipsychotics led to a worsening of the signal. However, adding SNPs in candidate genes for major depressive disorder (see Table 1 provided in Bosker et al.[33]) made the lambda significant again: 1.20 . Figure 2 shows the QQ plot of the 320 SNPs with the largest lambda of 1.26. (A full list of the SNPs is available on request from the first author).

None of the above investigated SNPs were localized in the 20p12.1 or the 9q33.1 regions that we highlighted from the results in our GWA meta-analysis. The minimum p-value from these SNPs was 0.002, and did not meet criteria for a suggestive finding.

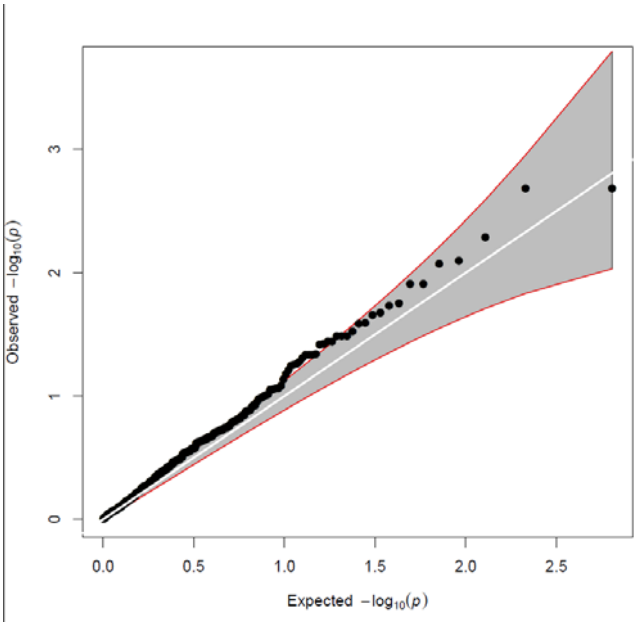


Figure 2. Quantile-Quantile plot for the joint effect on preschool Internalizing Problems of SNPs with p-values < 1x10-5 in prior genome-wide association meta-analyses of childhood and adulthood psychiatric disorders and treatment responses. The gray shaded area represents the 95% confidence intervals, and the white line represents equality between observed and expected – log10(p-values). The lambda value of 1.26 is significant (p = 0.03).

Table 3. Lambda's of the analysis of the joint effect on preschool Internalizing Problems of genetic variants with p-values of $1x10^{-5}$ in prior genome-wide association studies for internalizing disorders (depression), psychotic disorders (bipolar disorder and schizophrenia), disorders usually diagnosed in childhood (ADHD, autism and conduct disorder), treatment response, and genetic variants located in candidate genes for major depressive disorder.

Selected SNPs	No. SNPs in database	No. SNPs in filtered meta-analysis results	Lambda	Bootstrapped p-value
1) Internalizing disorders	61	59	1.18	0.29
2) Internalizing disorder, psychotic disorders and disorders usually diagnosed in childhood	368	320	1.26	0.03
3) Internalizing disorders, psychotic disorders, disorders usually diagnosed in childhood and treatment response on antidepressants, lithium and antipsychotics	472	412	1.17	0.08
4) 3 + SNPs in candidate genes for major depressive disorder	594	472	1.20	0.04

Discussion

Genome wide SNP data were used to investigate genetic factors influencing preschool internalizing symptoms (INT). First, the analysis of variance explained by all SNPs indicated that the common variation as measured with SNPs explain at least 22% of the genetic variance. Second, the meta-analysis of the results of the GWA analyses in the three cohorts did not yield a genome-wide significant effect, but two areas showed suggestive findings. Third, SNPs that were suggestively associated with childhood and adult psychiatric disorders in prior GWAS or candidate gene studies for MDD, showed a joint significant effect on INT at age 3 although none of these SNPs individually reached the level of suggestive significance ($p < 1 *10^{-5}$).

We note regarding our first result that the estimates vary quite widely (between 13% and 43% of the total phenotypic variance is explained) and that for two of the three cohorts, the estimates of the variance explained by all SNPs were not significant. The differences in estimates seem mostly due to the method; GCTA estimates were consistently lower than DE estimates. The lack of significance in Raine and NTR appears to be a matter of sample size, since, within one method, the estimates are rather similar for the three cohorts. Focusing on the (nearly) significant results for Generation R suggests that between 44% and 69% of the genetic variance is explained by common SNPs. Together with the significant joint effect of SNPs associated with other psychiatric phenotypes, these outcomes point to the conclusion that INT at age 3 is influenced by a large number of genetic variants, each with a small effect. Moreover, even INT measured as early as age 3 years genetically overlaps with adult psychiatric phenotypes.

These results add to the picture as painted by Visscher et al (2012)[34] reviewing five years of GWA studies. Complex phenotypes, psychiatric or somatic, seem to be highly polygenic and genetic variants can influence multiple traits, i.e. there is pleiotropy. Specifically focusing on psychiatric phenotypes, a polygenic architecture has been suggested for adult anxiety disorders, major depression, schizophrenia and bipolar disorder, with GCTA estimates of around one third to 50% of the genetic variance explained by common SNPs.[35-39] Pleiotropy has been detected for ADHD, autism and schizophrenia, bipolar disorder and major depression in studies using genome wide SNP or copy number variant data.[40, 41] Twin studies have also shown substantial genetic overlap within internalizing

symptoms or disorders[42, 43] and between internalizing disorders and mania and schizophrenia,[44, 45] in addition to stable genetic influences from childhood into adulthood for anxious depression and attention problems[12,46]

Although the sample had 80% power to detect an effect explaining ~ 0.90% of the variance and the INT measure was similar in the three cohorts and assessed at the same age, in the meta-analysis, there were no SNPs with a genome wide significant effect. Given the DE and GCTA results, this indicates that sample size was still insufficient for the, apparently, very small effect sizes. This is in agreement with the results for a continuous phenotype such as height in which a GWA analysis in over 180,000 individuals detected hundreds of genetic variants explaining jointly 10% of the phenotypic variance.[47]

In two regions, there were SNPs with a p-value below $1x10^{-5}$ in the meta-analysis. One is located in an intergenic region at chromosome 9. According to the latest results from ENCODE, this is also not a regulatory region of the genome.[48-50] The other region was at chromosome 20 and included SNPs of the PCSK2 gene. PCSK2 is an important protein in the processing of proinsulin to insulin[51] and PCSK2 variants have been correlated with insulin resistance,[52, 53] myocardial infarction[54] and age at menarche.[54] The link between depression and cardiovascular disease has long been recognized.

We compared our results with the results from the GWA analysis of anxiety-related behaviors in 7-year olds.[14] Similar to their own replication effort, we did not find significant effects for their reported top SNPs. The lowest p-value in the current study was 0.09 for SNP rs2772129, and the effect was in the same direction as the discovery sample. All other p-values were above 0.33. The estimates for the variance explained by all SNPs using GCTA were somewhat higher in our samples than in the sample analyzed by Trzaskowski et al,[13] in which the GCTA estimates varied between 1% and 19%. They analyzed four anxiety dimensions (negative affect, negative cognition, fear and social anxiety) and a general anxiety composite score consisting of the sum of the standardized scores of the four dimensions. The second highest GCTA estimate (0.16) was for the composite score. This could indicate that common SNPs explain more variance in a broader defined phenotype, as a narrower phenotype might be influenced by less SNPs.

To conclude, this study shows that a phenotype such as INT at age 3 is genetically similar to adult phenotypes, i.e. it is a polygenic disorder, influenced by a large number of SNPs each with a small effect. This signifies that with large enough samples it is possible to detect genetic variants influencing preschool INT. This is even more important given the overlap in results with GWA studies for other psychiatric disorders indicating that these genetic variants also increase the risk for later psychiatric disorders.

References

1. Wichstrom L., Berg-Nielsen T.S., Angold A., Egger H.L, Solheim E, Sveen T.H. (2012) Prevalence of psychiatric disorders in preschoolers. Journal of Child Psychology and Psychiatry *53*, 695-705.

2. Mesman J., Koot H.M. (2001) Early preschool predictors of preadolescent internalizing and externalizing DSM-IV diagnoses. Journal of the American Acadamy of Child and Adolescent Psychiatry *40*, 1029-1036.

3. Beyer T., Postert C., Muller J.M., Furniss T. (2012) Prognosis and continuity of child mental health problems from preschool to primary school: Results of a four-year longitudinal study. Child Psychiatry and Human Development *43*, 533-543.

4. Pihlakoski L., Sourander A., Aromaa M., Rautava P., Helenius H., Sillanpaa M. (2006) The continuity of psychopathology from early childhood to preadolescence: A prospective cohort study of 3-12-year-old children. European Child and Adolescent Psychiatry *15*, 409-417.

5. Bartels M., van den Oord E.J., Hudziak J.J., Rietveld M.J., van Beijsterveldt C.E., Boomsma D.I.(2004) Genetic and environmental mechanisms underlying stability and change in problem behaviors at ages 3, 7, 10, and 12. Developmental Psychology 40, 852-867.

6. Eley T.C., Bolton D., O'Connor T.G., Perrin S., Smith P., Plomin R.(2003) A twin study of anxiety-related behaviours in pre-school children. Journal of Child Psychology and Psychiatry *44*, 945-960.

7. Saudino K.J., Carter A.S., Purper-Ouakil D., Gorwood P. (2008) The etiology of behavioral problems and competencies in very young twins. Journal of Abnormal Psychology *117*, 48-62.

8. Van Hulle C.A., Lemery-Chalfant K., Goldsmith H.H. (2007) Genetic and environmental influences on socio-emotional behavior in toddlers: An

initial twin study of the infant-toddler social and emotional assessment. Journal of Child Psychology and Psychiatry *48,* 1014-1024.

9. Gregory A.M., Eley T.C., Plomin R. (2004) Exploring the association between anxiety and conduct problems in a large sample of twins aged 2-4. Journal of Abnormal Child Psychology *32*, 111-122.

10. Hettema J.M., Neale M.C., Kendler K.S. (2001) A review and meta-analysis of the genetic epidemiology of anxiety disorders. American Journal of Psychiatry *158,* 1568-1578.

11. Sullivan P.F., Neale M.C., Kendler K.S. (2000) Genetic epidemiology of major depression: Review and meta-analysis. American Journal of Psychiatry *157,* 1552-1562.

12. Nivard M.G., Dolan C.V., Kendler K.S., Kan K.J., Willemsen G., van Beijsterveldt C.E., Lindauer R.J.L., van Beek J.H.D.A., Geels L.M., Bartels M., Middeldorp C.M, & Boomsma D. I. (2014). Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. *Psychological medicine*, ePub ahead of print.

13. Trzaskowski M., Eley T.C., Davis O.S., Doherty S.J., Hanscombe K.B., Meaburn E.L., et al. (2013) First genome-wide association study on anxiety-related behaviours in childhood. PLoS One *8*, e58676.

14. Yang J., Benyamin B., McEvoy B.P., Gordon S., Henders A.K., Nyholt D.R., et al. (2010) Common SNPs explain a large proportion of the heritability for human height. Nature Genetics *42,* 565-569.

15. Reef J., van Meurs I., Verhulst F.C., van der Ende J. (2010) Children's problems predict adults' DSM-IV disorders across 24 years. Journal of the American Academy of Child and Adolecent Psychiatry *49*, 1117-1124.

16. Cannon M., Caspi A., Moffitt T.E., Harrington H., Taylor A., Murray R.M., et al. (2002) Evidence for early-childhood, pan-developmental

impairment specific to schizophreniform disorder: Results from a longitudinal birth cohort. Archives of General Psychiatry *59*, 449-56.

17. Kim-Cohen J., Caspi A., Moffitt T.E., Harrington H., Milne B.J., Poulton R. (2003) Prior juvenile diagnoses in adults with mental disorder: Developmental follow-back of a prospective-longitudinal cohort. Archives of General Psychiatry. 709-717.

18. Souery D., Oswald P., Massat I., Bailer U., Bollen J., Demyttenaere K., et al. (2007) Clinical factors associated with treatment resistance in major depressive disorder: Results from a european multicenter study.Journal of Clinical Psychiatry *68*, 1062-1070.

19. Jaddoe V.W., Mackenbach J.P., Moll H.A., Steegers E.A., Tiemeier H., Verhulst F.C., et al. (2006) The generation R study: Design and cohort profile. European Journal of Epidemiology 475-484.

20. Jaddoe V.W., van Duijn C.M., Franco O.H., van der Heijden A.J., van Iizendoorn M.H., de Jongste J.C., et al. (2012) The generation R study: Design and cohort update 2012. European Journal of Epidemiology *27,* 739-756.

21. Jaddoe V.W., Bakker R., van Duijn C.M., van der Heijden A.J., Lindemans J., Mackenbach J.P., et al. (2007) The generation R study biobank: A resource for epidemiological studies in children and their parents. European Journal of Epidemiology *22,* 917-923.

22. Boomsma D.I., de Geus E.J., Vink J.M., Stubbe J.H., Distel M.A., Hottenga J.J., et al. (2006) Netherlands twin register: From twins to twin families.Twin Research and Human Genetics *9*, 849-857.

23. van Beijsterveldt C.E., Groen-Blokhuis M., Hottenga J.J., Franic S., Hudziak J.J., Lamb D., et al. (2102) The young netherlands twin register (YNTR): Longitudinal twin and family studies in over 70,000 children. Twin Research and Human Genetics 1-16.

24. McKnight C.M., Newnham J.P., Stanley F.J., Mountain J.A., Landau L.I., Beilin L.J., et al. (2012) Birth of a cohort - the first 20 2012s of the raine study. Mediical Journal of Australia *197*, 608-610.

25. Newnham J.P., Evans S.F,. Michael C.A., Stanley FJ, Landau LI. (1993) Effects of frequent ultrasound during pregnancy: A randomised controlled trial. Lancet. *342,* 887-891.

26. Scheet P., Ehli E.A., Xiao X., van Beijsterveldt C.E., Abdellaoui A., Althoff R.R., et al. (2012) Twins, tissue, and time: An assessment of SNPs and CNVs. Twein Research and Human Genetics. *15,* 737-745.

27. International HapMap Consortium. (2005) A haplotype map of the human genome. Nature *437*, 1299-1320.

28. Li Y., Willer C.J., Ding J., Scheet P., Abecasis G.R. (2010) MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genetic Epidemiology *34,* 816-834.

29. Li Y., Willer C., Sanna S., Abecasis G. (2009) Genotype imputation. Annual review of Genomic and Human genetics *10,* 387-406.

30. Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., Bender D., et al. (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics *81*, 559-575.

31. So H.C., Li M., Sham P.C. (2011) Uncovering the total heritability explained by all true susceptibility variants in a genome-wide association study. Genetic Epidemiology *35*, 447-56.

32. Willer C.J., Li Y., Abecasis G.R. (2010) METAL: Fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190-2191.

33. Bosker F.J., Hartman C.A., Nolte I.M., Prins B.P., Terpstra P., Posthuma D., et al. (2011) Poor replication of candidate genes for major depressive

disorder using genome-wide association data. Molecular Psychiatry *16,* 516-532.

34. Visscher P.M., Brown M.A., McCarthy M.I., Yang J. (2012) Five years of GWAS discovery. American Journal of Humand Genetics *90*, 7-24.

35. Lee S.H., Wray N.R., Goddard M.E., Visscher P.M. (2011) Estimating missing heritability for disease from genome-wide association studies. American Journal of Human Genetics *88,* 294-305.

36. Lee S.H., DeCandia T.R., Ripke S., Yang J., Schizophrenia Psychiatric Genome-Wide Association Study Consortium (PGC-SCZ), International Schizophrenia Consortium (ISC), et al. (2012) Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. Nature Genetics *44,* 247-50.

37. Demirkan A., Penninx B.W., Hek K., Wray N.R., Amin N., Aulchenko Y.S., et al. (2011) Genetic risk profiles for depression and anxiety in adult and elderly cohorts. Molecular Psychiatry *16,* 773-83.

38. Lubke G.H., Hottenga J.J., Walters R., Laurin C., de Geus E.J., Willemsen G., et al. (2012) Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. Biological Psychiatry *72,* 707-709.

39. International Schizophrenia Consortium, Purcell S.M., Wray N.R., Stone J.L., Visscher P.M., O'Donovan M.C., et al. (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature *460,* 748-752.

40. Cross-Disorder Group of the Psychiatric Genomics Consortium, Smoller J.W., Craddock N., Kendler K., Lee P.H., Neale B.M., et al. (2013) Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis. Lancet *381,* 1371-1379.

41. Williams N.M., Zaharieva I., Martin A., Langley K., Mantripragada K., Fossdal R., et al. (2010) Rare chromosomal deletions and duplications

in attention-deficit hyperactivity disorder: A genome-wide analysis. Lancet *376,* 1401-1408.

42. Hettema J., Neale M.C., Myers J.M., Prescott C.A., Kendler K.S.. (2006) A population-based twin study of the relationship between neuroticism and internalizing disorders. American Journal of Psychiatry *163,* 857-864.

43. Franic S., Dolan C.V., Borsboom D., Hudziak J.J., Van Beijsterveldt C.E.M., Boomsma D.I. (2013) Can genetics help psychometrics? improving dimensionality assessment through genetic factor modeling. Psychological Methods *18,* 406-433.

44. Argyropoulos S.V., Landau S., Kalidindi S., Toulopoulou T., Castle D.J., Murray R.M., et al. (2008) Twins discordant for schizophrenia: Psychopathology of the non-schizophrenic co-twins. Acta Psychiatrica Scandinavica. *118,* 214-219.

45. McGuffin P., Rijsdijk F., Andrew M., Sham P., Katz R., Cardno A. (2013) The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. Archives of General Psychiatry *60,* 497-502.

46. Kan K.J., Dolan C.V., Nivard M.G., Middeldorp C.M., Van Beijsterveldt C.E.M., Willemsen G., et al. (2013) Genetic and environmental stability in attention problems across the lifespan: Evidence from the netherlands twin register. Journal of the American Academy of Child and Adolescent Psychiatry *52,* 12-25.

46. Lango Allen H., Estrada K., Lettre G., et al.(2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature *467,* 832-838.

47. ENCODE Project Consortium. (2004) The ENCODE (ENCyclopedia of DNA elements) project. Science *306,* 636-640.

48. ENCODE Project Consortium, Birney E., Stamatoyannopoulos J.A., Dutta A., Guigo R., Gingeras T.R., et al. (2007) Identification and analysis of

functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447,* 799-816.

49. ENCODE Project Consortium, Dunham I., Kundaje A., Aldred S.F., Collins P.J., Davis C.A., et al. (2012) An integrated encyclopedia of DNA elements in the human genome. Nature *489,* 57-74.

50. Malide D, Seidah N.G., Chretien M., Bendayan M. (1995) Electron microscopic immunocytochemical evidence for the involvement of the convertases PC1 and PC2 in the processing of proinsulin in pancreatic beta-cells.Journal of Histochemistry and Cytochemistry *43,* 11-19.

51. Jonsson A., Isomaa B., Tuomi T., Eriksson J.G., Groop L., Lyssenko V.(2012) Effect of a common variant of the PCSK2 gene on reduced insulin secretion. Diabetologia *55,* 3245-3251.

52. Leak T.S., Keene K.L., Langefeld C.D., Gallagher C.J., Mychaleckyj J.C., Freedman BI, et al. (2007) Association of the proprotein convertase subtilisin/kexin-type 2 (PCSK2) gene with type 2 diabetes in an african american population. Molecular Genetic and Metabolism *92,* 145-150.

53. Fujimaki T., Kato K., Yokoi K., Oguri M., Yoshida T., Watanabe S., et al. (2010) Association of genetic variants in SEMA3F, CLEC16A, LAMA3, and PCSK2 with myocardial infarction in japanese individuals. Atherosclerosis *210,* 468-473.

54. Elks C.E., Perry J.R., Sulem P., Chasman D.I., Franceschini N., He C., et al. (2010) Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. Nature Genetics *42,* 1077-1085.

Supplementary Material

The following presents additional methods, tables and figures that are designed to provide greater detail into the findings from our genome-wide meta-analysis of internalizing problems in children.

METHODS

Each of the three cohorts participating in this genome-wide association meta-analysis transformed the internalizing problem (INT) score by taking the square root. This was based on model diagnostics and results from data simulation. The INT score was right skewed so that a large density of scores was observed at zero to low values, while an increasingly smaller number of scores was observed at the more extreme values. We were thus interested in determining the best way to handle this non-normal distribution.

We first simulated data to mimic the INT score, and then compared type I error, power and bias for several modeling approaches. A score of 20 items with responses of 0, 1 or 2 and probability of 0.925, 0.05 and 0.025, respectively, were simulated; the items were correlated and summed to create a simulated score. The SNP frequency was set to 0.20 and the association of the summed score with the SNP was recorded. Analyses were repeated 1000 times for different combinations of SNP effect size (and thus percent variation explained) and sample size. Four different models were tested: 1) a generalized linear model (GLM) specifying a gamma distribution, 2) a GLM specifying a Gaussian distribution, 3) a GLM specifying a Gaussian distribution on the log transformation of the score, and 4) a GLM specifying a Gaussian distribution on the square root transformation of the score. Simulation results for the 20 items score that was distributed similarly to our INT trait are provided in Supplementary Table 3. Across a range of effect sizes and sample sizes, the square root transformation of the score generally resulted in the highest power.

Additionally, model diagnostics were carried out in the Raine cohort among 1737 participants with complete phenotype data, regardless of whether GWA data was available or not, and plots of the residuals by fitted values were examined, as well as normal plots of the residuals. Model diagnostics suggested a considerably improved fit by transforming the INT score, particularly for the normalized plots of the residuals (data not shown). Between the natural log, log based 10 and square root transformations, the square root was judged to be the best fit in the Raine data. Along with the results from the simulation above, these findings justified our decision to perform the GWA using a square root transformation of the INT score.

To perform the phasing and imputation steps, Generation R and Raine studies used MaCH[1, 2] and NTR used Beagle.[3] All positions reflect build 36 locations.

REFERENCES

1. Li Y., Willer C., Sanna S., Abecasis G. (2009) Genotype imputation. Annual Review Genomics Human Genetics *10,* 387-406.

2. Li Y., Willer C.J., Ding J., Scheet P., Abecasis G.R. (2010) MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genetic Epidemiology *34,* 816-834.

3. Browning B.L., Browning S.R. (2009) A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. American Journal of Human Genetics *84,* 210-223.

Table S1. Genotyping and Quality Control in each Cohort

	Generation R	**NTR**	**Raine**
Genotyping Center	Genetic Laboratory-Dept Internal Medicine - Erasmus MC The Netherlands	Avera Institute for Human Genetics, Sioux Falls, South Dakota, USA	Centre for Applied Genomics, University of Toronto, Toronto, Canada
Genotyping Platform	Illumina 610K Quad	Affymetrix 6.0	Illumina 660 Quad
Calling Algorithm	Genomestudio 2009 V.1.1.9		Beadstudio
Imputation Software	MACHqtl	BEAGLE	MACHqtl
Association Software	MACH	PLINK	MACH
Reference panel	HapMap CEU r22 b36	HapMap CEU r22 b36	HapMap CEU r22 b36
Population Stratification Adjustments	2 principal components used as covariates	10 principal components used as covariates	2 principal components used as covariates
Filtering Criteria:			
Pre-Imputation			
SNP level			
Callrate	95%	95%	95%
HWE	<1x10-6	<0.00001	<5.7x10-7
Sample level			
% missing	>2.5%	>1%	>3%
IBD	π>0.1875	NA	>3sd from HapMap CEU sample
heterozygosity %	PLINK h<0.30	0.10< F > -0.10	< 4 sd from mean
Mendelian inhertance check	na	yes	na
Gender check	yes	no	yes
Post-Imputation			
Quality metric (cutoff)	r2hat	info	r2hat
GC lambda	1.030319	1.028017	1.004965
MAF	0.01	0.01	0.01

Table S2. Items of the Child Behavior Check List Internalizing Problem Behavior Scale

Item	Subscale	Item	Subscale
Avoids looking others in the eye	Withdrawn	Self-conscious or easily embarrassed	Anxious/Depressed
Acts too young for age	Withdrawn	Too fearful or anxious	Anxious/Depressed
Doesn't answer when people talk to him/her	Withdrawn	Looks unhappy without good reason	Anxious/Depressed
Seems unresponsive to affection	Withdrawn	Unhappy, sad or depressed	Anxious/Depressed
Shows little affection toward people	Withdrawn	Feelings hurt	Anxious/Depressed
Shows little interest in things around him/her	Withdrawn	Nervous	Anxious/Depressed
Refuses active games	Withdrawn	Doesn't eat well	Somatic Complaints
Withdrawn	Withdrawn	Aches	Somatic Complaints
Disturbed by any change in routine	Emotionally reactive	Can't stand things out of order	Somatic Complaints
Upset by new people or situations	Emotionally reactive	Constipated	Somatic Complaints
Sulks	Emotionally reactive	Diarrhea	Somatic Complaints
Worrying	Emotionally reactive	Headaches	Somatic Complaints
Twitches	Emotionally reactive	Nausea	Somatic Complaints
Moody	Emotionally reactive	painful bowel movements	Somatic Complaints
Whining	Emotionally reactive	Stomach aches	Somatic Complaints
Clings to adults or too dependent	Anxious/Depressed	too concerned with neat/clean	Somatic Complaints
Gets too upset when separated from parents	Anxious/Depressed	Vomits	Somatic Complaints

Table S3. Power for Simulated Trait

N	Effect	No Transform, Gaussian	No Transform, Gamma	Log Transform, Gaussian	Square Root Tranform, Gaussian
1000	0.001	0.092	0.093	0.104	0.118
1000	0.0025	0.186	0.187	0.181	0.217
1000	0.01	0.618	0.617	0.505	0.642
1000	0.05	0.999	0.999	0.976	0.999
1500	0.001	0.137	0.137	0.143	0.176
1500	0.0025	0.262	0.262	0.239	0.306
1500	0.01	0.793	0.791	0.662	0.812
1500	0.05	1	1	0.999	1
2500	0.001	0.182	0.182	0.166	0.216
2500	0.0025	0.407	0.407	0.348	0.452
2500	0.01	0.933	0.933	0.877	0.971
2500	0.05	1	1	1	1
5000	0.001	0.36	0.36	0.304	0.371
5000	0.0025	0.716	0.715	0.566	0.712
5000	0.01	0.998	0.998	0.991	1
5000	0.05	1	1	1	1

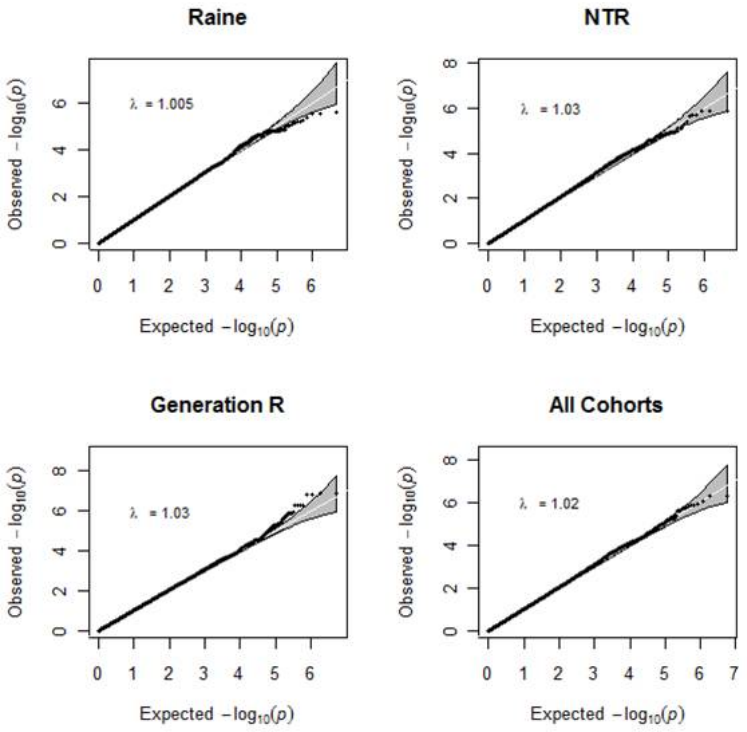


Figure S1. Quantile-Quantile plots for the association of genetic variants with Internalizing Problems. Plots are shown for each cohort separately, and for the results from the meta-analysis of all cohorts combined. Gray shaded areas represent the 95% confidence intervals and the white lines represent equality between observed and expected -log10(pvalues).

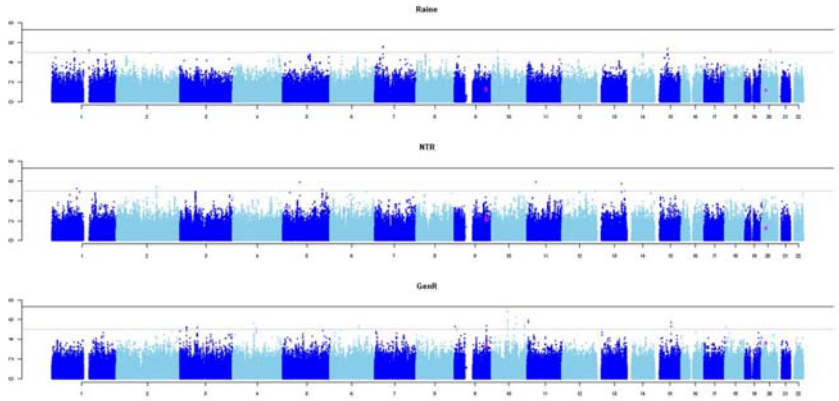


Figure S2. Manhattan plots for the genome-wide association analyses with Internalizing Problems. Plots are shown for each cohort separately. Chromosome is displayed on the x axis and the association statistic, expressed as –log10(pvalue), is given on the y axis. Points that fall within the black (-log10($5x10^{-8}$) and gray lines (log10($1x10^{-5}$) are suggestive, but do not reach genome-wide significance. Pink points denote the chr9 and chr20 SNPs that were of interest after meta-analysis.

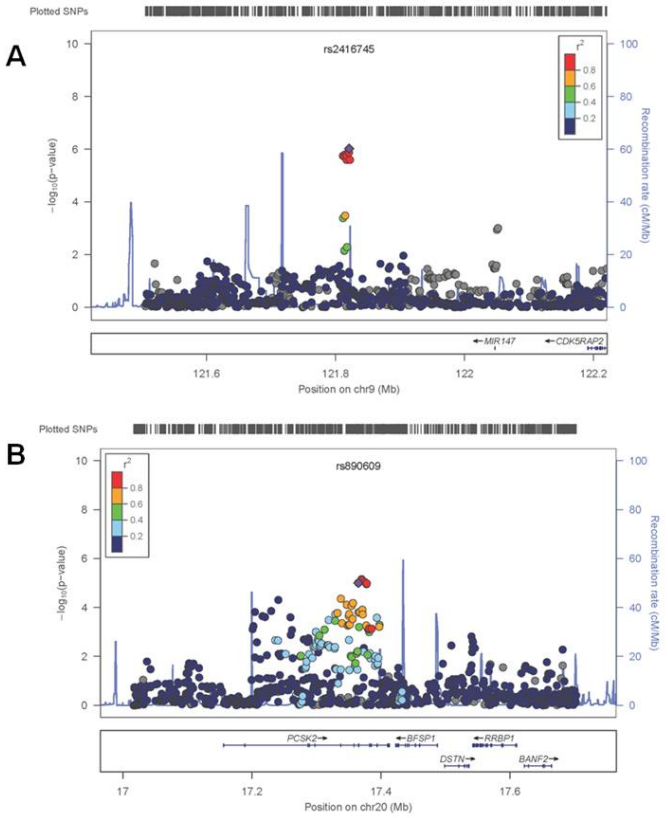


Figure S3A-B. Regional plots for the two regions that show suggestive signals. *P*-values reflect the meta-analysis results across the Netherlands Twin Registry (NTR), Generation R and Raine cohorts. Panel A shows the findings from chr9q33.1, an intergenic region that is not correlated with nearby genes. Panel B shows the findings from chr20p12.1, which is in the PCSK2 gene. Plots were made using LocusZoom (https://statgen.sph.umich.edu/loc)

Chapter 5: The genetic overlap between schizophrenia and childhood psychopathology.

Abstract

Objective:

The onset of schizophrenia is often preceded by a broad range of psychiatric disorders and symptoms, not only during adolescence but also during childhood. We investigated whether this association is explained by pleiotropy, i.e., genetic effects that influence both schizophrenia and childhood and adolescent psychiatric symptoms.

Methods

Polygenic risk scores (PRS), reflecting an individual's genetic risk for schizophrenia, were constructed for 1953 children from the Netherlands Twin Register (NTR) and 5665 children from the Avon Longitudinal Study of Parents And Children (ALSPAC). The association between the polygenic risk scores and DSM-IV based measures of anxiety, depression, Attention Deficit Hyperactivity Disorder (ADHD), Oppositional Defiant Disorder/Conduct Disorder (ODD/CD) was analyzed at age 7, 10, 12 and 15 years. The results were meta- analyzed across cohorts.

Results

The results revealed an FDR-corrected significant association between the PRS and anxiety at age 10 and nominal significant associations for anxiety at age 7and and depression at age 7 and 10. A post hoc analysis revealed stronger associations between the PRS and internalizing disorders than between the PRS and externalizing disorders.

Conclusion

In line with the earlier reported significant association between adult major depression and schizophrenia, these results suggest a common genetic etiology for schizophrenia and internalizing disorders. In contrast, genetic factors do not explain the association between externalizing disorders at childhood and the onset of schizophrenia later in life.

Introduction

The onset of schizophrenia generally occurs during adolescence or early adulthood[1], but it is well established that non-psychotic psychiatric symptoms can already be present in the period before the first psychotic episode. The prodromal phase is characterized by neurodevelopmental deficits,[2-4] cognitive learning and memory problems[5], and elevated psychiatric symptoms.[6] And even well before the prodromal phase, psychiatric symptoms or disorders are more prevalent in individuals who later develop schizophrenia. This has been demonstrated in population based cohorts that have been followed longitudinally,[7; 8] in retrospective assesment of schizophrenia cases,[9] as well as in people at risk for developing schizophrenia.[10] Both externalizing symptoms or disorders, including attention deficit hyperactivity disorder, conduct disorder, aggression, and anti-social behavior,[11; 12] and internalizing symptoms or disorders, including anxiety and depression, are associated with a higher risk of schizophrenia.[7; 8; 11; 13-15] These diverse findings indicate that an increase in non-specific (rather than specific) psychiatric symptoms precedes the onset of schizophrenia.
The early detection of schizophrenia can improve outcomes[8], and preventive treatment for individuals at risk for schizophrenia can reduce the risk of psychosis.[16; 17] Insight into the risk factors associated with the predictors of schizophrenia may facilitate early detection. Since schizophrenia is highly heritable with genetic factors explaining around 80% of the variance[18-20], we focused on the role of genetic risk factors, and investigate whether genetic factors that increase the risk of schizophrenia are also associated with psychiatric symptoms in childhood and adolescence.

A potential source of pleiotropy, i.e., a shared genetic etiology between childhood psychiatric symptoms and schizophrenia, is that a small portion of individuals suffering from childhood or adolescent psychiatric symptoms are in the process of developing schizophrenia, and, consequently, carry an increased genetic risk for schizophrenia. Another source is that distinct psychiatric disorders are genetically correlated. Pleiotropy has been observed between schizophrenia and major depressive disorder, bipolar disorder, and autism spectrum disorder, but not between schizophrenia and attention deficit hyperactivity disorder.[21; 22] This genetic correlation could extend to various childhood and adolescent disorders.

Other molecular genetic and twin studies have already suggested genetic overlap between childhood and adult psychopathology.[23-25]

We hypothesized that genetic risk factors for schizophrenia are associated with elevated childhood and adolescent psychopathology scores. Given that the incidence of the prodromal phase and of psychiatric disorders genetically correlated to schizophrenia (i.e., major depression and bipolar disorder)[26] show a marked increase during adolescence, we expected the association between genetic risk for schizophrenia and internalizing and externalizing psychopathology to be higher in adolescence compared to childhood.

We studied the genetic overlap between schizophrenia and childhood psychopathology using polygenic risk score analyses. For a review of the method see[27-29]. In brief, the results from a genome-wide association (GWA) (meta-) analysis, in this case for schizophrenia, are used to calculate polygenic risk scores (PRS) in individuals in an independent target sample.These PRS are obtained by taking a set of top single nucleotide polymorphisms (SNPs), e.g. all SNPs with p-values below 0.1, 0.2 and so on, and multiplying the individual's genotypic score (0, 1 or 2) by the effect of the SNP. If the PRS are significantly related to a second trait in the target sample, in this case childhood psychopathology, this indicates that the two traits are influenced by overlapping genetic risk factors. In the current study, the PRS are based on the recent schizophrenia GWA meta-analysis for schizophrenia that comprised 36,989 cases and 113,075 controls and yielded 108 significantly associated loci.[19] Since the statistical power of polygenic risk score analyses depends to a large extend on the sample size of the discovery set, these meta-analysis results provide an excellent starting point to investigate the genetic overlap between schizophrenia and other traits.

PRS for schizophrenia, based on earlier smaller GWA meta-analyses, have previously been associated with cannabis use,[30] cognitive decline,[31] immune disorders[32] , and quantitative measures of psychosis.[33] These results show that this method is suitable for testing genetic relationships between different traits and schizophrenia that have long been implied or suspected.

Data on childhood psychopathology were available from two large longitudinal population based cohorts as target samples. We tested the association between the schizophrenia PRS and DSM-IV[34] based assessments of anxiety, depression,

attention deficit hyperactivity disorder (ADHD) ,and oppositional deviant disorder and conduct disorder (ODD/CD).

Methods

Subjects

The Netherlands Twin Register (NTR) follows young and adult twins (YNTR and ANTR). In the YNTR, twins are registered by their parents and followed from birth onwards. Until age 12, parents complete surveys to report on their twins. From age 14 onwards, information is collected by means of self-report.[35] The ANTR was established in 1987. Adolescent and adult twins were recruited through city-councils. The minimum age to participate was age 12, surveys were sent around every two to three years and data collection is ongoing. [36] In the current study, the maternal ratings of childhood psychopathology at age 7, 10, and 12 collected in the YNTR were analyzed. For the adolescent period, self-report data collected in the YNTR and the ANTR at age 14, 15, or 16 years were combined. The number of genotyped children at each age group varied between 1758 and 1956 (Table 1).

Table 1: Sample sizes per age group for the NTR, ALSPAC and combined

	NTR	ALSPAC	total
Age 7	1953	5665	7618
Age 10	1865	5506	7371
Age 12	1744	5111	6855
Age 15	1758	4131	5889

The ALSPAC birth cohort consists of mothers and their children, born between 1990 and 1991 in the Avon area in the UK.[37] The ALSPAC cohort includes ratings of psychopathology at age 7, 10, 13, and 15. The number of genotyped children at each age group varied between 4131 and 5665 (Table 1). Ethics approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. The study website contains details of all the data available through a fully searchable data dictionary (http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary).

Measures

In the NTR, psychopathology was measured using the age appropriate versions of the Achenbach System of Empirically Based Assessment (ASEBA). These assessments include DSM-IV based symptom scales. For age 7 to 12, maternal Child Behavior CheckList (CBCL) ratings were analyzed on the anxiety disorder scale (anxiety), the affective disorder scale (depression), the attention deficit hyperactivity disorder scale (ADHD), the oppositional deviant disorder scale and the conduct disorder scale. The last two were combined into an externalizing scale (ODD/CD).[38] From age 14 onwards, self-ratings were analyzed .[38] As the self-report item set varied slightly across surveys conducted between 1991 and 2012, the mean item scores per subscale were calculated.

In ALSPAC, psychopathology was assessed using the development and wellbeing assessment (DAWBA), which measures the presence of symptoms required for a DSM-IV diagnosis.[41] Based on the scales available in the NTR, the following disorders were included in the analyses: any anxiety disorder (anxiety), major depression (depression), attention deficit hyperactivity disorder(ADHD), and combined oppositional deviant disorder and conduct disorder(ODD/CD).[42] Any anxiety disorder included generalized anxiety disorder, specific phobia, social phobia (at age 7, 10, 13, and 15), separation anxiety disorder (at age 7, 10, and 13), and panic disorder and agoraphobia (at age 15).As well as generating a dichotomous outcome of being affected or not, the DAWBA can generate ordered categorical measures. These DAWBA band scores, which range from 0 to 5, correspond to probabilities of <0.01%, 0.5%, 3%, 15%, 50%, and >70% of satisfying diagnostic criteria. At ages 7, 10, and 13 all ratings were maternal ratings. At age 15, ADHD and CD/ODD were rated by the mother while anxiety and depression were self-rated.

Genotyping

The NTR participants were genotyped on Affymetrix 6.0, Affymetrix-perlegen 5.0, Illumina 370 and 660. Omni express (1M) array specific calls and cleaning were performed. SNPS were removed if The allele frequency difference with the reference set was above .20, the MAF was < 1%, if the HWE p-value was < 0.00001 or if the call rate was <95%. Data from the different platforms were subsequently imputed to a single reference genome (1000 genomes build 37 HG19) using MACH and miniMACH. For a more extensive overview of the QC procedure and imputation steps performed procedure see elsewhere.[43]

In ALSPAC, children were genotyped using the Illumina HumanHap550 quad chip genotyping platforms. The resulting raw genome-wide data were subjected to standard quality control methods. Individuals were excluded on the basis of gender mismatches, minimal or excessive heterozygosity, disproportionate levels of individual missingness (>3%), and insufficient sample replication (IBD < 0.8). Population stratification was assessed by multidimensional scaling analysis, and compared with Hapmap II (release 22); all individuals of non-European ancestry were removed. SNPs with a minor allele frequency of < 1%, a call rate of < 95%, or evidence for violations of Hardy-Weinberg equilibrium (P < 5E-7) were removed. Cryptic relatedness was measured as proportion of identity by descent (IBD > 0.1). Related subjects that passed all other quality control thresholds were retained during subsequent phasing and imputation. Imputation of the target data was performed using Impute V2.2.2 against the 1000 genomes phase 1 version 3 reference panel, using all 2186 reference haplotypes (including non-Europeans).

Polygenic risk scores

The polygenic risk scores (PRS) indicating an individual's genetic risk of schizophrenia were based on the results obtained in the most recent genome wide association meta-analysis for schizophrenia[19]. The scores were based on 102,636 SNPs included in the polygenic risk training set created by the authors of the original study (available online: http://www.med.unc.edu/pgc/downloads). Following the supplemental material of the schizophrenia GWAS[19], the inclusion criteria for these SNPs are: "... minor allele frequency above 10%, high imputation quality (INFO > .9). Data were then LD pruned and clumped removing variants within 500kb and associated above r^2 > .1 of another variant of bigger effect."

For each NTR and ALSPAC participant, we calculated 5 PRS based on 5 sets of SNPs. Inclusion was based on the significance of the effect in the schizophrenia meta-analysis with p-value cutoffs of < .001, < .01, < .1, <.3, < .5. For all included SNPs, the log odds are multiplied by the dosage score and summed per individual.

Statistical analyses

Regression analyses were performed to test whether the schizophrenia PRS predicted the childhood and adolescent measures of anxiety, depression, ADHD, and ODD/CD. In the NTR sample, six principal components correcting for genotype platform, a second set of 3 principal components correcting for ancestry, and sex were included as covariates. As the NTR contained related individuals, the regression was performed using a generalized estimation equation with exchangeable background correlations within family, and robust standard errors. This procedure adequately corrects for the presence of related individuals in the sample[44]. As the ALSPAC sample is genetically homogeneous, no principal components were added, and only sex was included as covariate. An ordered logistic regression was performed since the DAWBA bands are ordered categorical variables.

For both NTR and ALSPAC, the variance of the traits and the PRS were standardized (i.e., unit variance). The regression coefficients (beta's) from the two samples were combined using inverse variance weighting. Our null hypothesis was that all inverse variance weighted beta's equal zero. Our alternative hypothesis was that childhood psychopathology measures are positively associated with the schizophrenia PRS, which implies a one-sided hypothesis test. The associations in the meta- analysis were tested at a False Discovery Rate corrected alpha < 0.05 (p_{fdr}) [45] and an uncorrected significant alpha < 0.05 (p_n < .05). The p_{fdr} was calculated following the procedure of Benjamini & Hochberg[45]. This procedure adjusts the p-values in such a way that the proportion of false discoveries is controlled for. FDR is less strict then Bonferroni correction, and thus provides better power to detect an effect. The uncorrected p-values are informative to observe patterns in the results, such as consistent, but not FDR corrected significant, signals over the ages, and can be informative for post hoc tests.

Results

Descriptives

Table 2 shows the mean scores on the DSM based scales of anxiety, depression, ADHD, ODD/CD for males and females in the NTR at the four ages. Table 3 provides the percentages of male and female ALSPAC participants with these diagnoses, defined as a score of 4 or 5 on the DAWBA. As previously reported, the prevalences of DAWBA diagnoses tend to be lower than in other general population studies. The 6 category DAWBA band was used as outcome variable since this is a more informative measure than the dichotomous DAWBA diagnosis.

As expected, at age 15, girls scored substantially higher on internalizing measures (anxiety & depression). Boys scored higher on the externalizing measures (ADHD, ODD/CD) at all ages, except for the NTR at age 15. Sex was included as a covariate in the final analysis to account for this difference.

Table 2 Descriptives (mean and SD) per age for the NTR standardized scales for females and males.

	Female mean	SD	Male mean	SD
ODD/CD 7	-0.12	0.89	0.13	1.10
OOD/CD 10	-0.13	0.86	0.15	1.13
ODD/CD 12	-0.09	0.92	0.10	1.07
OOD/CD 15	0.01	0.96	-0.01	1.06
ADHD 7	-0.10	0.96	0.11	1.03
ADHD 10	-0.12	0.96	0.14	1.03
ADHD 12	-0.10	0.95	0.12	1.05
ADHD 15	0.06	1.00	-0.09	0.99
Depression 7	0.08	1.07	-0.09	0.90
Depression 10	0.03	1.05	-0.04	0.94
Depression 12	0.07	1.07	-0.08	0.90
Depression 15	0.14	1.08	-0.21	0.82
Anxiety 7	0.01	0.98	-0.01	1.03
Anxiety10	0.02	0.99	-0.02	1.01
Anxiety 12	0.04	0.97	-0.05	1.03
Anxiety 15	0.23	1.03	-0.34	0.85

Table 3: Per age, the percentage of ALSPAC female and male participants affected with a psychiatric disorder based on a score of 4 or 5 on the DAWBA bands.

	Female	Male
ODD/CD 7	1.71	5.42
OOD/CD 10	1.73	4.51
ODD/CD 12	2.99	3.95
OOD/CD 15	3.37	4.02
ADHD 7	0.69	3.12
ADHD 10	0.55	2.43
ADHD 12	0.55	2.01
ADHD 15	0.42	1.25
Depression 7	0.44	0.74
Depression 10	0.85	0.85
Depression 12	0.71	0.89
Depression 15	2.27	0.98
Anxiety 7	1.34	1.86
Anxiety10	1.76	1.95
Anxiety 12	1.17	1.22
Anxiety 15	3.06	0.77

In the meta-analysis, based on the betas and their standard errors, an inverse variance weighted mean on the beta's across cohorts and a fixed effect p-value

were computed (Figure 1). The association between schizophrenia PRS and childhood and adolescent psychopathology was FDR corrected significant for anxiety at age 10 for PRS with inclusion cutoffs of $p < 0.01$ (beta= 0.064, $p_{fdr} = 0.02$) , $p < 0.1$ (beta= 0.057, $p_{fdr} = 0.03$), $p < 0.3$ (beta=0.063, $p_{fdr} = 0.02$), and $p < 0.5$ (beta=0.059, $p_{fdr} = 0.03$). There were also rather consistent patterns of uncorrected significant ($p_n < 0.05$) associations with anxiety at age 7, and with depression at age 7 and 10. Some uncorrected significant results were observed for anxiety and depression at age 12/13, tough in both cases only for a single risk score. The weighted mean standardized beta is equivalent to a semi partial correlation between PRS and the psychopathology score. Squaring of the weighted mean standardized beta's reflect the R^2, i.e., a measure of the proportion of variance explained. For the nominally and FDR corrected significant results, the proportion of variance explained varied between 0.11% and 0.41%.

The anxiety results were mainly driven by a stronger effect in the NTR sample, while the depression results appeared stronger in ALSPAC (See supplemental figures S1, S2 and S3).

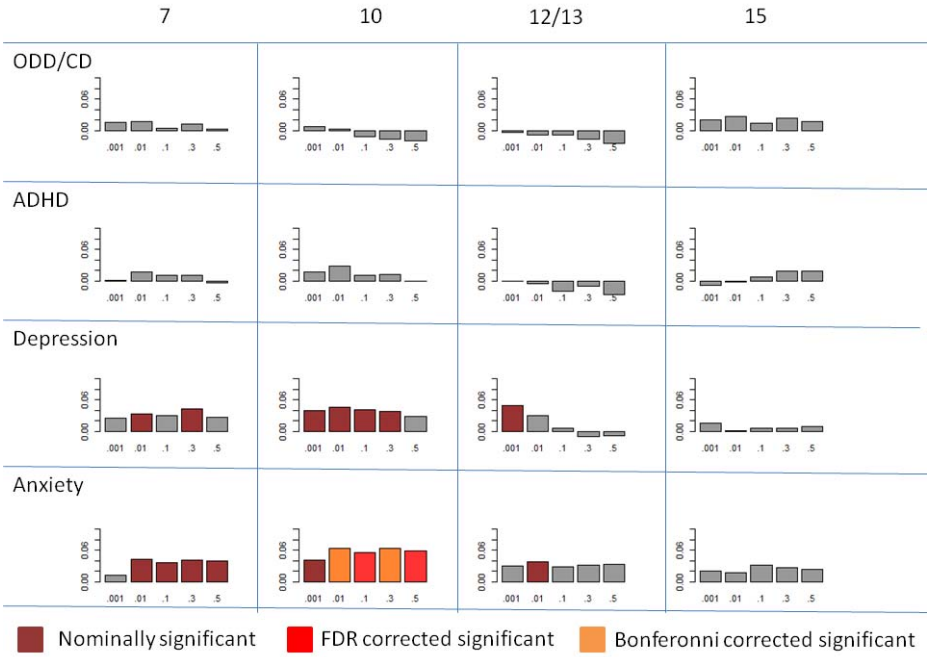


Figure 1: The standardized beta's indicating the effects of the 5 polygenic scores on the 4 DSM scales at age 7, 10, 12/13 and 15 in the meta-analysis.

As we observed larger beta's (mean = 0.030) for the regression of internalizing disorders (anxiety and depression,) on the schizophrenia PRS, than for the regression of externalizing disorders (ADHD and ODD/CD) on PRS (mean = 0.0035), a post-hoc test was performed to examine mean differences between the beta's for internalizing disorders and the beta's for externalizing disorders. We performed a two-way ANOVA with the meta-analyses beta's as the dependent variable. The independent variables were: a categorical variable coding internalizing versus externalizing disorders, a categorical variable coding any disorder versus anxiety disorders, and a categorical variable coding age (4 Bins: 7, 10 ,12/13 and 15). This test revealed that the inverse variance weighted beta's were higher for internalizing disorders than for externalizing disorders ($F = 21{,}148$, $df = 1$, $p < 0.0001$). The second fixed effect revealed that the means of the anxiety related effect sizes were even higher than the means of the depression related effect sizes ($F = 10.027$, $df = 1$, $p = 0.002$). A significant age effect was observed ($F = 7.706$, $df = 3$, $p < 0.001$). However, the effect was not the expected monotonic increase of effect sizes with increasing age. In Figure 2, the p-values for internalizing and externalizing disorders are plotted in seperate histograms. The expected distribution of p- values under the null hypothesis of a uniform distribution is indicated with a red line.

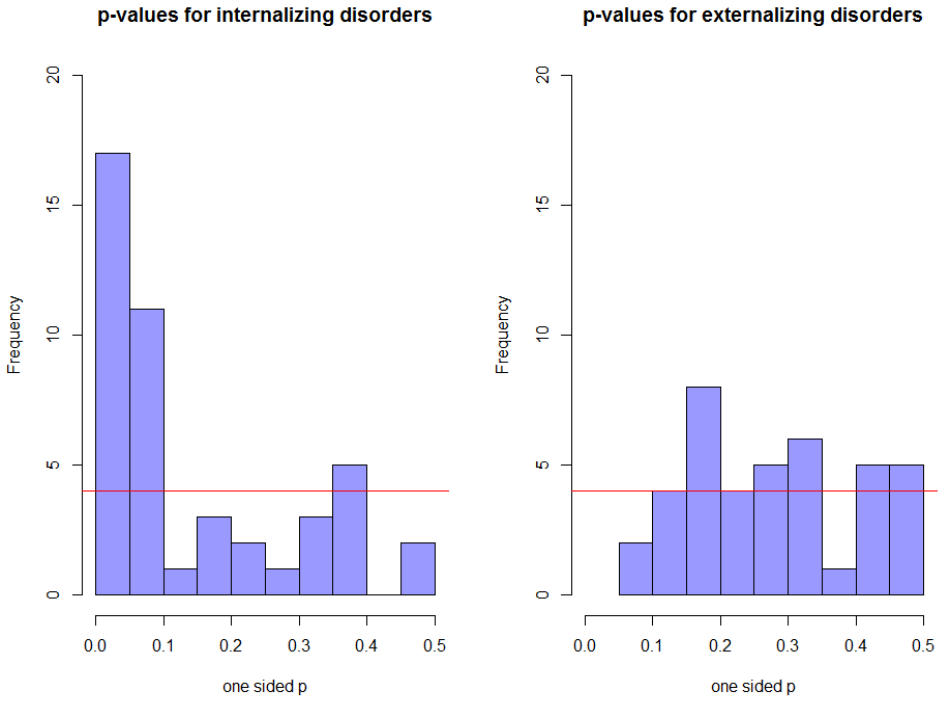


Figure 2a and 2b: The frequency of p-values of the regression of internalizing and externalizing disorders on the polygenic risk scores. The red line indicates the expected distribution of p-values under the null hypothesis.

Figure 2 shows that the p-values derived from the regression analyses of the externalizing phenotypes do not depart from the expectation under the null. The distribution of the 40 p-values associated with the regression of internalizing disorder scores on PRS show a steep over representation of low p-values when compared to the expected distribution of p-values under de null hypothesis .

Discussion

We investigated whether the phenotypic associations between schizophrenia and childhood psychopathology, as established in longitudinal, retrospective, and high risk studies[7-10], are explained by shared genetic risk factors. Our analyses yielded FDR corrected significant associations between anxiety at age 10 and schizophrenia PRS. Our analyses further revealed a pattern of uncorrected significant results for anxiety at age 7 and depression at age 7 and 10, but not at age 12/13 and 15. Post-hoc analyses revealed higher mean effect sizes for the internalizing phenotypes than for the externalizing phenotypes.

Therefore, we conclude that our results are suggestive of a genetic correlation between childhood internalizing psychopathology and schizophrenia. Since there was no evidence for genetic overlap between schizophrenia and externalizing childhood psychopathology, including ADHD and ODD/CD, we reject our hypothesis of a broad effect of schizophrenia risk on all childhood psychopathology measures. Our hypothesis of an increase in effect with age was also not supported.

The results presented here are partly consistent with previous findings from the psychiatric genomics consortium (PGC) cross disorder working group.[21] The PGC reported genetic correlations between schizophrenia and lifetime depression, but not between schizophrenia and ADHD.[22] This suggests that the association between externalizing psychopathology in childhood and the risk of schizophrenia, as found in previous studies[9; 11; 12] , is not explained by an overlap in genetic risk factors. Possibly, the externalizing symptoms preceding schizophrenia are influenced by other genetic variants than externalizing symptoms not associated with schizophrenia. This would amount to genetic heterogeneity, i.e., different genetic factors resulting in similar phenotypes.

We do not have a clear explanation for the absence of an effect at age 12/13 and 15, which would be expected based on the observed association with lifetime depression. For internalizing psychopathology, self-ratings were used at age 15, while at the other ages maternal ratings were used. But as the 12/13 year old ratings also did not show an association with the schizophrenia PRS, the difference in rater cannot (fully) explain the lack of significant results after age 10. The absence of signal for self-ratings at age 15 could be related to poor disease insight in people at risk for schizophrenia. However, poor insight is only related with actual schizophrenia in patients and not in people at risk. Furthermore, a small study found self-rating instruments to be valid, even in schizophrenia patient with poor disease insight[46]. Phenotypic heterogeneity could of course be an explanation with symptoms of anxiety and depression in early adolescence being a broader phenotype than lifetime depression in adults. However, it remains a question why this is different in childhood.

One strength of our study were the substantial sample sizes of the discovery and target samples. The schizophrenia PRS were based on a large discovery set and the GWA study had revealed more than 108 significant loci. The target sample varied between 6416 and 7610 for the different disorders at different ages, which is higher than the required number of ~2,000 subjects generally indicated as sufficient for PRS analysis[27]. The use of different measures of psychopathology in ALSPAC and NTR may be considered a limitation. However, these measures have both been successfully related to DSM-IV diagnoses[39-41]. Thus, it is reasonable to assume that they are associated to the same underlying construct. Associations between the schizophrenia PRS and childhood and adolescent psychopathology should therefore be consistent over these measures.

To conclude, we observed suggestive evidence for a genetic association between schizophrenia and internalizing psychopathology in childhood. This was in contrast with the results for externalizing psychopathology for which no genetic overlap with schizophrenia was detected. Future research is warranted to confirm our results, and further characterize the possible genetic association between internalizing disorders and schizophrenia.

References

1. Kessler R.C., Amminger G.P., Aguilar-Gaxiola S., Alonso J., Lee S., and Ustun T.B. (2007) Age of onset of mental disorders: a review of recent literature. Current opinion in psychiatry *20*, 359

2. Heinrichs R.W., Zakzanis K.K. (1998) Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. Neuropsychology *12*, 426

3. Kahn R.S., Sommer I.E. (2014) The neurobiology and treatment of first-episode schizophrenia. Molecular psychiatry

4. Fusar-Poli P., Bechdolf A., Taylor M.J., Bonoldi I., Carpenter W.T., Yung A.R., and McGuire P. (2013) At risk for schizophrenic or affective psychoses? A meta-analysis of DSM/ICD diagnostic outcomes in individuals at high clinical risk. Schizophrenia bulletin *39*, 923-932

5. Woodberry K.A., Giuliano A.J., and Seidman L.J. (2008) Premorbid IQ in schizophrenia: a meta-analytic review. The American journal of psychiatry *165*, 579-587

6. Cunningham Owens D.G., Johnstone E.C. (2006) Precursors and prodromata of schizophrenia: findings from the Edinburgh High Risk Study and their literature context. Psychological medicine *36*, 1501-1514

7. Cannon M., Caspi A., Moffitt T.E., Harrington H., Taylor A., Murray R.M., and Poulton R. (2002) Evidence for early-childhood, pan-developmental impairment specific to schizophreniform disorder: results from a longitudinal birth cohort. Archives of general psychiatry *59*, 449-456

8. Kim-Cohen J., Caspi A., Moffitt T.E., Harrington H., Milne B.J., and Poulton R. (2003) Prior juvenile diagnoses in adults with mental disorder:

developmental follow-back of a prospective-longitudinal cohort. Archives of general psychiatry *60*, 709-717

9. Rossi A., Pollice R., Daneluzzo E., Marinangeli M.G., and Stratta P. (2000) Behavioral neurodevelopment abnormalities and schizophrenic disorder: a retrospective evaluation with the Childhood Behavior Checklist (CBCL). Schizophrenia research *44*, 121-128

10. Miller P.M., Byrne M., Hodges A., Lawrie S.M., and Johnstone E.C. (2002) Childhood behaviour, psychotic symptoms and psychosis onset in young people at high risk of schizophrenia: early findings from the Edinburgh High Risk Study. Psychological medicine *32*, 173-179

11. Muratori F., Salvadori F., D–Arcangelo G., Viglione V., and Picchi L. (2005) Childhood psychopathological antecedents in early onset schizophrenia. European psychiatry *20*, 309-314

12. Keshavan M.S., Diwadkar V.A., Montrose D.M., Rajarethinam R., and Sweeney J.A. (2005) Premorbid indicators and risk for schizophrenia: a selective review and update. Schizophrenia research *79*, 45-57

13. Meyer S.E., Bearden C.E., Lux S.R., Gordon J.L., Johnson J.K., O'Brien M.P., Niendam T.A., Loewy R.L., Ventura J., and Cannon T.D. (2005) The psychosis prodrome in adolescent patients viewed through the lens of DSM-IV. Journal of Child & Adolescent Psychopharmacology *15*, 434-451

14. Fusar-Poli P., Nelson B., Valmaggia L., Yung A.R., and McGuire P.K. (2014) Comorbid depressive and anxiety disorders in 509 individuals with an at-risk mental state: impact on psychopathology and transition to psychosis. Schizophrenia bulletin *40*, 120-131

15. Gajwani R., Patterson P., and Birchwood M. (2013) Attachment: Developmental pathways to affective dysregulation in young people at ultra-high risk of developing psychosis. British Journal of Clinical Psychology *52*, 424-437

16. Perkins D.O., Gu H., Boteva K., and Lieberman J.A. (2005) Relationship between duration of untreated psychosis and outcome in first-episode schizophrenia: a critical review and meta-analysis. American Journal of Psychiatry *162*, 1785-1804

17. Stafford M.R., Jackson H., Mayo-Wilson E., Morrison A.P., and Kendall T. (2013) Early interventions to prevent psychosis: systematic review and meta-analysis. Bmj *346*, f185

18. Tsuang M. (2000) Schizophrenia: genes and environment. Biological psychiatry *47*, 210-220

19. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014) Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421-427

20. Sullivan P.F., Kendler K.S., and Neale M.C. (2003) Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Archives of general psychiatry *60*, 1187-1192

21. Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nature genetics *45*, 984-994

22. Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. The Lancet *381*, 1371-1379

23. Nivard M.G., Dolan C.V., Kendler K.S., Kan K.J., Willemsen G., van Beijsterveldt C.E., Lindauer R.J., van Beek J.H.D.A., Geels L.M., and Bartels M. (2014) Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. Psychological medicine, 1-11

24. Kan K.J., Dolan C.V., Nivard M.G., Middeldorp C.M., van Beijsterveldt C.E., Willemsen G., and Boomsma D.I. (2013) Genetic and environmental stability in attention problems across the lifespan: evidence from the

Netherlands twin register. Journal of the American Academy of Child & Adolescent Psychiatry *52*, 12-25

25. Benke K.S., Nivard M.G., Velders F.P., Walters R.K., Pappa I., Scheet P.A., Xiao X., Ehli E.A., Palmer L.J., and Whitehouse A.J. (2014) A genome-wide association meta-analysis of preschool internalizing problems. Journal of the American Academy of Child & Adolescent Psychiatry *53*, 667-676

26. Kessler R.C., Berglund P., Demler O., Jin R., Merikangas K.R., and Walters E.E. (2005) Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Archives of general psychiatry *62*, 593-602

27. Wray N.R., Lee S.H., Mehta D., Vinkhuyzen A.A., Dudbridge F., and Middeldorp C.M. (2014) Research Review: Polygenic methods and their application to psychiatric traits. Journal of Child Psychology and Psychiatry *55*, 1068-1087

28. Purcell S.M., Wray N.R., Stone J.L., Visscher P.M., O'Donovan M.C., Sullivan P.F., Sklar P., Ruderfer D.M., McQuillin A., and Morris D.W. (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature *460*, 748-752

29. Wray N.R., Yang J., Hayes B.J., Price A.L., Goddard M.E., and Visscher P.M. (2013) Pitfalls of predicting complex traits from SNPs. Nature Reviews Genetics *14*, 507-515

30. Power R.A., Verweij K.J.H., Zuhair M., Montgomery G.W., Henders A.K., Heath A.C., Madden P.A.F., Medland S.E., Wray N.R., and Martin N.G. (2014) Genetic predisposition to schizophrenia associated with increased use of cannabis. Molecular psychiatry *19*, 1201-1204

31. McIntosh A.M., Gow A., Luciano M., Davies G., Liewald D.C., Harris S.E., Corley J., Hall J., Starr J.M., and Porteous D.J. (2013) Polygenic risk for schizophrenia is associated with cognitive change between childhood and old age. Biological psychiatry *73*, 938-943

32. Stringer S., Kahn R.S., de Witte L.D., Ophoff R.A., and Derks E.M. (2014) Genetic liability for schizophrenia predicts risk of immune disorders. Schizophrenia research *159*, 347-352

33. Derks E.M., Vorstman J.A., Ripke S., Kahn R.S., Ophoff R.A., and Schizophrenia Psychiatric Genomic Consortium (2012) Investigation of the genetic association between quantitative measures of psychosis and schizophrenia: a polygenic risk score analysis. PloS one *7*, e37852

34. American Psychiatric Association (1994) Diagnostic and statistical manual of mental diseases. DSM-IV. 4th edn. Washington (DC): American Psychiatric Association

35. van Beijsterveldt C.E., Groen-Blokhuis M., Hottenga J.J., Franic S., Hudziak J.J., Lamb D., Huppertz C., de Zeeuw E., Nivard M., and Schutte N. (2013) The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. Twin Research and Human Genetics *16*, 252-267

36. Willemsen G., Vink J.M., Abdellaoui A., den Braber A., van Beek J.H., Draisma H.H., van Dongen J., van 't Ent D., Geels L.M., and van Lien R. (2013) The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. Twin Research and Human Genetics *16*, 271-281

37. Boyd A., Golding J., Macleod J., Lawlor D.A., Fraser A., Henderson J., Molloy L., Ness A., Ring S., and Smith G.D. (2012) Cohort profile: the `children of the 90s` the index offspring of the Avon Longitudinal Study of Parents and Children. International journal of epidemiology, dys064

38. Achenbach T.M., Rescorla L.A., McConaughey S., Pecora P., Wetherbee K., and Ruffle T. (2003) The Achenbach system of empirically based assessment. Handbook of psychological and educational assessment of children: Personality, behavior, and context *2*, 406-432

39. Ebesutani C., Bernstein A., Nakamura B.J., Chorpita B.F., Higa-McMillan C.K., Weisz J.R., and Research Network on Youth Mental Health (2010)

Concurrent validity of the Child Behavior Checklist DSM-oriented scales: Correspondence with DSM diagnoses and comparison to syndrome scales. Journal of Psychopathology and Behavioral Assessment *32*, 373-384

40. Bellina M., Brambilla P., Garzitto M., Negri G.A., Molteni M., and Nobile M. (2013) The ability of CBCL DSM-oriented scales to predict DSM-IV diagnoses in a referred sample of children and adolescents. European child & adolescent psychiatry *22*, 235-246

41. Goodman R., Ford T., Richards H., Gatward R., and Meltzer H. (2000) The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. Journal of child psychology and psychiatry *41*, 645-655

42. Goodman A., Heiervang E., Collishaw S., and Goodman R. (2011) The 'DAWBA bands` as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. Social psychiatry and psychiatric epidemiology *46*, 521-532

43. Nivard M.G., Mbarek H., Hottenga J.J., Smit J.H., Jansen R., Penninx B.W., Middeldorp C.M., and Boomsma D.I. (2014) Further confirmation of the association between anxiety and CTNND2: replication in humans. Genes, Brain and Behavior *13*, 195-201

44. Minicâ C.C., Dolan C.V., Kampert M.M., Boomsma D.I., and Vink J.M. (2014) Sandwich corrected standard errors in family-based genome-wide association studies. European journal of human genetics

45. 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 (Methodological), 289-300

46. Bell M., Fiszdon J., Richardson R., Lysaker P., and Bryson G. (2007) Are self-reports valid for schizophrenia patients with poor insight?

Relationship of unawareness of illness to psychological self-report instruments. Psychiatry research *151*, 37-46

Chapter 5 Supplemental figures.

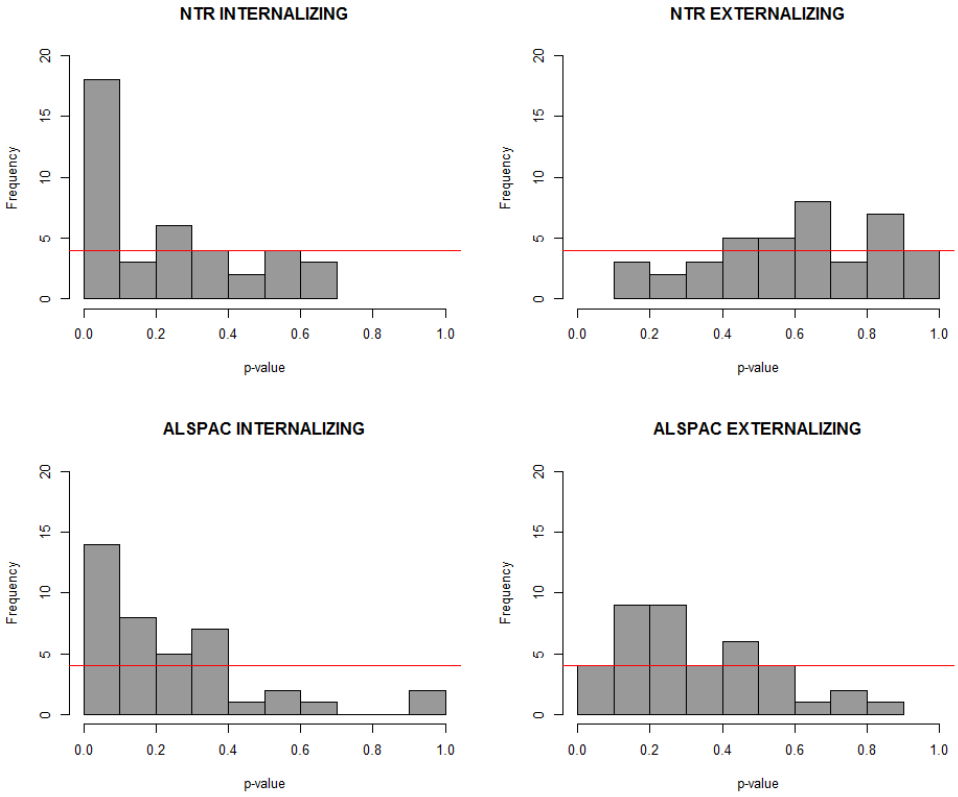


Figure S1: The p-values obtained in the regression analyses in NTR and ALSPAC predicting internalizing and externalizing scores by schizophrenia PRS. Expected (uniform) distribution is indicated by the red line. In the NTR sample there is a clear indication of an abundance of low p-values for the regression of internalizing scores on PRS. The results in ALSPAC show the same pattern.

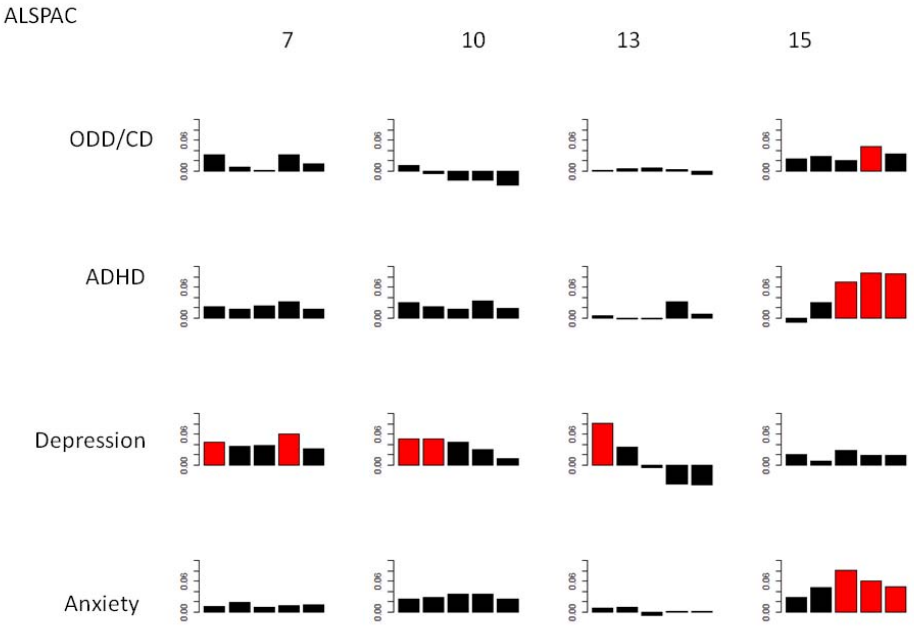


Figure S2: Effect sizes (Beta's) of the ordinal logistic regression of DAWBA band scores on the 5 schizophrenia PRS in the ALSPAC sample for age 7, 10, 13 and 15. Red bars are $p < 0.05$ in a one sided test.

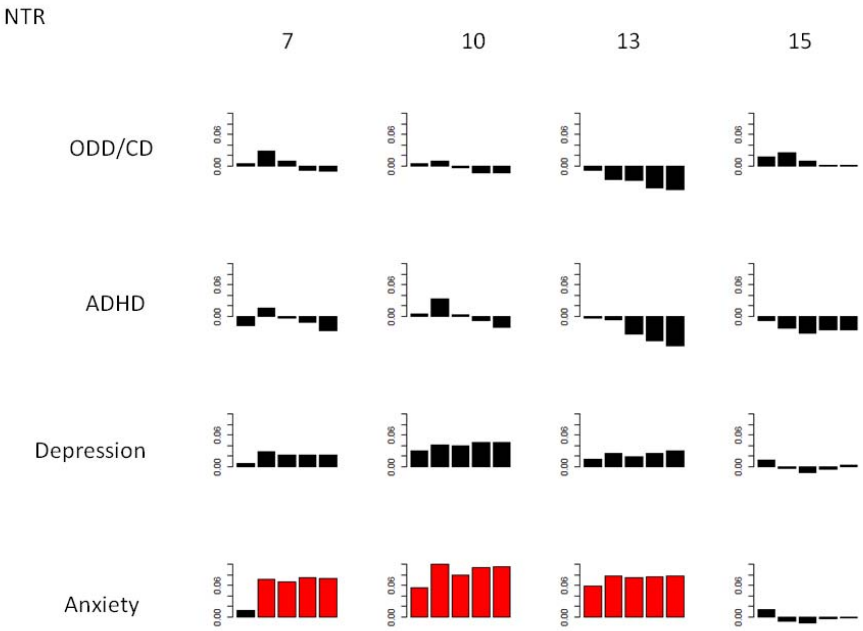


Figure S3: Effect sizes (Beta's) of the GEE regression of CBCL/YSR scores on the 5 schizophrenia PRS in the NTR sample for age 7, 10, 12 and 15. Red bars are $p < 0.05$ in a one sided test.

Chapter 6: Further confirmation of the association between anxiety and *CTNND2* : replication in humans*

This chapter is based on the publication:

Nivard, M. G., Mbarek, H., Hottenga, J. J., Smit, J. H., Jansen, R., Penninx, B. W., Middeldorp, C.M. & Boomsma, D. I. (2014). Further confirmation of the association between anxiety and CTNND2: replication in humans. *Genes, Brain and Behavior*, *13*(2), 195-201.

Abstract

The rat genome sequencing and mapping consortium found evidence for an association between the catenin-δ2 gene (*CTNND2)* and anxious behaviour. We replicated these results in humans by carrying out a genetic association test in patients with panic-disorder, social phobia, generalized anxiety disorder and/or agoraphobia (N=1714) and controls (N= 4125). We further explored the association between *CTNND2* and other psychiatric disorders based on publicly available genome-wide association results. A gene-based test showed that SNPs in *CTNND2* have a significantly increased signal ($p<1e^{-5}$) and decreased p-values. SNP rs1012176 showed the strongest association with any anxiety disorder (Odds ratio: 0.8128, SE = 0.063, p =.00099), but this effect was not significant after correction for multiple testing. In available genome -wide association results from the Psychiatric Genomics Consortium we found that SNPs in *CTNND2* collectively showed an increased signal for Schizophrenia ($p < 1e^{-5}$) and Major Depressive Disorder ($p < 1e^{-5}$), but not for Bipolar Disorder. These signals remained significant after correction for potential confounders. The association between *CTNND2* and anxiety was not strong enough to be picked up in the current generation of human genome-wide analyses, indicating the usefulness of and need for animal genetic studies to identify candidate genes for further study in human samples.

Introduction

The rat genome sequencing and mapping consortium recently published an extensive sequence based analysis of 160 complex phenotypes, including disease models for anxiety, diabetes, hypertension, aortic elastic lamina ruptures, multiple sclerosis and osteoporosis and measures of risk factors for common diseases such as lipid and cholesterol levels in outbred rats (Rat Genome Sequencing Consortium, 2013).[1] The Rat Genome Consortium paper reported 28 QTLs at which only a single gene contained candidate variants. At one QTL, a new gene for an anxiety-related phenotype was implicated. The catenin δ2 gene (*CTNND2,* encoding catenin δ2) was associated with anxiety related traits in rats. This QTL explained 5% of the variance (see their online supplementary Table 3). *CTNND2* has also been found to be related to reduced hippocampal volume and synaptic dysfunction but not to anxiety related traits in knockout mice (Israely et al. 2004).[2] *CTNND2* is highly expressed in the human and fetal brain, and has been implicated in neuronal functioning, adhesion, and migration (Lu et al. 1999).[3] The gene has not yet been studied as a candidate for anxiety disorders in humans.

We explored the effects of *CTNND2* on anxiety disorders in 1714 patients with an anxiety disorder and 4125 screened controls who take part in the Netherlands Study of Depression and Anxiety (NESDA) (Penninx et al. 2008)[4] and the Netherlands Twin Register (NTR) (Lubke et al. 2012; Willemsen et al. 2010; Willemsen et al. 2013).[5-7] Firstly, we tested in a gene-based test whether all SNPs collectively obtained significantly lower p-values than expected, indicating an association between the gene and the phenotype. Next, individual SNPs within *CTNND2* were regressed on anxiety status using logistic regression.

Secondly, the association between *CTNND2* and major depressive disorder (MDD) (Ripke et al. 2012),[8] Schizophrenia (SCZ)[9] (Ripke et al. 2011) and Bipolar Disorder (BIP)[10] (Sklar et al. 2011) was explored in published results based on the mega-analyses of these disorders by the Psychiatric Genomics Consortium (PGC). All three PGC genome-wide association (GWA) studies use a dichotomous case/control phenotype and logistic regression to determine the association of SNPs with each disorder. The PGC results were inspected for inflation in p-values for all SNPs in the *CTNND2* gene to test for an association between the gene and the phenotype and we again looked at the best individual SNP corrected for multiple testing.

Methods

Subjects

Anxiety disorder cases were derived from the Netherlands Study of Depression and Anxiety (NESDA) and from the Netherlands Twin Register (NTR). There were 1747 patients who met diagnostic criteria for a DSM-IV defined anxiety disorder (panic disorder, agoraphobia, social phobia and/or generalized anxiety disorder) as assessed with the Composite Interview Diagnostic Instrument (CIDI) (Ter Smitten, Smeets, & Van den Brink 1998).[11] Controls came from the NESDA and NTR studies . In NESDA, the absence of any lifetime depressive and anxiety disorder was assessed by the CIDI (lifetime version 2.1). In NTR, controls were selected in a similar way as described in Boomsma et al. (2008).[12] The selection was based on low scores on depression, anxiety or neuroticism scales in longitudinal surveys. These surveys contained the neuroticism and somatic anxiety scales of the Amsterdamse Biografische Vragenlijst (Wilde 1970),[13] the Beck Depression Inventory (Beck, Rial, & Rickels 1974),[14] the anxious depression scale of the Adult Self Report (Achenbach & Rescorla 2003)[15] and the State Trait Anxiety Inventory – Trait version (Van der Ploeg, Defares & Spielberger 1979).[16] In total, 4125 controls met the selection criteria.

Genotyping and imputation

Whole blood and /or buccal DNA samples were collected for various projects done by the NTR and NESDA studies (see: Boomsma et al. 2008;Scheet et al. 2012;Sullivan et al. 2009;Willemsen et al. 2010).[6,12,17,18] DNA extraction and purification of these samples have been performed at various stages in time, following several manufacturer specific protocols in order to obtain the best quality and concentration prior to SNP platform genotyping. Genotyping subsequently has been done on multiple chip platforms, for several partly overlapping subsets of the total sample collection. Chronologically the following platforms have been used Affymetrix Perlegen 5.0 (N=3840), Illumina 370 (N=290), Illumina 660 (N=1501), Illumina Omni Express 1M (N=445) and Affymetrix 6.0 (N=10412, 5 subsets). After array specific data analysis, genotype calls were made with the platform specific software (Genotyper, Beadstudio, Birdseed).

Quality control was done within and between chip platforms. For each platform the individual SNP markers were lifted over to build 37 (HG19) of the Human

reference genome, using the LiftOver tool (Kuhn et al. 2009).[19] SNPs that were not mapped at all, SNPs that had ambiguous locations, and SNPs that did not have matching - or strand opposite alleles were removed. Subsequently, the data were strand aligned with the 1000 Genomes phase 1 Integrated release version 3 ALL panel of March 2012 (McVean et al. 2012).[20] SNPs from each platform were removed if they still had mismatching alleles with this imputation reference set, if the allele frequencies differed more than 0.20 with the reference set, if the MAF was < 1%, if the HWE p-value was < 0.00001 or if the call rate was <95%. All samples were excluded from the data if their expected sex did not match their genotyped sex, if the genotype missing rate was above 10% or if the Plink F inbreeding value was either > 0.10 or < -0.10. After these steps the data of the individual chips were merged into a single dataset using the Plink 1.07 software(Purcell et al. 2007).[21]

Within the merged set IBD was calculated between all possible pairs of individuals and compared to the expected family structure of the NTR and NESDA studies. Samples were removed if the data did not match the expected IBD sharing, or if potentially consistent with biographic data, corrections were made to the family structure. DNA samples that were typed on multiple platforms were tested if the overlapping SNPs had a concordance rate above 99.0%. If this was not true, all data of these samples were removed. On the merged data, HWE and MAF SNP filters were re-applied, and the reference allele frequency difference <0.20 checks. As a final prior step to imputation SNPs with C/G and A/T allele combinations were removed if the MAF was between 0.35 and 0.50 to avoid wrong strand alignment.

Imputation was done using the two stage approach. Pre-imputation phasing and imputation of genotype platform specific SNPs was done using the MACH software(Li et al. 2010).[22] Subsequently, imputation of the reference set was done with Minimach. To avoid issues with monozygotic (MZ) twin pairs, prior to imputation a single person of a monozygotic twin with the highest SNP call rate of a pair was selected. Post imputation, the resulting imputed genotypes were duplicated back to the co-twin in the data. Based on phenotype, a single MZ twin was then selected for analysis in this study. From the imputed genotype data, for the cases and controls of this particular study all SNPS between the *CTNND2* gene borders as reported on genecards.org (Safran et al. 2010)[22] were included for analysis. In total 1349 SNPs in the *CTNND2* gene met the post imputation QC standards (MAF >.05, INFO>.8, INFO < 1.1 & HWE <$1e^{-3}$). Of these 1349 SNPs

475 were directly genotyped, and 874 are imputed.

Lookup in Psychiatric Genomics Consortium (PGC)
We retrieved results from the mega-analyses of genome-wide association studies by PGC on MDD, Bipolar Disorder and Schizophrenia. These data are available at https://pgc.unc.edu/Sharing.php#SharingOpp and more conveniently and completely at http://www.broadinstitute.org/mpg/ricopili/. Using the PGC summary statistics, the association between *CTNND2* and these psychiatric phenotypes was explored further. We applied the same QC standards to the SNPs reported by PGC as to the SNPs tested in the NTR/NESDA sample(MAF >.05, INFO>.8 and <1.1). We then retrieved the PGC p-values for SNPs in the *CTNND2* gene plus a 200kb area around the gene.
Statistical analyses.
For various significance tests, a correction for the number of individual signals in *CTNND2* was needed. To determine the number of independent signals, a PCA (Principal Component Analysis) was carried out on the pair-wise correlation matrix between all SNPs in the gene. The number of independent signals was determined to be the number of components needed to explain 95% of variance in the SNPs in *CTNND2*. Pair-wise LD was retrieved using the SNAP (SNP Annotation and Proxy Search) online tool (Johnson et al. 2008).[24] This tool can be used for SNPs to identify and to annotate nearby SNPs in linkage disequilibrium based on HapMap or 1000 genomes.
The anxiety disorder phenotype was regressed on the SNPs in *CTNND2* using logistic regression. Regression analyses were controlled for sex, study of origin within NTR/NESDA, genotyping platform and 3 principle components to control for population stratification (Abdellaoui et al. 2013).[25] A sandwich estimator was used to control for the presence of related individuals in the sample(Purcell et al. 2007).[21]
To determine if significant signals exist within the entire gene, we tested whether the p-values off SNPs in the *CTNND2* gene as found in the anxiety disorder results, or reported by the PGC, were significantly lower than expected. The effect size for this test is λ (lambda), the inflation of p-values over the expected distribution of p-values. The significance of λ was tested against different null hypotheses using three different approaches (see table 1).

Table 1: Models used to test the significance of λ

Method	p-values sampled from:	Null hypothesis
1: Uniform	Uniform distribution between 0 and 1	Observed λ is equal to the statistical expectation of λ (λ = 1)
2: Uniform controlled for population structure	Uniform distribition betweeen 0 and 1. λ then multiplied by the genome-wide λ	Observed λ is equal to the statistical expectation of λ in the case of no effect in the presence of population stratification (λ = genome wide λ)
3 Uniform independent signals and population structure	Sample from a uniform distribution the size of the number of independent signals in *CTNND2*. λ then multiplied by the genome-wide λ	Observed λ is equal to the expected λ for a set of p-values based on the number of independent signals in *CTNND2*, thereby taking into account LD, and pop stratification as in method 2 (λ for 1366 SNPs = λ 261 for signals)

We bootstrapped 10,000 distributions of p-values under these 3 null distributions. This yielded a mean λ and the variance of λ that reflects the distribution under the null distribution for each of these 10,000 bootstrapped sets of p-values . Next, the significance of λ observed in any anxiety disorder, MDD, Bipolar Disorder and Schizophrenia was tested against the expected λ obtained in each null distribution. To be certain the reported λ reflects actual effects, we step by step excluded other common causes of an inflated λ. First, we adapted the null hypothesis to reflect the fact that some inflation of λ is expected to be caused by population stratification and sample size. To counteract the effects of population stratification and sample size, the bootstrapped λ were drawn from samples that showed inflation by population stratification or sample size at the level of the genome wide λ for each trade (method 2 in table 1).The variance of the bootstrapped λ distribution is influenced by LD between the SNPs in *CTNND2*. Higher LD between the SNPs would result in fewer independent signals, if we base λ on sets of p-values representing a smaller number of signals the variance of these λ would increase. To take into account the LD observed between SNPs we created a null distribution in which the λ was based on the number of independent signals in *CTNND2* (method 3 table 1).

We also determined the significance of the best SNP in the gene for anxiety disorders as well as for SCZ, MDD and BIP based on the PGC results. To be clear about the significance of our results, we report p-values per best SNP which are uncorrected for multiple testing, and corrected for the number of independent signals in *CTNND2*.

Results

Test for the number of independent signals in CTNND2.
From the 1349 SNPs in *CTNND2* that met QC standards, 1256 were also found in the SNAP pair-wise LD tool. The pair-wise LD between these SNPs was entered into a PCA in R. The eigenvalues derived from this PCA suggested that 261 components are needed to explain 95% of variance and 446 components are needed to explain 99% of variance in the 1256 SNPs found in SNAP. The Kaiser criterion suggested 180 independent signals. Pairwise correlations between all 1349 snps in the *CTNND2* gene were also calculated in the NTR/NESDA sample

using a minimal cutoff of r=.10 (r^2=0.01). Principle component analysis was performed on the resulting pairwise correlation matrix. The Kaiser criterion indicated 124 independent components existed, the PCA suggested 254 components are needed to explain 95% of variance and 563 components are needed to explain 99% of variance. These results are similar to the ones obtained using SNAP suggesting the LD structures are similar, at least in dimensionality. The 3 criteria described did vary fairly widely in their conclusion on the number of independent signals, though they were in the same order of magnitude (i.e. hundreds of signals) and very similar in the 1000 genomes reference panel and the NTR/NESDA set. For all further significance testing we adjusted for 261 independent signals as found in SNAP, as adjusting for the Kaiser criterion would be too liberal and adjusting for 446 signals might be too strict.

Anxiety disorders in NTR/NESDA

All SNPs in the gene jointly showed significantly lower p-values than under the expected null distribution (λ = 1.34, P < 0.0001) using method 1 in table 1. Testing the λ using method 2 in table 1, taking into account population stratification and sample size, we again found the λ significantly exceeded the null (p < 0.0001). Testing the λ using method 3 in Table 1, taking into account the effects of LD, population stratification and sample size, the λ remained significant (p < 0.035). This inflation of p-values indicates that the gene shows a significantly stronger association to anxiety disorders than expected under the null hypotheses considered (table 1) indicating this gene has a significant association with anxiety disorders.

SNP rs1012176 in *CTNND2* showed the strongest association with anxiety disorders (Odds ratio: 0.8128, SE = 0.063, p =.00099) (figure 1). This individual SNP is not significant if corrected for the number of independent signals (p = .22). Table 2 provides the RS-numbers and locations for the 10 SNPs showing strongest association with Anxiety disorders. Some of these 10 best SNPs are in complete or high, but not total, LD with 2 SNPs in the coding region of *CTNND2*. 2 SNPs in the coding region, rs17802557 and rs1566622 are in LD with different SNPS from table 2 and are both synonymous mutations in the coding region of *CTNND2.* Complete LD indicates the SNPs associated with anxiety disorder and SNP in the coding region of *CTNND2* are completely dependent, the lack of total LD merely indicates the SNP would not be useful as a proxy for each other.

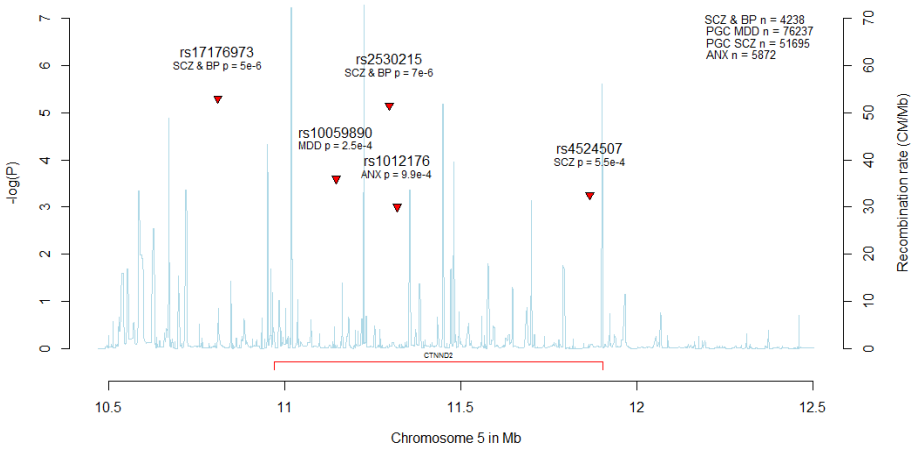


Figure 1: Regional plot for the *CTNND2* gene, with recombination rates plotted in blue. The best association results in the *CTNND2* gene are plotted in red for the GWA for anxiety in NESDA/NTR (ANX), and for 3 published GWA meta analyses for major depressive disorder (MDD) (Ripke et al. 2012), schizophrenia (SZC)[9] (Ripke et al. 2011) and schizophrenia and bipolar disorder (SZC&BP) combined (Wang, Liu, & Aragam 2010).[26]

Table 2: top 10 SNPs per disorder and their relation to functional SNPs.

Schizophrenia	RS-number	position	P	Relation to functional SNPs
	rs4524507	11919716	0.0005	
	rs4702836	11918658	0.0005	
	rs4379193	11920393	0.0006	
	rs4571470	11913452	0.0006	
	rs11741312	11911555	0.0006	
	rs12652699	11917077	0.0006	
	rs4330462	11913217	0.0007	
	rs6873547	11917482	0.0008	
	rs4257769	11913395	0.0010	
	rs4701924	11923554	0.0012	
MDD	RS-number	position	P	
	rs10059890	11199698	0.0002	SNP in complete LD (D'=1) but not total LD (r2 =0.087) with SNP rs2285975
	rs1859382	11200672	0.0002	SNP in complete LD (D'=1) but not total LD (r2 =0.087) with SNP rs2285975
	rs6859601	11200414	0.0002	SNP in complete LD (D'=1) but not total LD (r2 =0.087) with SNP rs2285975
	rs886527	11200814	0.0003	SNP in complete LD (D'=1) but not total LD (r2 =0.087) with SNP rs2285975
	rs6885587	11200114	0.0004	
	rs6880938	11200302	0.0005	
	rs10041627	11200082	0.0005	
	rs10073056	11199253	0.0005	
	rs730610	11192051	0.0007	
	rs2057795	11207597	0.0008	
Bipolar	RS-number	position	P	
	rs10044218	11466046	0.0130	
	rs7722906	11461708	0.0173	
	rs7728281	11469941	0.0274	
	rs16901579	11455272	0.0342	

	rs4702761	10955733	0.0367	
	rs7722560	11461527	0.0371	
	rs1860245	10956537	0.0371	
	rs2895578	10956553	0.0371	
	rs6884596	11456540	0.0372	
	rs1995364	10956781	0.0376	
Anxiety	RS-number	position	P	
	rs1012176	11320538	0.0009	
	rs17216753	11484112	0.0017	SNP in complete (D'=1) but not total LD (R2=0.015) with SNP rs17802557
	rs17805573	11488568	0.0018	SNP in complete (D'=1) but not total LD (R2=0.015) with SNP rs17802557
	rs11747109	11493331	0.0018	SNP in complete (D'=1) but not total LD (R2=0.015) with SNP rs17802557
	rs79213734	11495176	0.0019	SNP in complete (D'=1) but not total LD (R2=0.015) with SNP rs17802557
	rs10513094	11482045	0.0020	SNP in complete (D'=1) but not total LD (R2=0.015) with SNP rs17802557
	rs11948339	11378306	0.0036	
	rs2012187	11322037	0.0048	Moderate D' (0.681) and low R2 (0.038) with coding SNP rs1566622
	rs32128	11343384	0.0051	Moderate D' (0.668) and low R2 (0.096) with coding SNP rs1566622
	rs32129	11343200	0.0051	Moderate D' (0.668) and low R2 (0.096) with coding SNP rs1566622

PGC lookup

Major depressive disorder (Ripke et al. 2012)[8]: In total, 632 SNPs were present in the PGC MDD results that were in *CTNND2*, or in a 200kb window around, and met QC standards (MAF >.05, INFO>.8 and < 1.1). These SNPS showed inflation (λ =1.39, p <$1e^{-5}$) (figure 2, black). The λ was significant if tested using method 1 in Table 1 (p <0.0002) and remained significant when correcting for potential effects of population structure and sample sizes (p < 0.0013) (method 2 in Table 1) and after further correction for the effects of LD (p < 0.026) (method 3 in Table 1). These results indicate the p-values observed in *CTNND2* show a stronger association to MDD than expected, even when we took into account the effects of population stratification, sample size and LD within the gene.

The best SNP (rs10059890) in the published results of the mega-analysis of the PGC for MDD[8] (Ripke et al. 2012) had a nominally significant p-value of 0.00025 (Figure 1). Corrected for the number of independent tests this result was no longer significant (p = 0.06). Table 2 provides the RS-numbers and locations for the 10 SNPs showing strongest association with Major depressive disorder. The top SNP and several other SNPs present in table 2 are in complete, but not total, LD with SNP rs2285975. rs2285975 is a synonymous mutation in the coding region of CTNND2

Schizophrenia (Ripke, et al. 2011)[8]: A total of 614 SNPs in and around *CTNND2* were present in the PGC schizophrenia results and met QC standards (MAF >.05, INFO>.8 and < 1.1). These 614 SNPs showed large inflation (λ = 2.54; Figure 2). This inflation was significantly higher than 1 as tested using method 1 in Table 1 (p <0.0001). If we correct this test for the effects of population stratification and sample size, the observed inflation remained significant (method 2 in Table 1, p <0.0001). It also remained significant after further correction for the LD in *CTNND2* (method 3 in Table 1) (p <0.0001). These results indicate the p-values observed in *CTNND2* show a larger effect than expected, even when this expectation is corrected for population stratification, sample size and LD within the gene.

Uncorrected results for the single best SNP (rs4524507) showed an association (p = 0.00056; figure 1) with the *CTNND2* gene. Corrected for the number of independent signals the SNP effect is no longer significant (p = 0.13). Table 2 provides the RS-numbers and locations for the 10 SNPs showing strongest association with Schizophrenia.

Bipolar Disorder (Sklar et al. 2011)[10]: In total 1123 SNPs were present in the PGC results that are in and around *CTNND2* and met QC standards (MAF >.05, INFO>.8 and < 1.1). None of the individual SNPs were significant if corrected for multiple testing. These SNPs did not show an inflation over the expected null (figure 2). These results indicate that variants in *CTNND2* have no association to Bipolar Disorder. Table 2 provides the RS-numbers and locations for the 10 SNPs showing strongest association with Bipolar disorders.

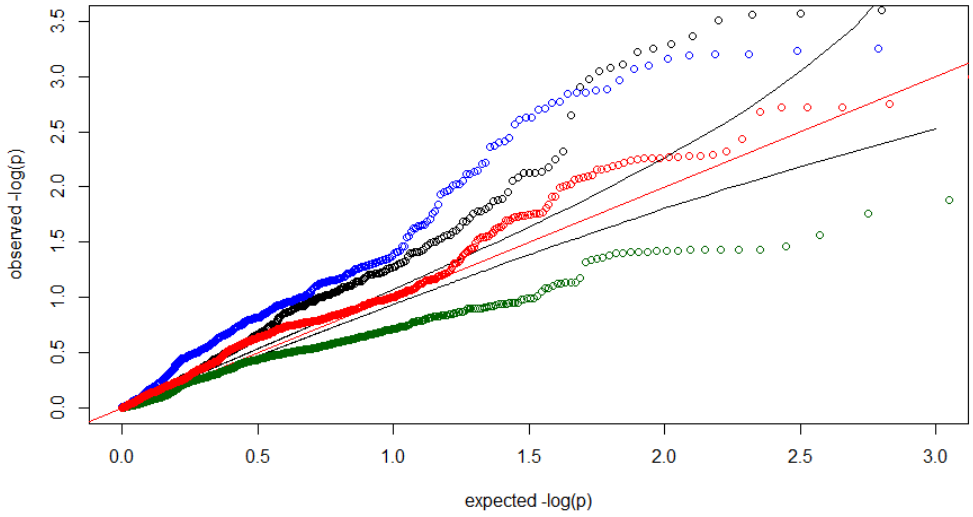


Figure 2: qq-plot for the *CTNND2* region in the PGC bipolar (dark green), Anxiety (red), Schizophrenia (blue) and MDD results (black). Expected qq values (red line) and their 95% confidence interval (black lines) are plotted for reference.

Discussion

Based on the genetic association results presented here, the association between the *CTNND2* gene and anxiety found by the rat sequencing and mapping consortium is also present in humans. We confirmed the association in a sample of patients with an anxiety disorder and controls from the Netherlands. The PGC GWAS results suggest an even broader role of *CTNND2* in psychiatric disorders, namely for schizophrenia and MDD. Our lookup in PGC results found no evidence for a role of *CTNND2* in Bipolar Disorder. The top SNPs for each of the 4 disorders are found in different areas of the CTNND2 gene. The top SNPs for MDD and Anxiety have strong LD with different SNPs in the coding regions of CTNND2. However, our strongest results are the inflated Lambda's for all SNPs in the gene. Those results currently do not point to a specific SNP within *CTNND2* but show a general association between the gene and psychiatric disorders. Further functional studies, for example based on sequence data and fine mapping are needed to see how *CTNND2* functionally relates to psychiatric phenotypes. Beyond the different types of evidence derived from association studies, there are additional studies indicating that the *CTNND2* gene functionally might be a plausible candidate. The gene is a sensor of synaptic activity and implements activity-related morphological changes at the synapse and cell adherence in adult brains(Kosik et al. 2005).[27] In developing brains delta-catenin gene expression is related to both cortical and cerebral development(Duparc et al. 2006).[28]

In addition to the results described here, there is more evidence for a role of *CTNND2* in psychiatry. Wang et al.[25] (Wang, Liu, & Aragam 2010) have also submitted associations between SNPs, in and near *CTNND2*, and bipolar disorder and schizophrenia (rs2530215, $p < 7e^{-6}$ and rs17176973, $p < 5e^{-6}$) (figure 1) to the catalogue of published GWAS studies (Hindorff et al. 2009).[29] The samples in this study were later included in the PGC mega-analyses. From the website, it cannot be derived whether this association was driven by schizophrenia cases and not by bipolar disorder cases, given the absence of an effect in the mega-analysis of bipolar disorder in PGC. Moreover, a rare CNV has been found to disrupt *CTNND2* in schizophrenia (Vrijenhoek et al. 2008)[30] and *CTNND2* hemizygozity is implicated in mental retardation and behavioral

symptoms in cri du chat syndrome (Cornish & Pigram 1996;Medina et al. 2000).[31;32]

The rat genome mapping and sequencing consortium reported little overlap between mouse and rat genomes at the gene or pathway level. This was attributed to the relatively limited amount of sequence variation segregating within the two heterogeneous stock mice populations. As a consequence, the inability to detect shared loci may result from sampling. This problem would be smaller when comparing rats and human populations as the amount of sequence variation within the human population would probably be larger. We would like to highlight that the involvement of *CTNND2* without the primary result presented by the rat genome sequencing and mapping consortium would not have stood out in a GWAS for any of these psychiatric disorders. This work can been seen as a successful synthesis between animal genetics studies and human genetics studies. Combining the results from the animal model with the replication for anxiety in humans and the previous findings in the literature provide us with ample reasons to further investigate the role of *CTNND2* in psychiatry.

References

1. Rat genome sequencing consortium. (2013) Combined sequence-based and genetic mapping analysis of complex traits in outbred rats. Nature genetics 45, 767-775

2. Israely I., Costa R.M., Xie C.W., Silva A.J., Kosik K. S., and Liu X. (2004) Deletion of the neuron-specific protein delta-catenin leads to severe cognitive and synaptic dysfunction. Current biology 14, 1657-1663.

3. Lu Q., Paredes M., Medina M., Zhou J., Cavallo R., Peifer M., Orecchio L., and Kosik K.S. (1999) Delta-Catenin, an adhesive junction associated protein which promotes cell scattering. The Journal of cell biology 144, 519-532.

4. Penninx B.W., Beekman A.T., Smit J.H., Zitman F.G., Nolen W.A., Spinhoven P., Cuijpers P. et al. (2008) The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. International Journal of Methods in Psychiatric Research 17, 121-140.

5. Lubke G. H., Hottenga J.J., Walters R., Laurin C., de Geus E.J., Willemsen G., Smit J.H., Middeldorp C.M., Penninx B.W., and Vink J.M. (2012) Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. Biological Psychiatry 72, 707–709.

6. Willemsen G., de Geus E. J.C., Bartels M., van Beijsterveldt C.T., Brooks A.I., Estourgie-van Burk G.F., Fugman D.A., Hoekstra C., Hottenga J.J., and Kluft K. (2010) The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. Twin Research and Human Genetics 13, 231.

7. Willemsen G., Vink J.M., Abdellaoui A., den Braber A., van Beek J.H.D.A., Draisma H.H.M., van Dongen J. et al. (2013) The Adult Netherlands Twin Register: Twenty-Five Years of Survey and Biological Data Collection. Twin Research and Human Genetics, 16, 271-281.

8. Ripke S., Lewis C.M., Lin D.Y., and Wray N., et al. (2012) A mega-analysis of genome-wide association studies for major depressive disorder. Molecular Psychiatry 18, 497-511.

9. Ripke S., Sanders A.R., Kendler K.S., Levinson D.F., Sklar P., Holmans P.A., Lin D.Y., Duan J., Ophoff R.A. and Andreassen O.A., et al. (2011) Genome-wide association study identifies five new schizophrenia loci. Nature genetics, 43, 969.

10. Sklar P., Ripke S., Scott L. J., Andreassen O.A., Cichon S., Craddock N., Edenberg H. J., Nurnberger J. I., Rietschel M. and Blackwood D. (2011) Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nature Genetics 43, 977-983.

11. Ter Smitten M.H., Smeets R.M.W., and Van den Brink W. (1998) Composite International Diagnostic Interview (CIDI), version 2.1. Amsterdam/Geneva: World Health Organization.

12 Boomsma D I., Sullivan P.F., de Geus E.J., Heutink P., Meijer P., Sondervan D., Kluft C. et al. (2008) Genome-wide association of major depression: description of samples for the GAIN major depressive disorder study: NTR and NESDA Biobank Projects. European Journal of Human Genetics, 16, 335-342.

13. Wilde G.J.S. (1970) Neurotische labiliteit gemeten volgens de vragenlijstmethode (The questionnaire method as a means of measuring neurotic instability). Van Rossen, Amsterdam.

14. Beck A.T., Rial W.Y., and Rickels K. (1974) Short form of depression inventory: cross-validation. Psychological Reports, 34, 1184-1186.

15. Achenbach T. M. and Rescorla L. A. (2003) Manual for the ASEBA Adult Forms and Profiles University of Vermont, Research Center for Children, Youth, & Families. Burlington, VT.

16. Van der Ploeg H, Defares P.B., and Spielberger C.D. (1979) Zelfbeoordelingsvragenslijst STAI, versie DY-1 en DY-2. Swets & Zeitlinger, Lisse.

17. Scheet P., Ehli E.A., Xiao X., van Beijsterveldt C.E., Abdellaoui A., Althoff R. R., Hottenga J. J., Willemsen G., Nelson K. A., and Huizenga P.E. (2012) Twins, Tissue, and Time: An Assessment of SNPs and CNVs. Twin Research and Human Genetics *15,* 737.

18. Sullivan P.F., de Geus E.J., Willemsen G., James M.R., Smit J.H., Zandbelt T., Arolt V., et al. (2009) Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. Molecular Psychiatry 14, 359-375.

19. Kuhn R.M., Karolchik,D., Zweig A.S., Wang T., Smith K.E., Rosenbloom K.R., Rhead B. et al. (2009) The UCSC genome browser database: update 2009. Nucleic acids research *37,* 755-761.

20. McVean G.A. et al. (2012) An integrated map of genetic variation from 1092 human genomes. Nature *491,* 56–65.

21. Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A., Bender D., Maller J., Sklar P., de Bakker P. I., Daly M.J., and Sham P.C. (2007) PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. American Journal of Human Genetics *81,* 559-575.

22. Li Y., Willer C.J., Ding J., Scheet P., and Abecasis G.R. (2010) MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genetic epidemiology. *34,* 816-834.

23. Safran M., Dalah I., Alexander J., Rosen N., Stein T.I., Shmoish M., Nativ N., Bahir I., Doniger T., and Krug H. (2010) GeneCards Version 3: the human gene integrator, Database: the journal of biological databases and curation.

24. Johnson A.D., Handsaker R.E., Pulit S L., Nizzari M.M., O'Donnell C. J., and de Bakker P. I. (2008) SNAP: a web-based tool for identification and

annotation of proxy SNPs using HapMap. Bioinformatics *24,* 2938-2939.

25. Abdellaoui A., Hottenga J.J., de Knijff P., Nivard M.G., Xiao X., Scheet P., Brooks A., Ehli E.A., Hu Y., and Davies G.E. (2013) Population structure, migration, and diversifying selection in the Netherlands. European Journal of Human Genetics. *advance online publication.*

26. Wang K.S., Liu X.F., and Aragam N. (2010) A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. Schizophrenia research *124,* 192-199.

27. Kosik K.S., Donahue C.P., Israely I., Liu X., and Ochiishi T. (2005) Delta-Catenin at the synaptic-adherens junction. Trends in Cell Biolology *15,* 172-178.

28. Duparc R.H., Boutemmine D., Champagne M. P., Treault N., and Bernier G. (2006) Pax6 is required for delta-catenin/neurojugin expression during retinal, cerebellar and cortical development in mice. Developmental biology *300,* 647-655.

29. Hindorff L. A., Sethupathy P., Junkins H.A., Ramos E.M., Mehta J.P., Collins F.S., and Manolio T.A. (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proceedings of the National Academy of Sciences 106, 9362-9367.

30. Vrijenhoek T., Buizer-Voskamp J. E., van der Stelt I., and Strengman E. (2008) Recurrent CNVs disrupt three candidate genes in schizophrenia patients, American journal of human genetics, *83,* 504.

31. Cornish K. M. and Pigram J. (1996) Developmental and behavioural characteristics of cri du chat syndrome. Archives of disease in childhood *75,* 448-450.

32. Medina M., Marinescu R.C., Overhauser J., and Kosik K.S. (2000) Hemizygosity of Delta 2 catenin (CTNND2) is associated with severe mental retardation in cri-du-chat syndrome. Genomics *63,* 157-164.

Chapter 7: Detection of gene –environment interaction in pedigree data using genome-wide genotypes

Abstract

We have developed statistical methods to model gene by environment interaction (GxE). Genome-wide single-nucleotide polymorphism (SNP)-derived genomic relationships are used to estimate the SNP heritability and the dependence of this heritability on a moderator. These methods can be applied to complex phenotypes assessed in population based samples that also include related individuals belonging to nuclear families or larger pedigrees. The method extends GxE models as implemented in GCTA, and allows the environmental exposure to be continuous, ordinal or appropriately coded categorical. Our method can be applied to multiple genetic effects concurrently, i.e. additive genetic effects as estimated based on family relations, and SNP effects. However the multiple genetic effects can also reflect different sets of SNPs based on biological pathways, genes or functional categories.

We apply these methods to genome-wide SNP data gathered in related individuals to explore GxE interaction in attention problems (AP), anxious depression (AnxDep), body mass index (BMI), and height. We consider the moderation of genetic and environmental variance components as a function of age or year of birth. Different GxE interaction models are fitted to the data for AP, AnxDep, BMI, and height. Results for all phenotypes except AnxDep reveal significant moderations of the variance explained by SNPs or familial influences by age or year of birth. These results show that the model can enhance our understanding of the genetic architecture of complex traits.

Gene-environment (GxE) interaction is an important issue in genetics, with potentially important empirical implications. In psychiatric genetics, the genetic contribution to psychiatric disorders, including attention deficit hyperactivity disorder (ADHD) and major depressive disorder (MDD), has been modeled as a function of environmental risk factors (for reviews see[1-5]). Further, GxE interaction has been show to play a role in the genetics of transcriptomics[6] and BMI.[7] Notwithstanding such results, studies of GxE interaction have been criticized for lack of statistical power,[8] poor choice of candidate genes, or markers,[9; 10] and poor replication.[11] To judge whether GxE interaction studies can inform complex trait genetics, and may play a role in explaining missing heritability,[4] more knowledge is required about the extent to which GxE interaction plays a role in explaining phenotypic variance. In addition to studies that employ genetically informative (twin) designs, [12; 13] studies including genome-wide genotype data also provide a means to evaluate GxE interaction effects.

One method involves establishing whether the effect of a polygenic risk score is moderated by the environment.[14-16] Such risk scores may be based on a weighted linear combination of SNPs that are found in a GWAS to satisfy some (not necessarily genome wide significant) alpha level. An alternative approach, which we adopted in this paper, is to model the genetic effects of a set of measured single nucleotide polymorphisms (SNPs), using genetic relatedness matrix restricted maximum likelihood (GREML), as implemented in the GCTA software package.[17] GCTA was developed to estimate 'SNP based' heritability in large groups of unrelated individuals, and has been extended to estimate GxE interaction, but only given a dichotomous environmental exposure. Vinkhuyzen & Wray[14] recently discussed the current options in GxE research based on either polygenic profile scores or GCTA methods with dichotomous exposures.

Here, we propose a model that allows the environmental exposure to be measured on a continuous, ordinal, or appropriately coded categorical scale. In addition, we extend the model to include genetic data of closely related subjects, such as twin pairs, or members of extended pedigrees. To this end, we adopted a re-parameterization of the model proposed by Zaitlen et al. [18] This allows us to evaluate GxE interaction in terms of the moderation by the environmental

exposure of the genetic variance attributable to the measured SNPs, the total additive genetic variance, and the residual (environmental) variance . We illustrate the GxE approach by analyzing the moderating role of age in attention problems, anxious depression, and body mass Index (BMI), and the moderating role of birth year in body height.

Methods

Phenotyping: Phenotypic data were collected from participants in the Netherlands Twin Register (NTR; Boomsma et al 2006) by mailed or online surveys, or during home visits. Adult participants received surveys in 10 consecutive waves over the past 25 years. Adolescent twin and their siblings received self-report questionnaires, from the age of 14 onwards. Anxious-depression (AnxDep) and attention-problems (AP) scores were obtained from the Youth or Adult Self Report[19] as part of the Achenbach system for empirical assessment (ASEBA).[19] AnxDep and AP were defined as in appendix II of this thesis, AnxDep as a sum score of all items available across multiple surveys, AP as a mean item score of all available items per survey.[20,21] Height and BMI were assessed during a home visit for the NTR biobank projects, adult (age 18 and above) height and BMI are used in the current analysis.[22] Self-reported height and BMI were analyzed if measured height and BMI were unavailable. AnxDep scores were available in 6881, AP scores in 6618, BMI in 6585, and height in 6409 individuals. We carried out a square root (i.e., normalizing) transformation of the AP and AnxDep scales. Subsequently we regressed the four phenotypes on the first 6 principal components that reflect the population structure,[23] and saved the residuals. These residuals were computed to reduce the number of fixed effects in the final modal and thus reduce computational burden. Covariates that are closely tied to the moderator M are included directly in the final model. Fixed effects covariates entered into the GxE analysis of AP, AnxDep and BMI are age, age squared, and sex. For height the covariates included were birth year, birth year squared and sex. All phenotypes, covariates, and moderators entered into the GxE analysis were standardized (zero mean, unit variance).

Genotypes: DNA samples were obtained in different projects of the Netherlands twin register (NTR).[22; 24] Genotyping in the different projects was performed on the Affymetrix 6.0 chip. SNPs that were genotyped in less than 95% of individuals were removed. Individuals with a contrast quality control (CQC) score below 0.40,

who had less than 90% of SNPs successfully genotyped, or had excess genome-wide heterozygosity /inbreeding levels (F < -.10 or F > .10) were removed. Further, individuals in whom genotyped sex did not match reported sex,whose IBD relationships were anomalous, or who did not match known pedigree structures were removed. Finally, individuals of non-European descent were excluded. The resulting sample included genotypes of 10,829 individuals. A genetic relatedness matrix (GRM) was computed on the basis of all autosomal SNPs with a minor allele frequency > 0.01, and Hardy-Weinberg Equilibrium test p-value > 1×10^{-6} using GCTA 1.24.2. [17]

Statistical Methods

Genome-wide genetic similarities given measured SNPs can be used to calculate the variance attributable to these measured SNPs (for the derivation, see [17]). We convey this model in the equations below. In Equation 1a, let *y* (nx1) be a multivariate random vector of phenotypic scores as observed in n individuals, let *X* (nxm) be the matrix of fixed covariates, and β (mx1) the vector of fixed effects. Let matrix *W* (nxp) be the matrix of *p* standardized SNPs, *u* (px1) a zero mean vector of random effects, and let *e* be the nx1 vector of zero mean residuals. The phenotype *Y* is a random multivariate normal vector with mean vector μ and covariance matrix *V*. The GRM is a matrix of pair wise genetic similarities computed WW'/p. The parameter σ^2_{snp} is an estimate of the variance explained by the SNPs tagged by, or included in *W*.

$$Y = X\beta + Wu + e \qquad (1a)$$

$$Y \sim N(\mu, V) \qquad (1b)$$

$$V = GRM_{n*n} \otimes \sigma^2_{snp} + I_{n*n} \otimes \sigma^2_e \qquad (1c)$$

$$\mu = X\beta \qquad (1d)$$

The GCTA software fits the model above, with extensions allowing for dichotomous GxE moderation, among other options.[17]

The continuous GxE model. The continuous GxE model presented here allows the effect of all SNPs on a trait *Y* (σ^2_{snp}) to vary with respect to a moderator M. Parameter β_g quantifies the effect of moderation by *M* of the genetic effects on *Y*. Parameter σ_{snp} quantifies the effect of measured SNPs on *Y* given $\beta_g = 0$. Parameter β_e quantifies the effect of moderator *M* on the residual variance of *Y*. Parameter σ_e quantifies the effect of residual variance given $\beta_e = 0$.

$$V(Y|M) = V = GRM_{n*n} \otimes \left(\sigma_{snp} + \beta_g * M\right)^2 + I_{n*n} \otimes (\sigma_e + \beta_e * M)^2 \quad (2)$$

This model includes interaction, i.e., moderation of the genetic and residual variances by *M*.[12] Note that we assume that the moderator *M* is also included in the matrix *X*, i.e., as a fixed covariate with a main effect on the phenotype. We have presented the moderator *M* as continuous, but it may be discrete (interval, or appropriately coded nominal, or binary).

Related individuals in the sample.

A recent extension of the model as implemented in GCTA, proposed by Zaitlen, et al. [18] allows for the estimation of variance explained by SNPs, as well as the total additive genetic variance of a trait. This extension allows for in the inclusion in the sample of closely related individuals. In Eq 3 below, the matrix GRM^{IBS} is equivalent to the *GRM* in Eq 1 and 2, but now includes closely related individuals. The matrix $GRM^{IBS > 0.05}$ equals the matrix GRM^{IBS} in which all relatedness coefficients below .05 set to zero. Note that the values of these coefficients in closely related individuals tend towards the expected proportion of alleles shared identically by decent (~IBD; we denote the expected proportion pi-hat) (i.e., full siblings are characterized by pi-hat = .5 IBD, and ~.5 in *the* $GRM^{IBS > 0.05}$). Using an IBD matrix or IBS > 0.05 matrix yields very similar results.[18]

$$V = GRM^{IBS}_{n*n} \otimes \sigma^2_{snp} + GRM^{IBS>0.05}_{n*n} \otimes (\sigma^2_{\sim IBD} - \sigma^2_{snp}) + I_{n*n} \otimes \sigma^2_e \quad (3)$$

In Eq 3, parameter σ^2_{snp} reflects the variance explained by SNPs, the term $\left(\sigma^2_{\sim IBD} - \sigma^2_{snp}\right)$ represents the difference between the total additive genetic variance $\sigma^2_{\sim IBD}$ and the variance explained by SNPs, and σ^2_e reflects the variance attributable to residual effects. Inspection of the parameter correlation matrix derived from the Hessian revealed very strong negative parameter correlations between σ^2_{snp} and $\sigma^2_{\sim IBD} - \sigma^2_{snp}$, which complicates any moderation of these terms. To ensure low parameter correlations, and allow for separate moderation of σ^2_{snp} and $\sigma^2_{\sim IBD}$, we re-parameterized the Zaitlen model (Equation 3) as shown in Equation 4.

$$V = GRM^{IBS<0.05}_{n*n} \otimes \sigma^2_{snp} + GRM^{IBS>0.05}_{n*n} \otimes \sigma^2_{\sim IBD} + I_{n*n} \otimes \sigma^2_e \quad (4)$$

In Equation 4, the first GRM $GRM^{IBS<0.05}_{n*n}$includes only values < 0.05 where other values including the diagonal elements are set to 0. This provides an estimate of variance attributable to SNPs exclusively based on the covariance between distantly related individuals. The second GRM, $GRM^{IBS>0.05}_{n*n}$ contains only values above 0.05, it reflects all genetic variance as a function of approximate IBD. Note that this model requires the presence of closely related individuals to reliably estimateσ^2_{IBD} . The re-parameterized model (Equation 4) and the Zaitlen model (Eq 3) produce the same estimates of $\sigma^2_{\sim IBD}, \sigma^2_{snp}$, and σ^2_e. This equivalence between the models was established empirically by simulating data for a wide range of $\sigma^2_{\sim IBD}, \sigma^2_{snp}$, and σ^2_e, under the Zaitlen model, and subsequently fitting both models, and obtaining the same -2*log-likelihood and parameter estimates (table S1). We note that in the unlikely scenario that $\sigma^2_{snp} = 0$ or ($\sigma^2_{\sim IBD} - \sigma^2_{snp}$) = 0, the equivalence does not hold. However, if $\sigma^2_{snp} = 0$ or ($\sigma^2_{\sim IBD} - \sigma^2_{snp}$) = 0 is true, separate moderation of σ^2_{snp} or $\sigma^2_{\sim IBD}$ is undesirable given the known absence of the variance component to be moderated.

We proceeded to extend Eq 4 to include moderation, as shown in Eq 4a. Parameter $\beta_{\sim IBD}$ in Eq 4a reflects the change in additive genetic variance as a function of the moderator M. Parameter β_{snp} reflects moderation of genetic variance attributable to SNPs as a function of *M*.

$$V(Y|M) = GRM_{n*n}^{IBS<0.05} \otimes \left(\sigma_{snp} + \beta_{snp} * M\right)^2 + GRM_{n*n}^{IBS>0.05} \otimes \left(\sigma_{\sim IBD} + \beta_{\sim IBD} * M\right)^2 + I_{n*n} \otimes \left(\sigma_e + \beta_e * M\right)^2 \quad (4a)$$

Model 4a requires a sufficient number of related individuals to be in the sample to adequately estimate $\sigma^2_{\sim IBD}$. We consider several possible outcomes (assuming for convenience $\beta_e = 0$):

1. Significant β_{snp} and $\beta_{\sim IBD}$ of same sign and equal magnitude implies the variance explained by genes is moderated by M. This implies a change in genetic variance in a trait with change in moderator *M*.
2. The parameter β_{snp} is zero and $\beta_{\sim IBD}$ is not zero implies the variance explained by additive genetics changes while the variance explained by SNPs remains constant. This could be due to a lack of power to detect moderation of the genetic effects attributable to SNPs.
3. Both β_{snp} and $\beta_{\sim IBD}$ are zero implies the genetic effects on the phenotype are not moderated by M.
4. Non-zero β_{snp} but zero $\beta_{\sim IBD}$ implies that the proportion of genetic variance explained by SNP changes with the moderator, but the variance explained by genes does not. If the SNPs measured are a random sample of the common genetic variance, we consider it unlikely their effect is moderated while the total additive genetic variance is not. We do not therefore explicitly test for this fourth scenario.

Finally all outcomes above could be observed in the presence of moderation of the residual variance ($\beta_e \neq 0$). If moderation is limited to this residual variance, this will necessarily result in a decrease or increase in the heritability of a trait.

Model estimation

All models were fitted in R[25] using full information maximum likelihood (FIML) optimization with exact derivatives. The implementation in R includes two FORTRAN routines to speed up calculation of the likelihood function and derivatives. Optimization using exact derivatives is done using the optim() function available in R. Scripts to perform the optimization and the required

FORTRAN routines are available online (URL). We tested the significance of parameters by means of the likelihood ratio test. We adopted an alpha of .05 for each test.

Standard practice in implementing GCTA is to exclude genetically closely related individuals to avoid confounding of the total heritability and the SNP heritability.[17; 26] Closely related in our analysis is defined as genetic relatedness greater than 0.05, as in Zaitlen et al.[18]

Results

Heritability and variance explained by SNPs. The total and SNP related genetic effects (Eq 3) on AP, AnxDep, BMI and Height are given in Table 1. All estimates are significant (p-values <0.05) except the variance explained by SNPs in AnxDep (p = 0.078, chisq = 3.11, df = 1). The total heritability is 41.6% (s.e. 2.0%) for AP and 40.6% (s.e 2.0%) for AnxDep, the variance explained by all SNPs is and 11.4% (s.e. 5.8%) for AP and 9.8%. (s.e. 5.7%) for AnxDep. In contrast, the total heritability of BMI is 75.3% (s.e. =1.3%) and for height 91.3% (s.e. 0.4%). The variance explained by all SNPs for BMI is 41.6% (s.e 6.4%) and for height 53.8%(s.e. 6.3%). Consequently, for AP and AnxDep SNPs explain 27.5% (s.e. 14.1%) and 24.3% (s.e. 14.1%) of the *genetic* variance and for BMI and height 56.2% (s.e. 8.6%) and 59% (s.e. 6.9%).

As the genetic variance explained by SNPs for BMI was higher than reported in the literature,[27] we repeated the analysis for BMI in GCTA, using data of the distantly related individuals only. The analysis carried out in GCTA was based on 3119 nominally unrelated individuals and resulted in a heritability estimate of 49% (s.e. 11.4%). Analysis carried out in GCTA based on 6395 individuals (closely related and nominally unrelated) resulted in a heritability estimate of 75.4% (s.e. 1.3%). These results are very close to those obtained with our method.

	AnxDep	AP	BMI	Height
Proportion of **phenotypic** variance explained by SNPs	9.8%% n.s. (se = 5.7%)	11.4%* (se 5.8%)	41.6%*** (se= 6.4%)	53.8%*** (se=6.3%)
Proportion of **genetic** variance explained by SNPs	24.3% (se = 14.1%)	27.5% (se = 14.1%)	56.2% (se = 8.6%)	58.9% (se= 6.9%)
Proportion of **phenotypic** variance explained by additive genetic influences	40.6%*** (se = 2.0%)	41.6%*** (se = 2.0%)	75.3% *** (se = 1.3%)	91.3%*** (se = 0.4%)
Proportion of **phenotypic** variance explained by residual influences	59.4% (se = 2.0 %)	56.8% (se = 2.0%)	24.8% (se = 1.3%)	9.7% (se = 0.4%)

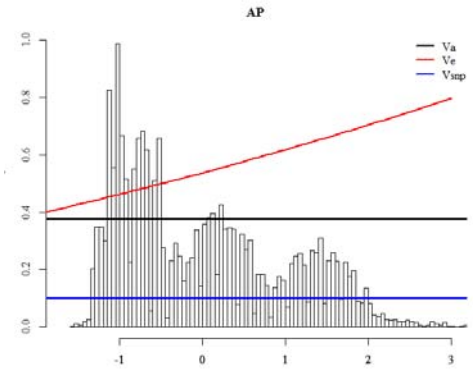
Table 1: estimates of the proportion of variance attributable to SNPs, familial genetic effects and the environment for AnxDep, AP, BMI and Height. Significance determined by likelihood ratio testing. Standard errors for the different ratios approximated using the delta rule. * $P < 0.05$, *** $P < 0.0001$,

Gene by Environment interaction. Next we fitted the model given in eq 4a to all four variables with age or birth year as the moderator of the variance components (σ^2_{snp}, $\sigma^2_{\sim IBD}$ and σ^2_e). Moderation of the variance components attributable to SNP effects was not significant for any of the phenotypes (Table 2). For AP, BMI, and height, moderation of the residual effects was significant as was moderation of the total genetic effects in BMI. In AnxDep, these effects were not significant.

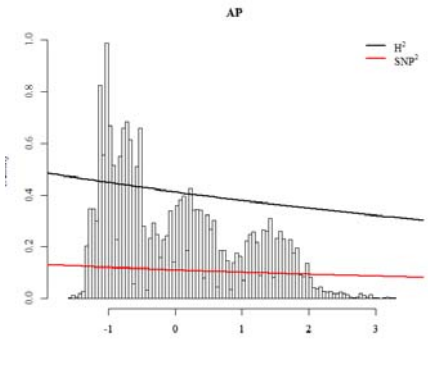
Table 2: Parameter estimates, -2 log likelihoods and significance tests for the GxE models fitted for AP, AnxDep, BMI and Height. σ indicates intercepts , i.e standard error at moderator = 0, β indicate magnitude of change given a standard deviation change in the age moderator.

AP									
	σ_{snp}	β_{snp}	σ_{IBD}	β_{IBD}	σ_e	β_e	-2ll	Δ-df	Likelihood ratio
Full-Moderation	.32	.096	.62	-.002	.73	.054	17830.78	-	-
Drop β_{snp}	.32	-	.61	-.003	.74	.054	17831.78	1	1
Drop β_{IBD}&β_{snp}	.32		.61		.74	.053	17831.82	1	0.04
Dropβ_{snp}, β_{IBD}& β_e	.32		.62		.73		17859.96	1	28.14
AnxDep									
	σ_{snp}	β_{snp}	σ_{IBD}	β_{IBD}	σ_e	β_e	-2ll	Δ-df	Likelihood ratio
Full-Moderation	.27	-0.20	.63	.03	.76	-.005	18977.33	-	-
Drop β_{snp}	.31		.63	.03	.76	-.004	18978.76	1	1.43
Drop β_{IBD}&β_{snp}	.31	-	.63	-	.76	.012	18981.273	1	2.51
Dropβ_{snp}, β_{IBD}& β_e	-	-	.64	-	.77	-	18982.684	1	1.39
BMI									
	σ_{snp}	β_{snp}	σ_{IBD}	β_{IBD}	σ_e	β_e	-2ll	Δ-df	Likelihood ratio
Full-Moderation	.62	.048	.80	.045	.48	.124	16480.76		
Drop β_{snp}	.61		.80	.043	.50	.125	16482.06	1	1.30
Drop β_{IBD}&β_{snp}	.61		.78		.52	.148	16491.55	1	9.50
Dropβ_{snp}, β_{IBD}& β_e	.61	-	.82	-	.48		16684.90	1	193.35
Height									
	σ_{snp}	β_{snp}	σ_{IBD}	β_{IBD}	σ_e	β_e	-2ll	Δ-df	Likelihood ratio
Full-Moderation	.53	0.047	.69	0.005	.21	0.02	11132.06	-	-
Drop β_{snp}	.53	-	.69	0.003	.21	0.02	11134.03	1	1.97
Drop β_{IBD}&β_{snp}	.52	-	.69	-	.21	0.02	11134.29	1	0.26
Dropβ_{snp}, β_{IBD}& β_e	.53	-	.68	-	.21	-	11154.57	1	20.28

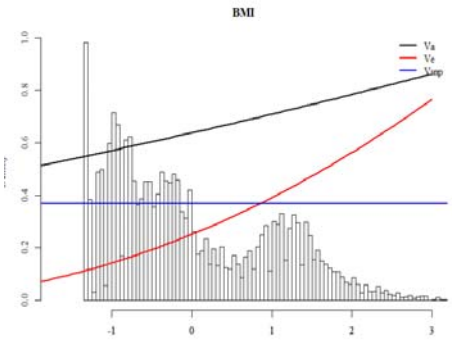
1a



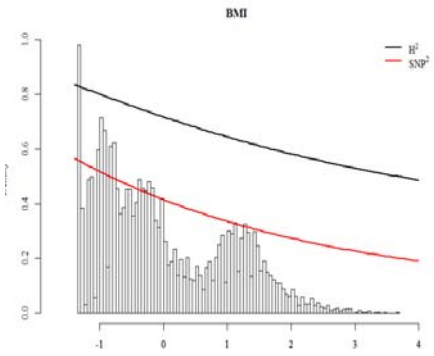
1b



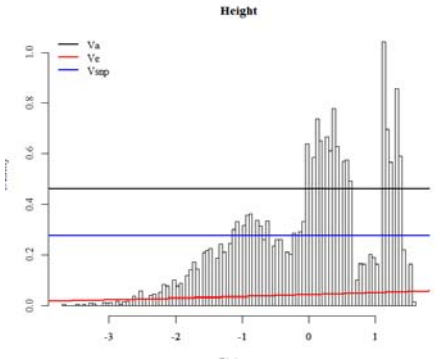
2a



2b



3a



3b

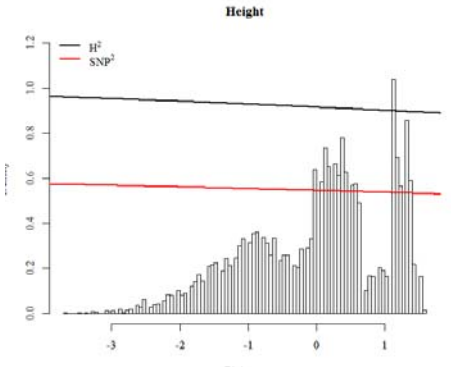


Figure 1: A. σ^2_{snp}, $\sigma^2_{\sim IBD}$, and , σ^2_e in AP as a function of age. B. the heritability($\frac{\sigma^2_{\sim IBD}}{\sigma^2_{\sim IBD} + \sigma^2_e}$) and proportion of phenotypic variance attributable to SNPs ($\frac{\sigma^2_{snp}}{\sigma^2_{\sim IBD} + \sigma^2_e}$) in AP as a function of age. A histogram visualized the distribution of the standardized moderator age.

Figure 2: A: σ^2_{snp}, $\sigma^2_{\sim IBD}$, and , σ^2_e in BMI as a function of age. B: the heritability($\frac{\sigma^2_{\sim IBD}}{\sigma^2_{\sim IBD} + \sigma^2_e}$) and proportion of phenotypic variance attributable to SNPs ($\frac{\sigma^2_{snp}}{\sigma^2_{\sim IBD} + \sigma^2_e}$) in BMI as a function of age. A histogram visualized the distribution of the standardized moderator age.

Figure 3: A: σ^2_{snp}, $\sigma^2_{\sim IBD}$, and , σ^2_e in height as a function of birthyear. B: the heritability($\frac{\sigma^2_{\sim IBD}}{\sigma^2_{\sim IBD} + \sigma^2_e}$) and proportion of phenotypic variance attributable to SNPs ($\frac{\sigma^2_{snp}}{\sigma^2_{\sim IBD} + \sigma^2_e}$) in height as a function of birth year. A histogram visualized the distribution of the standardized moderator: birth year.

Discussion

We presented a model to estimate the moderation of SNP genetic, total additive genetic, and residual effects by (continuous) moderators. The model can be used in samples of unrelated individuals to moderate the genetic variance attributable to SNPs. In addition, the model can be used in pedigree data to separately moderate SNP genetic effects and total additive genetic effects. The GxE interaction models fitted in this paper show differences in genetic architecture between attention problems, anxious depression, height, and BMI that are not limited to differences in $\sigma^2_{\sim IBD}$, σ^2_{snp} and σ^2_e between traits. Our results indicate that for anxious depression, none of the genetic and non-genetic effects are moderated by age,while for AP, BMI, and height the residual variance increases with age or birth year. For BMI, there is also a positive moderation effect of the additive genetic variance by age. We found no evidence of moderation of σ^2_{snp} in any of the phenotypes considered.

The findings for AnxDep and AP are in line with previous findings in longitudinal twin studies.[20; 21] Note that where we find an increase in σ^2_e for AP with time, Kan et al.[20] do not. Differences in study design, and estimated parameters between Kan et al.[20] and the current work could explain the different results. The findings for BMI (decreasing heritability with age) are in line with a recent meta-analysis into the heritability of BMI.[28] Other recent studies into the genetics of BMI have found evidence for GxE at the SNP level. Rosenquist et al.[7] reported evidence for an interaction between the FTO gene and birth cohort on BMI.

The results of fitting the Zaitlen model to AP, AnxDep, BMI and height confirmed previously observed differential genetic architectures for different disorders. In line with previous twin studies, both AP and AnxDep are moderately heritable,[20; 21] while BMI and height are strongly heritable.[28; 29] The pattern for the SNP heritability estimates also agreed with the picture emerging from other studies with lower SNP heritability estimates for quantitative psychiatric traits than for other psychiatric and non-psychiatric traits.[30] The relatively low SNP heritability estimates found for AP and AnxDep are consistent with the presence of disease heterogeneity.[31] In contrast, a substantial part of *phenotypic* variance of height and BMI is explained by SNPs. For height, our estimate of variance explained by SNPs (53.8%) was comparable to the estimate reported by the GIANT consortium

(49.8%, Wood, et al 2014[32]). The only unexpected result was the higher estimate for BMI (SNP heritability 41.6%), which is higher than those of previous studies (17%, Yang et al 2011[33]; 23%, Zaitlen et al. 2013[18]). We validated our BMI result by running GCTA on the maximum number of unrelated in the sample and on all data, resulting in comparable estimates. These results suggest the higher SNP heritability is not a function of the software but of the sample. A review of the heritability of BMI by Elks et al.[28] found the type of measurement of BMI (measured vs. self report) and age of the participants to influence the heritability. As our sample has many young individuals and is almost completely based on measured BMI and these factors are associate with a higher heritibillity by Elks et al this could play a role.

The method presented here has some limitations. Under certain conditions the GxE interaction for genetic variance as a function of IBD can result in false positives. If the current method is applied in genotype data collected in twins and/or families, we recommend the user to determine whether the relationship between moderator and phenotype is of the kind van der Sluis et al.[34] find to inflate the false positive rate. When carefully considering the relationship between moderator and phenotype and making plausible assumptions, this model will perform as expected. Any GxE model is sensitive to scaling[35]. The GxE interaction terms found here are conditional of the scale of the variable; different scaling may yield different results. Where BMI and height have a definite scale (i.e., kg/m^2 and m or cm), the scale of the data based on psychiatric questionnaires is generally arbitrary. This problem is however not limited to the current model and its solution is beyond the scope of this article (potential solutions are discussed elsewhere[13; 35]).

Besides the flexibility of this method to include unrelated as well as related individuals and the inclusion of a continuous moderator, specific tests that were already possible in GCTA can still be applied. Generally the genetic effects attributable to SNPs are assumed to be a random subset of all additive genetic variance. Moderation of the variance attributable to SNPs in absence of moderation of the additive genetic variance is therefore not expected. A result suggesting moderation of the variance attributable to SNPs in the absence of moderation of the additive genetics variance should be treated as anomalous. However, the model does allow for separate testing of the moderation of variance attributable to SNPs. Hypotheses that explicitly formulate the expectation that

variance attributable to SNPs is moderated , while the additive genetic variance is not, can be tested. This type of hypothesis test can be interesting in the context of disease heterogeneity.[31] In the presence of disease heterogeneity the explicit expectation is that of low SNP heritability while additive genetic variance remains high.

The addition of multiple genetic effects (GRMs) further allows for the separate moderation of different subsets of the genome. One could for example limit the SNPs in a GRM to a single biological pathway (i.e. SNPs in genes in the serotonin pathway), to a single class of SNPs (i.e. coding variants), or to specific regions of the genome (i.e. regulatory elements, the exome, etc.).The moderator can be a genetic variant (GxG i.e. a known risk variant), biological (i.e. a gene expression; gut microbiota) or environmental (i.e. early childhood trauma experiences).

The current application of the model revealed differences in genetic architecture between attention problems, symptoms of anxiety and depression, BMI and height. The model shows promise in being able to reveal biologically informative GxE interactions.

References

1. Caspi A., Hariri A.R., Holmes A., Uher R., and Moffitt T.E. (2010) Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. American journal of Psychiatry *167*, 509-527

2. Carpenter R.W., Tomko R.L., Trull T.J., and Boomsma D.I. (2013) Gene-environment studies and borderline personality disorder: a review. Current psychiatry reports *15*, 1-7

3. Dunn E.C., Uddin M., Subramanian S.V., Smoller J.W., Galea S., and Koenen K.C. (2011) Research Review: Gene–environment interaction research in youth depression–a systematic review with recommendations for future research. Journal of Child Psychology and Psychiatry *52*, 1223-1238

4. Uher R. (2014) Gene–environment interactions in common mental disorders: an update and strategy for a genome-wide search. Social psychiatry and psychiatric epidemiology *49*, 3-14

5. Young-Wolff K.C., Enoch M.A., and Prescott C.A. (2011) The influence of gene–environment interactions on alcohol consumption and alcohol use disorders: a comprehensive review. Clinical psychology review *31*, 800-816

6. Buil A., Brown A.A., Lappalainen T., Vinuela A., Davies M.N., Zheng H.F., Richards J.B. et al. (2015) Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. Nature genetics *47*, 88-91

7. Rosenquist J.N., Lehrer S.F., O'Malley A.J., Zaslavsky A.M., Smoller J.W., and Christakis N.A. (2014) Cohort of birth modifies the association between FTO genotype and BMI. Proceedings of the National Academy of Sciences, 201411893

8. Munafo M.R., Flint J. (2009) Replication and heterogeneity in genex environment interaction studies. The International Journal of Neuropsychopharmacology *12*, 727-729

9. Bosker F.J., Hartman C.A., Nolte I.M., Prins B.P., Terpstra P., Posthuma D., Van Veen T. et al. (2010) Poor replication of candidate genes for major depressive disorder using genome-wide association data. Molecular psychiatry *16*, 516-532

10. Sullivan P.F., Daly M.J., and O'Donovan M. (2012) Genetic architectures of psychiatric disorders: the emerging picture and its implications. Nature Reviews Genetics *13*, 537-551

11. Duncan L.E., Keller M.C. (2011) A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. American journal of Psychiatry *168*, 1041-1049

12. Purcell S. (2002) Variance components models for gene–environment interaction in twin analysis. Twin research *5*, 554-571

13. Molenaar D., Dolan C.V. (2014) Testing systematic genotype by environment interactions using item level data. Behavior genetics *44*, 212-231

14. Vinkhuyzen A.A.E., Wray N.R. (2014) Novel directions for Gx E analysis in psychiatry. Epidemiology and psychiatric sciences, 1-8

15. Wray N.R., Yang J., Hayes B.J., Price A.L., Goddard M.E., and Visscher P.M. (2013) Pitfalls of predicting complex traits from SNPs. Nat Rev Genet *14*, 507-515

16. Peyrot W.J., Milaneschi Y., Abdellaoui A., Sullivan P.F., Hottenga J.J., Boomsma D.I., and Penninx B.W. (2014) Effect of polygenic risk scores on depression in childhood trauma. The British Journal of Psychiatry *205*, 113-119

17. Yang J., Lee S.H., Goddard M.E., and Visscher P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics *88*, 76-82

18. Zaitlen N., Kraft P., Patterson N., Pasaniuc B., Bhatia G., Pollack S., and Price A.L. (2013) Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. PLoS genetics *9*, e1003520

19. Achenbach T.M., Rescorla L.A., McConaughey S., Pecora P., Wetherbee K., and Ruffle T. (2003) The Achenbach system of empirically based assessment. Handbook of psychological and educational assessment of children: Personality, behavior, and context *2*, 406-432

20. Kan K.J., Dolan C.V., Nivard M.G., Middeldorp C.M., van Beijsterveldt C.E., Willemsen G., and Boomsma D.I. (2013) Genetic and environmental stability in attention problems across the lifespan: evidence from the Netherlands twin register. Journal of the American Academy of Child & Adolescent Psychiatry *52*, 12-25

21. Nivard M.G., Dolan C.V., Kendler K.S., Kan K.J., Willemsen G., van Beijsterveldt C.E., Lindauer R.J. et al. (2014) Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. Psychological medicine, 1-11

22. Willemsen G., Vink J.M., Abdellaoui A., den Braber A., van Beek J.H., Draisma H.H., van Dongen J. et al. (2013) The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. Twin Research and Human Genetics *16*, 271-281

23. Abdellaoui A., Hottenga J.J., de Knijff P., Nivard M.G., Xiao X., Scheet P., Brooks A. et al. (2013) Population structure, migration, and diversifying selection in the Netherlands. European journal of human genetics *21*, 1277-1285

24. Willemsen G., De Geus E.J., Bartels M., Van Beijsterveldt C.E.M., Brooks A.I., Estourgie-van Burk G.F., Fugman D.A. et al. (2010) The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. Twin Research and Human Genetics *13*, 231-245

25. R Core Team (2013)
R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

26. Lee J.J., Chow C.C. (2014) Conditions for the validity of SNP-based heritability estimation. Human genetics, 1-12

27. Yang J., Manolio T.A., Pasquale L.R., Boerwinkle E., Caporaso N., Cunningham J.M., de Andrade M. et al. (2011) Genome partitioning of genetic variation for complex traits using common SNPs. Nature genetics *43*, 519-525

28. Elks C.E., Den Hoed M., Zhao J.H., Sharp S.J., Wareham N.J., Loos R.J., and Ong K.K. (2012) Variability in the heritability of body mass index: a systematic review and meta-regression. Frontiers in endocrinology *3*

29. Silventoinen K., Sammalisto S., Perola M., Boomsma D.I., Cornes B.K., Davis C., Dunkel L. et al. (2003) Heritability of adult body height: a comparative study of twin cohorts in eight countries. Twin research *6*, 399-408

30. Wray N.R., Lee S.H., Mehta D., Vinkhuyzen A.A., Dudbridge F., and Middeldorp C.M. (2014) Research Review: Polygenic methods and their application to psychiatric traits. Journal of Child Psychology and Psychiatry *55*, 1068-1087

31. Wray N.R., Maier R. (2014) Genetic Basis of Complex Genetic Disease: The Contribution of Disease Heterogeneity to Missing Heritability. Current Epidemiology Reports *1*, 220-227

32. Wood A.R., Esko T., Yang J., Vedantam S., Pers T.H., Gustafsson S., Chu A.Y. et al. (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. Nature genetics *46*, 1173-1186

33. Yang J., Manolio T.A., Pasquale L.R., Boerwinkle E., Caporaso N., Cunningham J.M., de Andrade M. et al. (2011) Genome partitioning of genetic variation for complex traits using common SNPs. Nature genetics *43*, 519-525

34. van der Sluis S., Posthuma D., and Dolan C.V. (2012) A note on false positives and power in GxE modelling of twin data. Behavior genetics *42*, 170-186

35. Eaves L., Verhulst B. (2014) Problems and Pit-Falls in Testing for Gx E and Epistasis in Candidate Gene Studies of Human Behavior. Behavior genetics *44*, 578-590

table S1: Parameter estimates and log likelihood for the Zaitlen model and the model in eq 4 given a range of simulated values for σ^2_{snp}, $\sigma^2_{\sim\text{IBD}}$, and , σ^2_{e} . Table reveals the models converge to the same *-2LL and same estimate of* σ^2_{snp}.

σ^2_{snp},	$\sigma^2_{\sim\text{IBD}}$ - σ^2_{snp},	$\sigma^2_{\sim\text{IBD}}$	σ^2_{e}	Zaitlen LL	Nivard LL	Simulated h2 snp	Estimated h2snp Zaitlen	Estimated h2snp Nivard
0.1	0.9	1	1	17681.27544	17681.2754	0.05	8.62306E-05	1.42263E-06
0.2	0.8	1	1	17629.37754	17629.37753	0.1	0.103484528	0.103214272
0.3	0.7	1	1	17656.17652	17656.17652	0.15	0.233336336	0.23338012
0.4	0.6	1	1	17570.96095	17570.96095	0.2	0.306630753	0.306637925
0.5	0.5	1	1	17651.7169	17651.71689	0.25	0.307512002	0.307599153
0.6	0.4	1	1	17625.01804	17625.01804	0.3	0.334809229	0.334829735
0.7	0.3	1	1	17603.86846	17603.86846	0.35	0.306268825	0.30627982
0.8	0.2	1	1	17624.51335	17624.51335	0.4	0.373124729	0.373106123
0.1	0.9	1	0.5	17145.9007	17145.9007	0.0666667	5.44948E-10	1.6286E-09
0.2	0.8	1	0.5	17094.9891	17094.9891	0.1333333	0.06232238	0.062400028
0.3	0.7	1	0.5	17114.8934	17114.8934	0.2	0.211478331	0.211555235
0.4	0.6	1	0.5	17116.72428	17116.72428	0.26666667	0.419951483	0.419968402
0.5	0.5	1	0.5	17102.29316	17102.29316	0.33333333	0.293760135	0.293758324
0.6	0.4	1	0.5	17142.75462	17142.75462	0.4	0.37675785	0.376759459
0.7	0.3	1	0.5	16996.54725	16996.54725	0.46666667	0.516114696	0.51612226
0.8	0.2	1	0.5	17128.18242	17128.18241	0.53333333	0.489751374	0.489637655
0.1	0.9	1	2	17938.80008	17938.80008	0.03333333	1.64593E-09	7.09134E-10
0.2	0.8	1	2	17911.46342	17911.4635	0.06666667	8.65775E-10	5.13991E-06
0.3	0.7	1	2	17949.52296	17949.52296	0.1	0.158529894	0.158487064
0.4	0.6	1	2	17907.82245	17907.82245	0.133333	0.089440064	0.089457644
0.5	0.5	1	2	17862.7044	17862.7044	0.166666667	0.146489071	0.146528413
0.6	0.4	1	2	17938.71755	17938.71755	0.2	0.275167092	0.275169226
0.7	0.3	1	2	17923.36645	17923.36644	0.233333333	0.236952853	0.237019414
0.8	0.2	1	2	17851.47076	17851.47076	0.266666667	0.320172587	0.320173591

Chapter 8: Rapid genetic (co)variance component estimation.

Abstract:

Estimating the genetic (co)variance between traits based on genotyped markers can reveal biologically insightful relations. However, procedures used to estimate genetic covariance are computationally intensive. As sample sizes and the number of traits increase, multivariate analysis becomes progressively less tractable. We propose a method that reduces multivariate (co)variance estimation into a series of univariate variance estimations. The method is based on the observation that the variance of the sum of two random variables equals the sum of their variances, plus twice their covariance. Our procedure drastically reduces the dimensions of the matrix that is repeatedly inverted. We show that we can obtain unbiased estimates of the genetic covariance between traits on the basis of a series of genetic variance decompositions of the individual variables and the sum of each pair of variables.

Our method allows for the concurrent modeling of multiple genetic covariance matrices. Separate covariance matrices can be estimated for different genetic pathways, genes, or functional categories of markers. The possibility of modeling multiple genetic covariance matrices allows for the estimation of the covariance attributable to the effect of genotyped markers in samples containing closely related subjects.

We applied the procedure to item data on extraversion and neuroticism. Sum scores for these personality traits tend to have relatively low SNP heritabilities, while heritability as estimated from twin data is moderately high, Here, an estimate of the genetic covariance matrix of a set of 24 items measuring extraversion and neuroticism was obtained.

Introduction.

Genetic relationship matrix restricted maximum likelihood (GREML) is used to estimate the variance in traits and diseases that is attributable to, or tagged by, common genetic variants[1]. The method, as implemented in the software Genetic Complex Trait Analysis (GCTA), has been extended to allow for estimation of the bivariate genetic covariance between two diseases or traits[2]. The software suit GEMMA extends this further and implements multivariate (co)variance estimation in the context of linear multivariate mixed modelling[3]. While high dimensional multivariate extension of GREML is conceptually easy, it is computationally heavy. We propose a method to estimate a high dimensional genetic covariance matrix by reducing the multivariate problem to a series of univariate problems. As these univariate problems can be executed independently, this method is suitable for parallelization.

Univariate GREML estimates the effects of all measured genotypes as random effects in a mixed linear model (equation 1).

$$y = X\mathrm{b} + W\mathrm{u} + e \qquad (1)$$

In this model: *y* is an (nx1) multivariate normal vector of observed phenotypes, with n being the sample size; *X* is an (nxm) matrix of random variables, i.e. fixed covariates, where m is the number of fixed covariates. *b* (mx1) is a vector of fixed effects. W is a matrix of (nxp) standardized genetic variants, where p is the number of genetic variants. *u* (px1) is a vector of zero mean random genetic effects, e (nx1) is a random normal vector of zero mean residual values. The genetic variants are usually single nucleotide polymorphisms (SNP) coded 0, 1, 2. If S_{ik} is the value of SNP for individual i then the standardized genetic value Z_{ik} equals $(S_{ik} - 2 * F_k) / \sqrt{2F_k * (1 - F_k)}$ where F_k is the minor allele frequency of SNP *k*.

Yang et al. (2011) model the covariance matrix (*V*) of the multivariate normal vector *y* as:

$$y \sim N(\mu, V)$$

$$V_{(\mathrm{nxn})} = A_{(\mathrm{nxn})} \otimes \sigma_g^2 + I_{(\mathrm{nxn})} \otimes \sigma_e^2 \qquad (2),$$

where *A* is an nxn matrix of genetic similarities between subjects defined as *WW'/p,* and σ^2_g as the genetic variance explained by all standardized genetic variants in *W*. This operationalization of σ^2_g has been shown to be valid under a wide array of conditions[4].

Estimation of σ^2_g and σ^2_e by means of (restricted) maximum likelihood requires repeated inversion of *V*. As the dimensions of *V* may be large (e.g., > 5000) repeated inversion is computationally burdensome. The computational burden is further increased if the phenotype of interest is multivariate, including k phenotypes. Letting $\Sigma_{\mathrm{g\ (kxk)}}$ denote the genetic covariance matrix and $\Sigma_{\mathrm{e\ (kxk)}}$ the environmental covariance matrix, the dimensions of V increase to $(\mathrm{k} * \mathrm{n})\ \mathrm{x}\ (\mathrm{k} * \mathrm{n})$.

$$V_{((\mathrm{k}*\mathrm{n})\ \mathrm{x}\ (\mathrm{k}*\mathrm{n}))} = A_{(\mathrm{nxn})} \otimes \Sigma_{\mathrm{g\ (kxk)}} + \mathrm{I}_{(\mathrm{nxn})} \otimes \Sigma_{\mathrm{e\ (kxk)}} \qquad (3)$$

Multivariate GREML optimization may become intractable with the increase in dimension n (subjects) and dimension k (phenotypes). One solution is to diagonalize *A* using eigenvalue decomposition. This renders the matrix to be inverted diagonal, which is computationally simple[5; 6]. This solution is implemented in the GEMMA software suite[3]. However this solution is limited to models in which a single genetic covariance matrix is estimated. Kostem & Eskim[7] proposed a constraint that allows for multiple genetic covariance matrices to be estimated subject to a diagonalization algorithm. However, a multivariate version of their method would constrain the multiple genetic covariance matrices to be proportional. This proportionality renders the diagonalization algorithm feasible, but may be substantively undesirable.

Here we propose a method to reduce the computational burden associated with a multivariate analysis by reducing the multivariate problem to multiple (tractable) univariate problems. We exploit the fact that the variance of the sum of 2 random variables is the sum of their variances plus twice their covariance. We estimate the genetic variance of both variables and the genetic variance for their sum. Given these results, we can derive the genetic covariance. We estimate the environmental variance for both variables and their sum, allowing us to derive the environmental covariance.

The method is further capable of estimating separate effects of multiple genetic covariance matrices for separate sets of SNPs. We can partition the covariance matrix V of a set of random variables Y into multiple genetic and/or environmental effects:

$$V = A_1 \otimes \Sigma_{g1} + A_2 \otimes \Sigma_{g2} + I_{(\mathrm{nxn})} \otimes \Sigma_e \quad (4)$$

For example, suppose a genetic relatedness matrix (A_1) is estimated using all SNPs in genes in the serotonergic pathway, and a second genetic relatedness matrix is estimated based on all other SNPs (A_2). The SNPs in A_1 and A_2 can be correlated due to LD. The validity of GREML estimates based on specific pathways and regulatory categories of SNPs has been studies extensively[8]. We can partition covariance matrix Σ of a set of k phenotypes Y into covariance matrices,Σ_{g1}, Σ_{g2}, and Σ_e. This model can, in addition, be applied to estimate separate covariance matrices conditioned on separate chromosomes, pathways, or functional categories of SNPs. If related individuals are present in the sample, one can accommodate this by modelling separate genetic covariance matrices for the closely and distantly related subjects separately (see[9]).

We simulated data to establish that the method performs as expected, and applied the method to an empirical, multivariate dataset of neuroticism and extraversion personality items. A recent genome-wide association study (GWAS) of a neuroticism sum score revealed a genome wide significant genetic variant (de Moor et al. 2015, in press[10]), and estimated the variance explained by all SNPs for neuroticism to be ~15% . For extraversion, the estimated SNP heritability is: ~12% [10; 11].

Methods.

Statistical methods

Let Y_i and Y_j be two correlated zero mean random variables (phenotype). The variance of the sum of two correlated random variables equals:

$$\mathrm{Var}(Y_1 + Y_2) = \mathrm{Var}(Y_1) + \mathrm{Var}(Y_2) + 2 * \mathrm{Cov}(Y_1, Y_2) \quad (5)$$

We assume that the variances of Y_1, Y_2, the covariance between Y_1 and Y_2, and the variance of the sum $Y_1 + Y_2$ are the sum of a genetic variance component and a residual variance component:

$$\text{Var}(Y_1) = \sigma^2_{g1} + \sigma^2_{e1} \quad (6a)$$

$$\text{Var}(Y_2) = \sigma^2_{g2} + \sigma^2_{e2} \quad (6b)$$

$$Var(Y_1 + Y_2) = \sigma^2_{g3} + \sigma^2_{e3} \quad (6c)$$

where

$$Var(Y_1 + Y_2) = \sigma^2_{g1} + \sigma^2_{g2} + \sigma^2_{e1} + \sigma^2_{e2} + 2*(\sigma_g(Y_1,Y_2) + \sigma_e(Y_1,Y_2)) \quad (6d)$$

It follows that:

$$\sigma^2_{g3} = \sigma^2_{g1} + \sigma^2_{g2} + 2*(\sigma_g(Y_1,Y_2)) \quad (6e)$$

This gives an unbiased estimate of the genetic and environmental covariance between trait Y1 and Y2:

$$\sigma_g(Y_1,Y_2) = (\sigma^2_{g3} - (\sigma^2_{g1} + \sigma^2_{g2}))/2 \quad (7a)$$

$$\sigma_e(Y_1,Y_2) = (\sigma^2_{e3} - (\sigma^2_{e1} + \sigma^2_{e2}))/2 \quad (7b)$$

Estimation of the genetic and environmental variance of Y_1, Y_2 and of the sum (Y_1+Y_2) using univariate GREML, and applying equation 7 gives the genetic covariance matrix Σ_g and environmental covariance matrix Σ_e. The resulting covariance matrix is not guaranteed to be positive definite due to sampling fluctuation. If the matrix is not positive definite, Higham's algorithm[12] can be applied to approximate the nearest positive definite matrix. From the genetic covariance matrix one can readily obtain the genetic en environmental correlation matrix (equation 8).

$$r_g = \frac{\sigma_{g(y_1y_2)}}{\sigma_{gy1} * \sigma_{gy2}} \quad (8a)$$

$$r_e = \frac{\sigma_{e(y_1y_2)}}{\sigma_{ey1} * \sigma_{ey2}} \quad (8b)$$

To expand from the bivariate to the multivariate, we sequentially estimate the genetic variance of each random variable and the sum of each pair of random variables. Given k phenotypes, this implies k*(k+1)/2 univariate analyses. As the genetic variance estimations are executed independently, this procedure can be parallelized.

In the present procedure the genetic correlation is not directly estimated but derived from other estimates, therefore no standard error is available. A Taylor approximation of the standard error of r_g was derived by Visscher et al[13] and can be used if a standard error is required. Extension of the model to include multiple genetic relatedness matrices (Equation 5) requires an adaptation of the functions for the genetic correlations and the standard errors (see Appendix A). Standard errors for genetic covariance were approximated using an extension of the Taylor approximation given in Visscher et al[13] .

Application to simulated data.

We simulated data to establish that the technique outlined above accurately estimates the genetic and environmental correlations for different true values for the genetic and environmental correlations between traits. In the first simulation, we retrieved the genetic correlation between two traits given a single GRM containing all SNPs. In the second simulation, we divided the genetic effects over 2 GRM's, one constructed with 4046 SNPs in the serotonin pathway, and one constructed with all other SNPs. The phenotype data were simulated for 4000 genotyped unrelated individuals in the Netherlands twin register (NTR). Subjects were genotyped on the Affymetrix 6.0 platform (described in chapter 7). Quality control selected SNPs with MAF > 0.01 and HWE < $1* 10^{-5}$ and all three GRMs were calculated in GCTA 1.24.1.

Application to neuroticism and extraversion data

The method was applied to neuroticism and extraversion item data collected from participants in the Netherlands Twin Register (NTR)[14]. Data on twelve neuroticism items and twelve extraversion items from the NEO PI questionnaire[15] were available for 5578 individuals in 2790 extended families. Items are scored by participants on a 5 point Likert scale, with categories ranging from "strongly disagree" to "strongly agree". We estimate the genetic covariance matrix based

on measured SNPs of the 24 item scores, the genetic covariance matrix based on pedigree data and the environmental covariance matrix. The genetic (co)variance structure of the NEO personality inventory items is known from a recent twin study[16]. The neuroticism items are expected to correlate positively, the extraversion items are expected to correlate positively, and the correlations between neuroticism and extraversion items are expected to be low and negative. This pattern of correlations has been confirmed across cultures[17]. Previous analysis of the neuroticism and extraversion sum scores using GCTA has revealed a low (15% for neuroticism, lower for extraversion) contribution of genetic effects measured by SNPs[10; 11]. All individuals in the represent study were genotyped on the Affymetix 6.0 platform (see chapter 7). Individuals were of European decent. We regressed the items on 3 ancestrally relevant principal components[18] and sex.

The Netherlands Twin Register sample contains a substantial number of twins and their extended family, and application of GREML to obtain the SNP heritability generally would require removing the closely related subjects. However, as mentioned above, our method can estimate multiple variance components, and we adopted the methodology suggested by Zaitlen et al.[9] to estimate the genetic effect attributable to SNPs and the genetic effect attributable to the total additive genetic variance. The first genetic effect is based on the genetic relatedness for all pairs of individuals (A_{n*n}^{full}) The second genetic effect is based on pairs of closely related individuals only ($A_{n*n}^{>0.05}$, relationships below 0.05 were all set to 0).

$$V = A_{n*n}^{full} \otimes \Sigma_{snp} + A_{n*n}^{>0.05} \otimes (\Sigma_a - \Sigma_{snp}) + \mathrm{I_{n*n}} \otimes \Sigma_e \quad (9)$$

For the 24 items, 3 covariance matrices were estimated: Σ_{snp}, Σ_e and $(\Sigma_a - \Sigma_{snp})$ based on the model in equation 9. The sum of Σ_{snp} and $(\Sigma_a - \Sigma_{snp})$ results an estimate of the total additive genetic (co)variance.

All models were estimate in R using full information maximum likelihood (FIML) optimization with exact derivatives in the function optim(). The implementation in R includes two FORTRAN routines to speed up essential steps in the likelihood function and the parameter derivatives. Scripts to perform the optimization and the required FORTRAN routines are available online (URL).

Results

Simulation

We simulated under the model in equation 4, in which we chose k to equal 2, and based A on the realized genetic relatedness matrix of 4000 distantly related genotyped individuals from the NTR. We simulated 3 sets of values forΣ_g, and Σ_e (see table 1).

Table 1: Simulated r_g, mean retrieved r_g ,mean standard error for r_g and the standard deviation for r_g over 100 replications.

Scenario	Simulated $\sum_{(kxk)g}$	Simulated $\sum_{(kxk)e}$	Simulated r_g	Estimated r_g	r_g (Visscher et al)	Mean SE(r_g)	SD(r_g)	SD(r_g) (Visscher et al)
1	1, .4 .4, 1	1, 0 0, 1	.4	.391	.391	0.1052	0.1148	0.1140
2	1, .8 .8, 1	1, 0 0, 1	.8	.806	.811	0.0882	0.0871	0.0837
3	1, .8 .8, 1	1, .4 .4, 1	.8	.802	.805	0.0623	0.0638	0.0615

We generated 100 instances of Y_1 and Y_2 given the covariance V, and means of 0 using mvrnorm() in R. We estimated the genetic correlation in 100 replications and present the mean of the retrieved correlations in table 1. As shown, the method produces unbiased estimates, in that the mean correlation of 100 replicates is close to the true population correlation. The standard deviations of the estimates of r_g are close to the theoretical standard errors. The proposed estimator of r_g correlates .99, .95 and .98 under scenario 1, 2, and 3, respectively, with the estimator of r_g as described in Visscher et al supplemental text S2[13].

We simulate 2 correlated traits, where the correlation between the traits is a function of 2 GRM's and a residual covariance matrix. We chose the number of traits k to equal 2. A_1 is the genetic relatedness matrix based on 4046 SNPs in the KEGG serotonine pathway[19] for 4000 distantly related genotyped individuals from the NTR, and A_2 is a GRM based on all other SNPs (not in the KEGG serotonine pathway) in the same 4000 NTR participants. We chose 2 sets of values for Σ_g, $\Sigma_{serotonine}$ and Σ_e as described in table 2.

Table 2: Parameters for two sets of $\sum_{(kxk)g}$ $\sum_{(kxk)serotonine}$ and $\sum_{(kxk)e}$, simulating two correlated traits for which the genetic correlation is different between SNPs in serotonin and SNPs in the rest of the genome.

	$\sum_{(kxk)g}$	Y1	Y2	$\sum_{(kxk)serotonine}$	Y1	Y2	$\sum_{(kxk)e}$	Y1	Y2
Set 1	Y1	.9	.23	Y1	.10	.07	Y1	1	.3
	Y2	.23	.9	Y2	.07	.10	Y2	.3	1
Set 2	Y1	.9	.4	Y1	.10	.10	Y1	1	.6
	Y2	.4	.9	Y2	.08	.08	Y2	.6	1

Note that the SNPs in the serotonin pathway were simulated to contribute 5% of variance in trait Y_1 and Y_2, and the correlation between serotonergic effects on Y_1 and Y_2 was chosen to equal 0.7. We generated 100 instances of Y_1 and Y_2 given the covariance V and means of 0 using mvrnorm() in R. Subsequently we estimated the genetic correlations r_g and $r_{serotonine}$ and the environmental correlation r_e.

Table 3: simulated $r_{serotonine}$, r_g and r_e, mean retrieved $r_{serotonine}$, r_g and r_e in 100 simulations, the standard deviations of the 100 estimates of $r_{serotonine}$, r_g and r_e and the approximate standard errors of $r_{serotonine}$, r_g and r_e.

	r(Y1 Y2)	Simulated r	Mean Estimated r	SD(r)	Mean SE (aproximated)
Set 1	$r_{serotonine}$	0.7	0.696	0.134	0.139
	r_g	0.2556	0.233	0.138	0.128
	r_e	0.3	0.312	0.119	0.108
Set 2	$r_{serotonine}$	0.8	0.792	0.107	0.101
	r_g	0.44444	0.431	0.120	0.111
	r_e	0.6	0.602	0.091	0.083

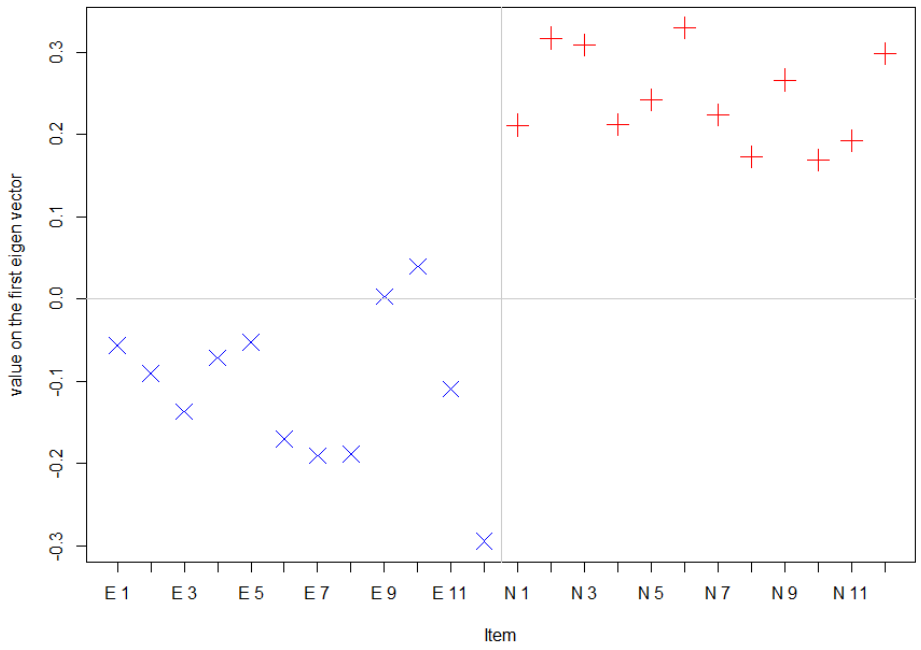


Figure 1: Loadings on the first principle component of $\mathbf{\Sigma_{snp}}$show a clear separation between the extraversion(x, blue) and neuroticism items(+, red).

The covariance matrices Σ_{snp} , Σ_a, and Σ_e were transformed to the nearest positive definite covariance matrices using Higham's algorithm[12]. The nearest positive definite matrix was subsequently standardized to obtain the correlation matrix. This procedure resulted in 3 correlation matrices: the correlation matrices between 24 items based on the effects of SNPs (r_{snp}), the correlation matrix based on familial genetic effects (r_a) and the environmental effect(r_e). We further computed the observed correlation between items (*r* observed). The 4 correlation matrices are plotted in figure 2.

As Figure 2 clearly shows the item correlation based on measured SNPs follow the expected pattern. We found high correlations between neuroticism items, mostly positive correlations within extraversion items, negative correlations across the two traits. The pattern of correlations is in line with our previously formulated expectations.

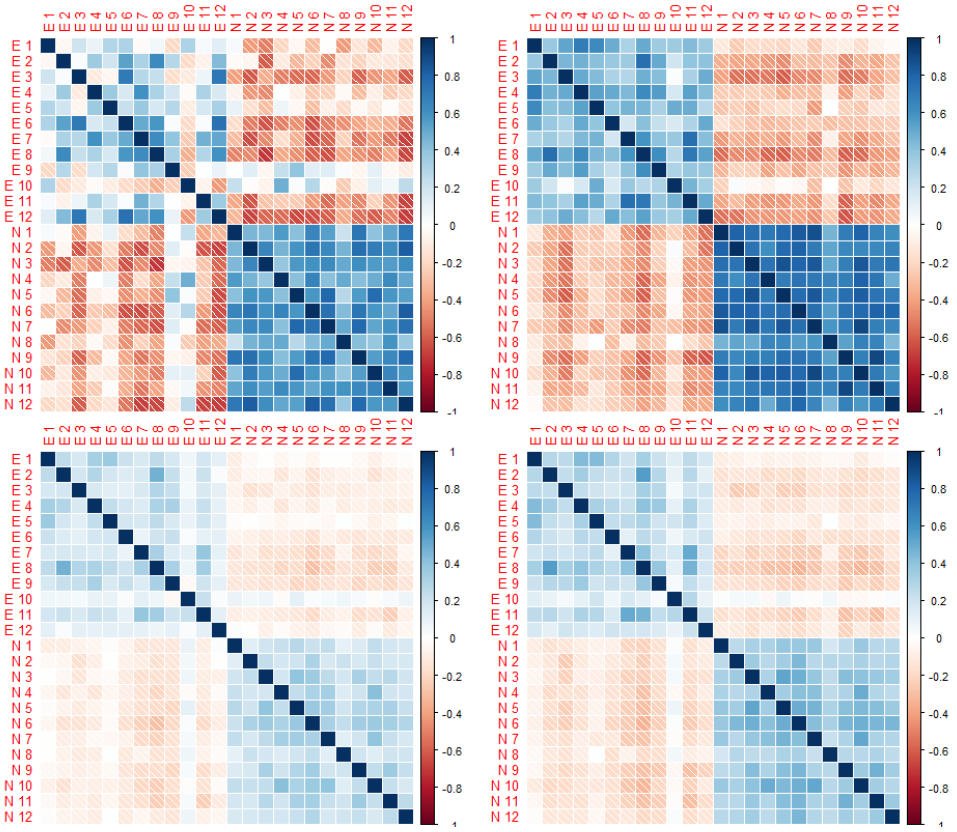


Figure 2: clockwise from top left the four 24x24 correlation matrices r_{snp}, r_{twin}, the observed correlation and r_e.

Discussion.

The method proposed here involves reducing a multivariate estimation problem to a series of univariate estimation problems. The method has the benefit of rapidly estimating the genetic covariance between traits given genome wide data. The (genetic) covariance matrices are estimated in a series of univariate analyses. The method can further estimate separate genetic covariance matrices for distinct sets of SNPs, or other distinct sources of genetic information (i.e., pedigree and SNPs). The estimates of genetic covariance were found to be unbiased. The standard errors based on extension of the work of Visscher et al.[13] are very close to the observed standard deviation of the parameters in our simulations.

The multivariate analysis of personality data revealed the expected covariance structure based on measured genotypic information. The SNP variances of three extraversion items were close to zero. The removal of these items from the scale followed by the recalculation of the SNP heritability of the scale revealed a higher SNP heritability. The present results show that our method is capable of detecting a specific genetic covariance structure, even if the power to detect individual genetic (co)variance parameters is low. While this particular covariance structure is well known and has often been replicated, this type of analysis lends itself well for exploration of large -omics datasets.

Our method has three limitations. First, the covariance matrices, constructed piecemeal from univariate analysis, are not necessarily positive definite. Various techniques exist to determine the nearest positive definite matrix should the user require a positive definite covariance matrix. Here we employed Hingham's algorithm [12], however given the dimensions of the covariance matrix or specific constraints one might want to place on the covariance matrix other solutions might be preferred. Second, the method is currently performed on pair wise complete data. All parameters (e.g h^2 and $r_{g,}$ $SE_{(rg)}$) could be estimated in the presence of missingness, however their performance in the presence of missingness in unknown. Further investigation is needed to extend the current method to handle missing data.

A third limitation is that this method still requires access to raw genotypes, where heritability and genetic correlation can be estimated from summary statistics from GWAS analysis using "LD score regression"[20; 21]. LD-score regression is further

capable of partitioning the genetic effect into multiple genetic effects, each reflecting the effect of a subsets of SNPs[22]. LD-score regression is computationally less burdensome than GREML. However, GREML is flexible in that it allows for random effects not only to reflect the effects of SNPs but can be extended to accommodate other effects. These effects include concurrently modelling SNP and additive effects as discussed above, gene by environment interaction (van Dongen et al; in preparation, Nivard et al. Submitted), and maternal genetic effects that influence traits via the environment[23].

The proposed technique allows the estimation of a genetic covariance matrix for a large number of traits of interest, and further allows concurrent estimation of multiple genetic covariance matrices based on distinct genetic effects. The technique can therefore be implemented for functional category based, pathway based or gene based tests of covariance between traits

References

1. Yang J., Lee S.H., Goddard M.E., and Visscher P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics *88*, 76-82

2. Lee S.H., Yang J., Goddard M.E., Visscher P.M., and Wray N.R. (2012) Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics *28*, 2540-2542

3. Zhou X., Stephens M. (2013) Efficient Algorithms for Multivariate Linear Mixed Models in Genome-wide Association Studies. arXiv preprint arXiv:1305. 4366

4. Lee J.J., Chow C.C. (2014) Conditions for the validity of SNP-based heritability estimation. Human genetics *133*, 1011-1022

5. Listgarten J., Lippert C., and Heckerman D. (2013) FaST-LMM-Select for addressing confounding from spatial structure and rare variants. Nature genetics *45*, 470-471

6. Lippert C., Listgarten J., Liu Y., Kadie C.M., Davidson R.I., and Heckerman D. (2011) FaST linear mixed models for genome-wide association studies. Nature Methods *8*, 833-835

7. Kostem E., Eskin E. (2013) Improving the Accuracy and Efficiency of Partitioning Heritability into the Contributions of Genomic Regions. The American Journal of Human Genetics *92*, 558-564

8. Gusev A., Lee S.H., Neale B.M., Trynka G., Vilhjalmsson B.J., Finucane H., Xu H., Zang C., Ripke S., and Stahl E. (2014) Regulatory variants explain much more heritability than coding variants across 11 common diseases. bioRxiv, 004309

9. Zaitlen N., Kraft P., Patterson N., Pasaniuc B., Bhatia G., Pollack S., and Price A.L. (2013) Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. PLoS genetics *9*, e1003520

10. De Moor, M. H. M. Genome-wide association study identifies novel locus for neuroticism and shows polygenic association with Major Depressive Disorder. 2015.
Ref Type: Unpublished Work

11. Vinkhuyzen A.A., Pedersen N.L., Yang J., Lee S.H., Magnusson P.K., Iacono W.G., McGue M., Madden P.A.F., Heath A.C., and Luciano M. (2012) Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. Translational psychiatry *2*, e102

12. Higham N.J. (2002) Computing the nearest correlation matrix-a problem from finance. IMA journal of Numerical Analysis *22*, 329-343

13. Visscher P.M., Hemani G., Vinkhuyzen A.A., Chen G.B., Lee S.H., Wray N.R., Goddard M.E., and Yang J. (2014) Statistical power to detect genetic (co) variance of complex traits using SNP data in unrelated samples. PLoS genetics *10*, e1004269

14. Willemsen G., Vink J.M., Abdellaoui A., den Braber A., van Beek J.H., Draisma H.H., van Dongen J., vanGÇÿt Ent D., Geels L.M., and van Lien R. (2013) The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. Twin Research and Human Genetics *16*, 271-281

15. Costa PT, MacCrae RR (1992) Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI): Professional Manual. Psychological Assessment Resources,

16. Franic S., Borsboom D., Dolan C.V., and Boomsma D.I. (2014) The big five personality traits: psychological entities or statistical constructs? Behavior genetics *44*, 591-604

17. McCrae R.R., Costa J., Del Pilar G.H., Rolland J.P., and Parker W.D. (1998) Cross-cultural assessment of the five-factor model: The Revised NEO Personality Inventory. Journal of Cross-Cultural Psychology *29*, 171-188

18. Abdellaoui A., Hottenga J.J., de Knijff P., Nivard M.G., Xiao X., Scheet P., Brooks A., Ehli E.A., Hu Y., and Davies G.E. (2013) Population structure, migration, and diversifying selection in the Netherlands. European journal of human genetics *21*, 1277-1285

19. Kanehisa M., Goto S. (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic acids research *28*, 27-30

20. Bulik-Sullivan B., Loh P.R., Finucane H., Ripke S., Yang J., Psychiatric Genomics Consortium S.W.G., Patterson N., Daly M.J., Price A.L., and Neale B.M. (2014) LD Score Regression Distinguishes Confounding from Polygenicity in Genome-Wide Association Studies. bioRxiv

21. Bulik-Sullivan B., Finucane H.K., Anttila V., Gusev A., Day F.R., Consortium R., Genomics Consortium, of the Wellcome Trust Consortium, Perry J.R.B., Patterson N. et al. (2015) An Atlas of Genetic Correlations across Human Diseases and Traits. bioRxiv

22. Finucane H.K., Bulik-Sullivan B., Gusev A., Trynka G., Reshef Y., Loh P.R., Anttilla V., Xu H., Zang C., Farh K. et al. (2015) Partitioning heritability by functional category using GWAS summary statistics. bioRxiv

23. Eaves L.J., Pourcain B.S., Smith G.D., York T.P., and Evans D.M. (2014) Resolving the effects of maternal and offspring genotype on dyadic outcomes in genome wide complex trait analysis (`M-GCTA`). Behavior genetics *44*, 445-455

Appendix A

If we wish to approximate standard errors for r_g in:

$$V_{(\mathrm{kxk}\otimes\ \mathrm{nxn})} = A_{(\mathrm{nxn})} \otimes \Sigma_{\mathrm{g}\ (\mathrm{kxk})} + I_{(\mathrm{nxn})} \otimes \Sigma_{\mathrm{e}\ (\mathrm{kxk})}$$

Visscher et al. (2014) provide a first order Taylor approximation of the sampling variance of r_g. However, if we wish to approximate standard errors for r_{ga} and r_{gb} in:

$$V = A_{\mathrm{a}} \otimes \Sigma_{\mathrm{ga}} + A_b \otimes \Sigma_{gb} + I_{(\mathrm{nxn})} \otimes \Sigma_{\mathrm{e}}$$

we need to extend he approximation provided by Visscher et al.

Let Y_1 and Y_2 be two traits of interest with unit variance and means of 0. Let the variance of Y_1 and Y_2 be a function of 2 genetic factors and a residual.

$$\mathrm{Var}(Y_1) = \sigma^2_{ga\ Y1} + \sigma^2_{gb\ Y1} + \sigma^2_{eY1}$$

$$\mathrm{Var}(Y_2) = \sigma^2_{ga\ Y2} + \sigma^2_{gb\ Y2} + \sigma^2_{eY2}$$

Let us then construct trait $z_1 = y_1 + y_2$ and $z_2 = y_1 - y_2$ which implies, $cov(z_1,z_2) = 0$. If we decompose the variance for z_1 and z_2 as below:

$$Var(Z_1) = \sigma^2_{ga\ Z1} + \sigma^2_{gb\ Z1} + \sigma^2_{e\ Z1}$$

$$Var(Z_2) = \sigma^2_{ga\ Z1} + \sigma^2_{gb\ Z2} + \sigma^2_{e\ Z2}$$

We can determine r_{ga} the correlation between trait Y_1 and Y_2 given the first genetic effect a and r_{gb} the correlation between Y_1 and Y_2 given the second genetic effect *b*.

$$Cov(Y_1,Y_2)_{ga\ =} = (\sigma^2_{ga\ Z1} - (\sigma^2_{ga\ Y1} + \sigma^2_{ga\ Y2}))\ /2$$

$$Cov(Y_1,Y_2)_{gb\ =} = (\sigma^2_{gb\ Z1} - (\sigma^2_{gb\ Y1} + \sigma^2_{gb\ Y2}))\ /2$$

$$r_{ga} = Cov(Y_1,Y_2)_{ga} / (\sigma_{ga\ Y1} * \sigma_{ga\ Y2})$$

$$r_{gb} = Cov(Y_1,Y_2)_{gb} / (\sigma_{gb\ Y1} * \sigma_{gb\ Y2})$$

We can then determine the variances explained by each of the genetic effects in trait Y_1 and trait Y_2

$$g^2_{a\,Y1} = \sigma^2_{ga\,Y1} / var(Y_1)$$

$$g^2_{b\,Y1} = \sigma^2_{gb\,Y1} / var(Y_1)$$

$$g^2_{a\,Y2} = \sigma^2_{ga\,Y2} / var(Y_2)$$

$$g^2_{b\,Y2} = \sigma^2_{gb\,Y2} / var(Y_2)$$

We extend this function to obtain the sampling variance for the genetic correlation between traits Y_1 and Y_2 given the SNPs in GRM A_a (r_{ga}) and the correlation for SNPs in GRM A_b (r_{gb}).

We empirically estimate $var(\sigma^2_{ga\,Z1})$, $var(\sigma^2_{ga\,Z2})$, $var(\sigma^2_{gb\,Z1})$, $var(\sigma^2_{gb\,Z2})$, $g^2_{a\,Y1}$, $g^2_{b\,Y1}$, $g^2_{a\,Y2}$, and $g^2_{b\,Y2}$. The empirical estimates of sampling variance account for the fact that the parameters $g^2_{a\,Y1}$ and $g^2_{b\,Y1}$, and $g^2_{a\,Y2}$ and $g^2_{b\,Y2}$ are correlated. Based on these estimates we express the sampling variance for the genetic correlation between traits Y_1 and Y_2 given the SNPs in GRM A_a, var(r_{ga}) as:

$$\mathrm{Var}(r_{ga}) = \left(\frac{1 - r_{ga}}{4 * \sqrt{g^2_{a\,Y1} * g^2_{a\,Y2}}}\right)^2 * \mathrm{var}(\sigma^2_{ga\,Z1}) + \left(\frac{1 + r_{ga}}{4 * \sqrt{g^2_{a\,Y1} * g^2_{a\,Y2}}}\right)^2 * \mathrm{var}(\sigma^2_{ga\,Z2})$$

And the sampling variance for the genetic correlation between traits Y_1 and Y_2 given the SNPs in GRM A_b is given by:

$$\mathrm{Var}(r_{gb}) = \left(\frac{1 - r_{gb}}{4 * \sqrt{g^2_{b\,Y1} * g^2_{b\,Y2}}}\right)^2 * \mathrm{var}(\sigma^2_{gb\,Z1}) + \left(\frac{1 + r_{gb}}{4 * \sqrt{g^2_{b\,Y1} * g^2_{b\,Y2}}}\right)^2 * \mathrm{var}(\sigma^2_{gb\,Z2})$$

Note that determining the sampling variances for correlations in multiple GRM's requires us to use estimated sampling variances of gz_{1a} ,gz_{1b} ,gz_{2a} and gz_{2b}. Therefore we are required to estimate a univariate model for variable Z_1 and Z_2. Estimation of the genetic correlations only require estimation of the univariate model for variable Z_1. The estimation of standard errors for multiple genetic correlations increases the computational burden of our method.

Chapter 9: Summary

This thesis focused on the genetics of psychopathology across the lifespan. Genetic contributions to measures of psychopathology were estimated based on twin and family models, and on measured genotypes (i.e., single nucleotide polymorphisms). **Chapter 1** broadly outlined the methods used in behavior genetics, and genetic epidemiology as applied in this thesis, and discussed how measured genetic variants and environmental exposures can be incorporated into genetic studies.

In **chapter 2,** the results were presented of a cohort sequential study of the genetic and environmental influences on Symptoms of Anxiety and Depression (SxAnxDep) between age 3 and 63. Symptoms of anxiety and depression were measured in twins participating in research of the Netherlands twin register (NTR).[1-3] Young twins were rated by their mother at age 3, 7, 10, and 12. Self-report data for these twins was available at age 14, 16 and 18 years, and for adolescent and adult twins of 14 years and older, who participated in up to 8 waves of data collection in the Adult Netherlands Twin Register. SxAnxDep were assessed using an age appropriate version of the Anxious-Depression subscale of the Child Behavior Check List (CBCL; ages 3 through 12) and the Youth or Adult self report (YSR and ASR) inventories (ages 14 through 63).[4] The availability of twin data allowed us to estimate the proportion of variance in SxAnxDep that was explained by genetic and the environment at different ages. The availability of repeated measures (up to a maximum of 8 repeated measures on some participants) allowed us to estimate to what extent the genetic and environmental factors at one age played a role at a later age (transmission), and to what extent novel genetic and environmental factors (innovation) were important. Specifically, after organizing the data into 2 year age bins, we fitted a genetic simplex model to obtain this information concerning stability and innovation.

Results showed a decrease in the heritability of SxAnxDep between childhood and adulthood. The heritability was around .60 in childhood, and decreased to around .40 in adolescents and adults. This decrease was caused by an increase in environmental variance that outpaces a simultaneous increase in genetic variance between the ages of 3 and 18. After age 18 the genetic variance in SxAnxDep remained very stable, genetic influences are highly correlated between ages (around .90), and new genetic variance (innovation) is absent after age 18, except for time point specific sources of genetic variance. The environmental variance in

SxAnxDep was also transmitted from age to age in adulthood. However at each age in adulthood, environmental innovation was present, as new environmental sources of variance (environmental innovation) were present. With increasing age, the transmission of environmental effects rose, as the new environmental influences on SxAnxDep reduced somewhat. This process resulted in increasing environmental correlations between subsequent SxAnxDep scores with increasing age.

Chapter 3 focused on internalizing and externalizing psychopathology. Internalizing psychopathology included measures of depressive and anxiety disorders, and externalizing psychopathology included measures of attention deficit hyperactivity disorder (ADHD), oppositional defiant disorder (ODD) and conduct disorder (CD). The aim of the analyses was to describe the transition between childhood and adolescence and the co-morbidity of internalizing and externalizing psychopathology over development.

Data from the Avon Longitudinal study of parents and children (ALSPAC) were analyzed in Bristol (UK). Psychopathology was measured at age 7, 10, 13, and 15 using the DAWBA diagnostic interview.[5] The DAWBA provide both a diagnosis and an ordered categorical score, with higher scores indicating higher risk to fulfill the criteria for the disorders. The ordered categorical scores, denoted DAWBA bands, are statistically more informative than the binary diagnosis.[5]

Growth mixture models were fitted to identify categorically distinct trajectories for both internalizing and externalizing psychopathologies. For both internalizing and externalizing disorders, we expected an increasing trajectory, a decreasing trajectory, a stable high category and a large and stable low category. Finally, a single model was fitted modeling the co-occurrence between trajectories of internalizing and externalizing psychopathology.

The results of the analyses revealed that internalizing psychopathology was best captured by a model with 5 distinct trajectories. In addition to two trajectories with a consistently low and very low risk for internalizing psychopathology, the other 3 trajectories were a decreasing risk for internalizing psychopathology, an increasing risk of internalizing psychopathology, and an adolescent onset risk of internalizing psychopathology. Externalizing psychopathology was also characterized by 5 distinct trajectories. Four were similar to internalizing psychopathology. However, instead of an adolescent onset risk trajectory, a

trajectory of stable high risk of externalizing disorders through childhood and adolescence was identified.
Combined analysis of internalizing and externalizing categories revealed that increasing internalizing and increasing externalizing trajectories co-occur, as did decreasing internalizing and externalizing trajectories. However, the adolescent onset internalizing trajectory was independent of high externalizing trajectories, and the persistent high externalizing trajectory was mainly associated with the decreasing internalizing trajectories. Sex and early life environmental risk factors predicted externalizing and, to a lesser extent, internalizing trajectories. The analysis reveals the need to screen for co-morbidity in the case of either early onset externalizing or internalizing problems. The only exception seems to be adolescent onset internalizing problems, which are not related to a high risk for externalizing problems.
Thus, Chapters 2 and 3 reveal both genetic and phenotypic continuity between childhood psychiatric problems and adolescent and adult psychiatric outcomes.

Chapter 4 reports on a genome-wide association study (GWAS) on preschool internalizing problems. Results from 3 cohorts, including NTR, were meta-analyzed, and the variance explained by all measured genetic variants (SNPs) was estimated. Three cohorts participated: the NTR, The Western Australian Pregnancy Cohort Study (Raine), and the generation R from Rotterdam. Internalizing scores in 2-3 year olds, based on the Child Behavior Check List[6], were harmonized. Genotypes were imputed against HAPMAP 2.[7] After post imputation quality control, 2.4 million SNPs were available for analysis. A total of 2037 children had genotype and internalizing scores available in generation R; 1475 children from 1031 families in the NTR; and 1084 children in the Raine cohort. In each cohort, the association between Internalizing and genotypes was tested, with the inclusion of principle components to correct for population stratification and sex. This was followed by meta-analysis of 4566 children. The variance explained by all SNPs for internalizing behavior was estimated using 2 methods, Genomic Relationship Matrix Restricted Maximum Likelihood (GREML) as implemented in Genomic Complex Trait Analysis (GCTA) and density estimation (DE).[8; 9]
The SNPs were found to explain between 13 and 43% of the total variance in internalizing problems. As the heritability in twin studies was estimated at 59%, this implies that the genetic variants analyzed in this study captured between 22

and 72% of the genetic variance. The meta-analysis revealed no SNPs associated with Internalizing problems at a genome-wide significant p-value < $1x10^{-8}$. In two regions, there were SNPs, which reached a p-value below $1x10^{-5}$ in the meta-analysis. One SNP was located in an intergenic region on chromosome 9. The other region was on chromosome 20, and included SNPs of the PCSK2 gene. PCSK2 is an important protein in the processing of pro-insulin to insulin, and PCSK2 variants are correlated with insulin resistance,[10; 11] myocardial infarction[12] and age at menarche[13]. The link between depression and cardiovascular disease has long been recognized. Post hoc analysis of SNPs that were previously associated with adult internalizing psychopathology, psychopathology that usually presents in childhood (ADHD, conduct disorder), or psychotic disorders, and of SNPs in candidate genes[14] did not show a significant association of any of these SNPs with internalizing problems in preschool children. Collectively, the SNPs previously associated with adult internalizing disorders did not show lower p-values than expected by chance. However, the SNPs previously associated with adult internalizing disorder, adult or childhood psychiatric disorders usually diagnosed in childhood or psychotic disorders did collectively show lower p-values than expected by chance in the GWAS of preschool internalizing problems. Adding SNPs associated with treatment response diminished this signal, while subsequently adding SNPs in candidate genes slightly strengthened the signal. The analyses performed in chapter 4 show that childhood preschool internalizing problems are heritable, and that a substantial part of this heritability can be explained by common genetic variation. The results further show that childhood internalizing problems are a complex trait, and no single genetic variant explains a substantial part of the phenotypic variation. The significant signal of SNPs previously associated with adult and other psychiatric disorders was suggestive of common genetic causes.

Chapter 5 looked at polygenic score prediction of childhood psychopathology. The most recent schizophrenia GWAS meta-analysis included 36,989 cases and 113,075 controls and revealed 108 loci significantly associated with schizophrenia.[15] This study provided the starting point to test for associations between genetic risk for schizophrenia and childhood psychopathology directly at the molecular genetics level. Polygenic risk scores were calculated based on the schizophrenia GWAS to predict childhood psychopathology scores at ages 7, 10, 12, and 15 years. The analysis was performed in samples from the Netherlands Twin

Registrer (NTR) and the Avon longitudinal study of parents and children (ALSPAC). In both cohorts, DSM based measures of anxiety, depression, attention deficit hyperactivity disorder (ADHD), and oppositional deviant disorder, and conduct disorder (ODD & CD) were available. The NTR scores were based on the DSM oriented CBCL or YSR scales[16] and the scores in ALSPAC on the DAWBA bands.[5] The regression of the psychopathology phenotype on the polygenic risk score included as covariates principle components to control for population stratification and sex. Meta analysis of the results of both studies revealed an false discovery rate (FDR) corrected significant association between schizophrenia risk and anxiety at age 10. This result seemed mainly to be driven by results in the NTR. The analysis further revealed associations at uncorrected $p < 0.05$ between schizophrenia polygenic risk scores and anxiety at age 7 and depression at age 7, age 10, and age 12 to 13. Based on these results the initial hypothesis of a broad positive association with childhood psychopathology was not confirmed. Post hoc test revealed a stronger effect on internalizing psychopathology than on externalizing psychopathologiy. Note that the results were consistent with the PGC cross disorder study[17]. The PGC cross disorder group found a genetic correlation between adult MDD and schizophrenia, but not between schizophrenia and ADHD[17].

In **chapter 6** the aim was to replicate a finding reported by the rat genome sequencing and mapping consortium. This consortium obtained evidence for an association between the CTNND2 gene and anxiety in rats.[18] Replication was sought in a sample of adult participants from NTR and from the Netherlands Study of Depression and Anxiety (NESDA). The phenotype was based on the CIDI anxiety Diagnosis. All individual SNPs in the CTNND2 gene were tested for association with anxiety in the NTR/NESDA sample. To test for an association between CTNND2 and Major Depressive Disorder (MDD), Bipolar Disorder, and Schizophrenia lookups in the results of the PGC mega and meta-analyses of MDD, bipolar disorder and schizophrenia were performed. No SNPs reached significance for any disorder corrected for the number of SNPs tested. A gene-based test for enrichment of all P-values in the gene was performed, and revealed tentative evidence for enrichment of the CTNND2 gene in anxiety, MDD, and schizophrenia, but not bipolar disorder. This chapter shows that follow up of findings in animal studies can reveal potential associations in human data, and may provide a useful addition tool to explore genetic associations.

The current REML model as implemented in GCTA allows for genotype by environment (GxE) moderation in the case that the environmental exposure is dichotomous.[8] In **chapter 7** the model specified in GCTA was extended to include continuous moderation of genetic and environmental effects, given a sample of closely related and nominally unrelated individuals. This involved a reparameterization of the model proposed by Zaitlen et al.[19]. This resulted in a model in which the concurrent moderation of the variance specifically attributable to SNPS, and the total additive genetic variance can be tested. We applied this model to symptoms of anxiety and depression (AnxDep), attention problems (AP), height, and body mass index (BMI). The analysis revealed that the (genetic) variance components for the different phenotypes were differently moderated by age (or birth year in the case of height).

Fitting the Zaitlen model to the four phenotypes revealed moderated additive genetic effects (~40%) for AnxDep and A) and strong additive genetic effects for Height (90%) and BMI (75%). The portion of variance explained by measured SNPs was moderate for AnxDep (~10%) and AP (~11%), but larger for height (~55%) and BMI (~40%).

We proceeded to fit moderation models. The variance explained by SNPs, additive variance and residual variance for AnxDep were not moderated. For BMI the additive genetic variance and the residual variance were moderated, but the variance explained by SNPs was not. In the analysis of height and AP, the residual variance was moderated, but the additive variance or variance explained by SNPs was not. The analysis revealed differences in the way age and or birth year moderated these 4 phenotypes.

chapter 8 presented a model that allows for the estimation of genetic (co)variance between traits based on measured genotypes. This in itself is not new. In bivariate GCTA[20] and GEMMA[21], it is also possible to estimate the genetic covariance between traits given all SNPs. However, increasing the number of traits or the number of separate genetic variance components will increase computational burden. In this chapter, a method was developed that breaks the multivariate analysis up in to a series of univariate analysis. This method relies on the fact that the variance of the sum of two variables equal to the variance of each individual variable and twice the covariance between two variables. Simulations showed that this method yields unbiased estimates of genetic (co)variance. Moreover, approximate standard errors were obtained using a Taylor approximation, first used by Visscher et al.[22] The model was extended to

allow for multiple genetic effects. We simulated data to show that the model produced unbiased estimates of separate genetic covariance matrices for each genetic relatedness matrix, and that their standard errors are correctly.
The method was applied to 24 items derived from the NEO PI personality inventory.[23; 24] Twelve items indicate the construct of neuroticism, 12 items indicate the construct of extraversion. We estimated the genetic correlation between the items based on SNPs in a sample containing related individuals. Based on previous results[25] neuroticism items were expected to correlate positively, extraversion items were expected to correlate positively. Negative correlations were expected between the neuroticism and extraversion items. Only a modest proportion of variance in the individual items was attributable to SNPs (0 to 14.8% for extraversion, 5 to 16.7% for neuroticism). Variance explained by SNPs in the total scale was 6.3% for extraversion and 22.6% for neuroticism. Despite these moderate SNP heritability's we were able to retrieve the expected covariance structure. The first principle component of the genetic covariance matrix of the items separated the neuroticism and extraversion items. A second feature of the model is the possibility to estimate a separate genetic (co)variance for multiple sets of SNPs. This would allow for estimation of separate co-variances between traits for multiple distinct sets of SNP. Each of these sets of SNPs could be selected to reflect a set of genes in a biological pathway, a specific chromosome or any other biologically interesting subset of all measured SNPs.

In **chapter 10** I discuss the developments in the field of behavior genetic in the periode that I was writing this dissertation. I discuss how current (methodological) developments will allow a deeper understanding of the genetics of complex traits in the coming few years.

References

1. Willemsen G., De Geus E.J., Bartels M., Van Beijsterveldt C.E.M., Brooks A.I., Estourgie-van Burk G.F., Fugman D.A., Hoekstra C., Hottenga J.J., and Kluft K. (2010) The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. Twin Research and Human Genetics *13*, 231-245

2. Willemsen G., Vink J.M., Abdellaoui A., den Braber A., van Beek J.H., Draisma H.H., van Dongen J., van't Ent D., Geels L.M., and van Lien R. (2013) The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. Twin Research and Human Genetics *16*, 271-281

3. van Beijsterveldt C.E., Groen-Blokhuis M., Hottenga J.J., Frani-ç S., Hudziak J.J., Lamb D., Huppertz C., de Zeeuw E., Nivard M., and Schutte N. (2013) The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. Twin Research and Human Genetics *16*, 252-267

4. Achenbach T.M., Rescorla L.A., McConaughey S., Pecora P., Wetherbee K., and Ruffle T. (2003) The Achenbach system of empirically based assessment. Handbook of psychological and educational assessment of children: Personality, behavior, and context *2*, 406-432

5. Goodman A., Heiervang E., Collishaw S., and Goodman R. (2011) The 'DAWBA bands` as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. Social psychiatry and psychiatric epidemiology *46*, 521-532

6. Achenbach TM, Rescorla LA (2000) Manual for the ASEBA Preschool Forms & Profiles: An Integrated System of Multi-informant Assessment; Child Behavior Checklist for Ages 1 1/2-5; Language Development Survey; Caregiver-Teacher Report Form. University of Vermont,

7. Frazer K.A., Ballinger D.G., Cox D.R., Hinds D.A., Stuve L.L., Gibbs R.A., Belmont J.W., Boudreau A., Hardenbol P., and Leal S.M. (2007) A second generation human haplotype map of over 3.1 million SNPs. Nature *449*, 851-861

8. Yang J., Lee S.H., Goddard M.E., and Visscher P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics *88*, 76-82

9. So H., Li M., and Sham P.C. (2011) Uncovering the total heritability explained by all true susceptibility variants in a genomeΓÇÉwide association study. Genetic epidemiology *35*, 447-456

10. Leak T.S., Keene K.L., Langefeld C.D., Gallagher C.J., Mychaleckyj J.C., Freedman B.I., Bowden D.W., Rich S.S., and Sale M.l.M. (2007) Association of the proprotein convertase subtilisin/kexin-type 2 (PCSK2) gene with type 2 diabetes in an African American population. Molecular genetics and metabolism *92*, 145-150

11. Jonsson A., Isomaa B., Tuomi T., Eriksson J.G., Groop L., and Lyssenko V. (2012) Effect of a common variant of the PCSK2 gene on reduced insulin secretion. Diabetologia *55*, 3245-3251

12. Fujimaki T., Kato K., Yokoi K., Oguri M., Yoshida T., Watanabe S., Metoki N., Yoshida H., Satoh K., and Aoyagi Y. (2010) Association of genetic variants in SEMA3F, CLEC16A, LAMA3, and PCSK2 with myocardial infarction in Japanese individuals. Atherosclerosis *210*, 468-473

13. Elks C.E., Perry J.R., Sulem P., Chasman D.I., Franceschini N., He C., Lunetta K.L., Visser J.A., Byrne E.M., and Cousminer D.L. (2010) Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. Nature genetics *42*, 1077-1085

14. Bosker F.J., Hartman C.A., Nolte I.M., Prins B.P., Terpstra P., Posthuma D., Van Veen T., Willemsen G., Derijk R.H., and De Geus E.J. (2011) Poor replication of candidate genes for major depressive disorder using genome-wide association data. Molecular psychiatry *16*, 516-532

15. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014) Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421-427

16. Achenbach T.M., Dumenci L. (2001) Advances in empirically based assessment: Revised cross-informant syndromes and new DSM-oriented scales for the CBCL, YSR, and TRF: Comment on Lengua, Sadowski, Friedrich, and Fisher (2001).

17. Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. The Lancet *381*, 1371-1379

18. Sequencing R.G., Mapping Consortium (2013) Combined sequence-based and genetic mapping analysis of complex traits in outbred rats. Nature genetics *45*, 767-775

19. Zaitlen N., Kraft P., Patterson N., Pasaniuc B., Bhatia G., Pollack S., and Price A.L. (2013) Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. PLoS genetics *9*, e1003520

20. Lee S.H., Yang J., Goddard M.E., Visscher P.M., and Wray N.R. (2012) Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics *28*, 2540-2542

21. Zhou X., Stephens M. (2014) Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nature methods *11*, 407-409

22. Visscher P.M., Hemani G., Vinkhuyzen A.A., Chen G.B., Lee S.H., Wray N.R., Goddard M.E., and Yang J. (2014) Statistical power to detect genetic (co) variance of complex traits using SNP data in unrelated samples. PLoS genetics *10*, e1004269

23. McCrae R.R., Costa J., Del Pilar G.H., Rolland J.P., and Parker W.D. (1998) Cross-cultural assessment of the five-factor model: The Revised NEO Personality Inventory. Journal of Cross-Cultural Psychology *29*, 171-188

24. Costa PT, MacCrae RR (1992) Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI): Professional Manual. Psychological Assessment Resources,

25. Franic S., Borsboom D., Dolan C.V., and Boomsma D.I. (2014) The big five personality traits: psychological entities or statistical constructs? Behavior genetics *44*, 591-604

Chapter 10: General discussion.

The field of psychiatric genetics, and human genetics more broadly, has undergone rapid progress in the last decade. To appreciate this progress, consider schizophrenia. In 2009[1], just before the onset of the work leading to this dissertation, 3 regions in the genome were found to be associated with schizophrenia. The number of associated independent risk loci for schizophrenia has since grown to 108[2]. While schizophrenia, for now, remains the biggest success story in psychiatric genetics, there is little doubt that sufficiently large samples will result in similar success in identifying genetic loci associated with other psychiatric disorders. The results from genome-wide association study (GWAS) in psychiatric genetics have confirmed that common psychiatric diseases are not Mendelian disorders, i.e. influenced by one or a few major genes, but are influenced by many genetic variants, each with a very small effect on risk for disease.[3] In other words: most common psychiatric diseases are polygenic. The GWAS of preschool internalizing problems, as presented in this dissertation, and other work in childhood psychiatry (ADHD, ASD, internalizing)[4-6] have shown that childhood psychiatric disorders are equally polygenic. This signifies that the lack of genome-wide significant loci in these GWAS projects is due to a limited sample size.

What have GWAS studies told us so far? Findings related to the genetic architecture of psychiatric disorders.

An early criticism of GWAS, which still is heard today, concerns the "missing heritability", i.e., the discrepancy between the heritability as established in twin and family studies, and the heritability attributable to measured genetic variants, which are associated with the phenotype of interest. Specifically, the genetic variants associated with psychiatric disorders, as identified in GWAS, tend to explain only a small part of the phenotypic variance, e.g. 7% on the liability scale for schizophrenia and 16% for height.[2; 7] This applies to psychiatric as well as non-psychiatric disorders and traits. One of the main causes of "missing heritability" is a lack of power to detect all genetic variants related to a trait. The polygenic model implies that large sample sizes are required to detect individual genetic variants. Also, one has to bear in mind that the significance level in a GWAS is typically set at $5*10^{-8}$. Given the small individual effects sizes, a few associations (or even a substantial number of associations) cannot be expected to explain

much variance. From the power perspective, an arguably more realistic question is how much of the phenotypic variance the genetic variants collectively explain. This question has been addressed by means of a statistical method called genomic relationship matrix restricted maximum likelihood, or GREML as implemented in the software genome-wide complex trait analysis (GCTA), which was developed by Yang et al[8]. GCTA provides an estimate of the variance explained by the total set of genetic variants, without identifying the individual associated variants. GCTA has been applied to psychiatric disorders by different groups[9; 10] and consortia,[11] and has shown that a substantial part of the "missing heritability" can be accounted for by genetic variants, as genotyped on modern GWAS platforms, i.e., common single nucleotide polymorphisms (SNPs). This supports the notion that the success of GWAS heavily depends on adequate sample size. While significantly associated SNPs explain only a modest proportion of variance in height and schizophrenia, GREML analyses showed that all measured SNPs explain a more substantial 23% of variance in schizophrenia and 49.8% of variance in height.

The substantive proportion of genetic variance accounted for by measured SNPs (as estimated using GCTA) does not completely account for the "missing heritability", as there remains a substantial discrepancy between the heritability estimates of twin and family studies and the genetic variance accounted for by SNPs . This accounts for adult and childhood psychiatric traits.[12-14] The fact that a discrepancy remains, i.e., the part of the heritability remains "missing", is not surprising. Genotyping platforms were never designed to capture all variation in the genome, but to quickly and cheaply tag a substantial portion of the genetic variation.[15] By design there is an upper limit to the portion of genetic variation SNPs can explain. The remaining "missing heritability" may be attributed to poor tagging of common causal variants other than SNPs, for example Variable Numbers of Tandem Repeats (VNTRs), the role of genetic rare variants (which are invariably poorly tagged), and structural genetic variants (copy number variation).[11]

Three recent papers discussed additional methodological and phenotypic causes of the residual missing heritability. Golan et al. point out that ascertainment bias will influence results when applying GCTA to a dichotomous trait.[16] Ascertainment of cases for GWAS at rates substantially higher than in the population leads to an underestimation of the variance explained by all common SNPs. This may have

resulted in an underestimation of the variance explained by SNPs for less common psychiatric disorders, such as schizophrenia and bipolar disorder, where the degree of ascertainment is more severe. A second potential cause of missing heritability is phenotypic or disease heterogeneity.[17] Here we define disease heterogeneity as multiple genetically distinct diseases causing the same or similar configurations of symptoms. Within a family a single distinct genetic disease would likely cause increased symptoms in multiple family members, therefore the heritability estimated in family and twin studies may actually be the mean heritability of the different genetic diseases observed in the population of twins and families in the study. Estimation of the variance explained by all genetic variants, using techniques like GCTA, may underestimate the SNP heritability, as unrelated cases do not necessarily suffer from the same underlying genetic disease. Heterogeneity is also evident at the phenotypic level when considering longitudinal data. As Chapter 3 of this dissertation reveals, there are many distinct ways to develop psychopathologies throughout childhood and adolescence. It is reasonable to ask whether developmentally distinct trajectories of maladaptive behavior display only partial genetic overlap, and thus may contribute to missing heritability when clustered as a single disorder or cluster of disorders. Lubke et al.[10] showed that different subscales of a common borderline questionnaire are characterized by substantial differences in the amount of variance explained by SNPs.

Not only can GCTA be applied to estimate the genetic variance attributable to measured SNPs (i.e., the SNP heritability) of individual traits. It has been extended to bivariate analysis allowing estimation of co-heritability based on measured SNPs.[11] The psychiatric genomics consortium (PGC) has shown that bipolar disorder, major depressive disorder (MDD) and schizophrenia are significantly genetically correlated. MDD and Attention Deficit Hyperactivity Disorder (ADHD) are also genetically correlated, while ADHD and Bipolar, and ADHD and Schizophrenia do not seem to be genetically correlated[11]. These studies explain substantial parts of missing heritability, and provide estimates of genetic correlations, independent of those obtained in twin and family studies.

The continuing relevance of twin studies in the era of GWA studies

While computational facilities to model measured genetic variants are quickly expanding, twin studies currently remain the most cost effective and computationally feasible to capture all genetic variance. Twin studies also allow for elaborate phenotypic modeling, because a wealth of data is available[18; 19] and computationally complex models can be used. Chapter 2 of this dissertation sheds light on the longitudinal genetic etiology of anxious depression symptoms, another study similarly shed light on the genetic and environmental etiology of attention problems.[20] These twin studies identified longitudinal genetic covariance in a way that cannot yet be replicated using measured genetic variants. Development of increasingly informative twin models is still ongoing.[21; 22]

Research performed in genotyped twins can also add a wealth of information to studies. Two specific examples show the continued value of twin studies. Van Dongen, Heijmans, Nivard et al. (*in preparation*) concurrently estimated in a genotyped cohort of twins and their families the twin heritability and the SNP heritability. This quantifies the ratio of variance tagged by SNPs to the total additive genetic variance. Van Dongen et al. went beyond this, and fitted GxE models, developed in chapter 7 of this thesis, in which the heritability of over 400 thousand methylation sites is modeled as a function of age and sex. This type of analysis reveals a wealth of information that is of interest to researchers who study the link between gene methylation and psychiatric disorders. The study by Buil et al.[23] is a second study that reveals the continued usefulness of twin studies in the genomic era. In a complex twin design, they identify different genetic influences on allele specific gene expression. Their model decomposes the genetic variance in allele specific gene expression into a part explained by known loci (eQTLs), a part explained by cis effects or local genetic effects, a part explained by trans or distant genetic effects, and a part attributable to the interaction between cis and trans genetic effects. This study could in principle have been executed in a sample of distantly related genotyped individuals. However, this would result in an extreme loss in power relative to the twin design. By leveraging the twin design, Buil et al.[23] managed to thoroughly explore the genetic influences on gene expression. Twin research will remain valuable, not only as a way to estimate and model all genetic variance, but also to enrich molecular genetics research, as discussed above.

The future use of identified loci: polygenic risk scores.

Psychiatric genetics has witnessed many successes in the past few years. The biggest success in terms of identifying loci has concerned adult psychopathology, but few doubt that increased sample sizes will eventually yield loci for childhood psychiatric disorders. The SNP heritabilities provide an upper limit of what can be extracted from genome-wide genotype data. Work in this dissertation has shed light on the link between childhood and adult psychopathology. Evidence from twin studies points to genetic causes of continuity of psychiatric problems, and other traits between childhood and adulthood.[20; 24] Future work will identify risk variants that influence childhood and (or) adult psychopathologies. But what is the ultimate relevance of these findings? The increased sample sizes of present consortia will certainly result in the identification of novel genetic variants associated with disease. But how can we utilize these variants?

One use lies in the computation of polygenic scores, i.e., weighted linear combinations of subsets of genetic variants, which are associated with the phenotype of interest (but not necessarily at the genome wide alpha level). Using polygenic scores, we have explored the link between adult schizophrenia and childhood and adolescent psychiatric disorders. As discussed in chapter 5, the results of our meta-analysis, obtained in 2 large population based cohorts, revealed that the schizophrenia risk score was associated at nominal significant p-values ($p < 0.05$) with measures of anxiety or depression and at an FDR corrected level of significance (adjusted $p < 0.05$) with anxiety at age 10. We concluded that an elevated genetic risk for schizophrenia is not broadly associated with all scales measuring the risk of childhood psychiatric disorders. The results suggest a relationship between schizophrenia risk and childhood internalizing problems (depression and anxiety) , but this issue requires further study. Others have established the genetic link between schizophrenia and cannabis use,[25] , schizophrenia and MDD, bipolar disorder, and autism spectrum disorder[5], adult educational attainment and childhood school performance, and attention problems.[26] These studies are a small selection from the broad literature (for review see[13]) that has revealed, using polygenic scores, a genetic overlap between psychiatric disorders and many other disorders and traits. Polygenic risk scores based on tagging SNPs, or in the future causal SNPs, can also facilitate the search for copy number variations (CNV) related to psychiatric disorders. An additive

model of genetic effects suggests that low polygenic scores in cases would increase the chance of identifying deleterious CNV's. Specifically targeting cases with a low polygenic risk score for the disorder they are suffering from might heighten the likelihood of finding deleterious CNV's.

Another future use of identified genetic variants, or polygenic risk scores, is in the study of environmental risk factors for disease. For example, teenage and early adult cannabis use has repeatedly been implicated in the risk of developing schizophrenia.[27; 28] At the same time risk loci for schizophrenia are associated with cannabis use.[25] This suggests that the predisposition to cannabis use and the predisposition to schizophrenia are genetically linked. This genetic link could explain (part of) the association between cannabis use and schizophrenia, without the need to infer a causal role for cannabis use in the onset of schizophrenia. To fully understand the risk cannabis use poses in the onset of schizophrenia, it is important to avoid confounding by controlling for genetic risk for schizophrenia. If genetic effects that increase cannabis use, also increase the risk for schizophrenia, these genetic effects could be included in the model when estimating the effect of cannabis use on later risk to develop schizophrenia. Similar confounding relationships possibly exist in adult and childhood psychiatry, and their discovery relies in part on identification of risk variants, or sufficiently accurate polygenic risk scores.

An often-mentioned application of genetic risk factors, either polygenic scores or identified loci, is prediction of disease status in clinical settings. While accurate clinical risk prediction would clearly be greatly useful, it is beyond the possibilities offered by currently available genetic data. The best possible genetic risk prediction remains the concordance between monozygotic twins. As monozygotic twins share all of their genetic variance barring rare *de novo* mutations,[29; 30] MZ twin concordance is very near the upper limit of genetic prediction. GWAS chips cover a substantial, but incomplete, part of all genetic variation. Prediction based on GWAS platforms will therefore necessarily underperform compared to MZ concordance rates. Clinical genetic risk prediction in psychiatry will for the foreseeable future not exceed the predictive value of an exhaustive family history. This is all the more true for child and adolescent psychiatry where GWAS samples are generally smaller, and thus predictive accuracy is lower.

A different downstream application of GWAS results is evident in the incorporation of GWAS results pertaining to MDD, bipolar disorder, and schizophrenia in an integrated pathway analysis. This analysis suggested specific biological pathways in the etiology of psychiatric disorders.[31] Specifically, this pathway analysis implicated the neuronal synapse, histone methylation, and immune pathways in psychiatric disease. Targeted investigation of these pathways will improve our understanding of the biology and the etiology of psychiatric disease. However, the number of distinct biological pathways that influence psychiatric illness is likely to be large. Just like there is likely no "gene for disorder X"[32] there is likely to be no "Pathway for disorder X".

Future developments in modeling genetic data: systems biology.

Genetic data have broader application than the identification of risk loci. Above I discussed methods for polygenic analyses and its application to psychiatric disorders. In chapter 7 I further discussed an extension of the model implemented in GCTA that allows for the moderation of the genetic effect of all SNPs on a trait. These moderating variables can be environmental (family environment, trauma exposure), biological (metabolite levels, gene expression intensity), or genetic (known risk loci).

In chapter 8, I presented a method to estimate genetic variance and covariance in the context of GREML. The aims of this method are not new as others have estimated genetic covariance in the GREML context.[33; 34] However the method I employed is amenable to parallelization, and is computationally lighter than bivariate GCTA or GEMMA.[34; 35] For example, as demonstrated in chapter 8, the estimation of a genetic covariance matrix of 24 items from a personality questionnaire does not pose a great computational problem. The results revealed that, although the individual items have a very low SNP heritability, one can still obtain a good estimate of the genetic covariance structure of these items based on measured genotypes. The method allows for the genetic effect to be separated into multiple genetic covariance matrices that reflect the effect of a subset of genes, pathways or SNPs from different functional categories (i.e. coding, non-coding in the gene body, in the gene promoter). Recently, a method has been developed to estimate SNP heritability on the basis of GWAS meta analysis summary statistics.[36] This method, LD score regression, can also be extended to

multivariate analysis[37] and allows for partitioning of the genome into multiple genetic effects based on functional categories.[38] The first analysis using LD regression has already revealed novel genetic correlations, and biological insights into the genetic etiology complex traits. LD regression has the benefit that it is computationally less burdensome than GREML and does not require the raw genotype data. However, if the genetic data are available GREML has some benefits that might well be worth the extra computational burden. For example, the model underlying GCTA can be used to improve predictive accuracy of risk loci by modeling multiple correlated psychiatric traits concurrently.[39]

The linear mixed model underlying the GREML model and our version of the model are flexible. Currently relatedness matrices used in these models are based on SNPs. However, there is no reason to limit the variables considered to SNPs. For example, researchers within the NTR are working on including genome wide measures of methylation and of gene expression into the linear mixed model. This type of modeling will allow for the integration of genetic and proteomic information into a single model, facilitating further understanding of the etiology of psychiatric disorders and symptoms.

Final thoughts.

The work in this dissertation was conducted in a period of rapid change in the field of psychiatric genetics. For schizophrenia and bipolar disorder the last few years have seen identification of large amounts of risk loci. In the next generation of GWAS mega- and meta-analysis, other psychiatric disorders will follow. These risk variants will help us better understand the genetic relationships between traits and disorders. Progress in genetic linear mixed modeling, and "LD score regression" will allow for ever more complex models of behavior to be fitted. These techniques will allow analysis of sub-disorder constructs such as symptoms and indicators. At the same time further integration of genetic, epigenetic, proteomic, and metabolomic data will help our understanding of the mechanisms, which relate risk variants to biochemistry, brain chemistry, and ultimately disease.

What is the ultimate goal of these ever increasingly complex models of psychiatric disorders? As our understanding of the pathways between gene and behavior grow, genetics might begin to inform our definition of disorders. If a distinct set of variants is found to influence a distinct set of pathways and these in turn

influence some, but not all symptoms of a disorder, then a disorder definition should be changed. This process of improving disorder definitions based on rigorous empirical genetic findings and subsequent genetic analysis of the improved disorder construct can form an iterative process bringing us closer to more homogeneous disorder definitions.[40; 41]

References

1. Stefansson H., Ophoff R.A., Steinberg S., Andreassen O.A., Cichon S., Rujescu D., Werge T. et al. (2009) Common variants conferring risk of schizophrenia. Nature *460*, 744-747

2. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014) Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421-427

3. Kendler K.S., O'Donovan M.C. (2014) A Breakthrough in Schizophrenia Genetics. JAMA psychiatry *71*, 1319-1320

4. Neale B.M., Medland S.E., Ripke S., Asherson P., Franke B., Lesch K.P., Faraone S.V. et al. (2010) Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. Journal of the American Academy of Child & Adolescent Psychiatry *49*, 884-897

5. Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. The Lancet *381*, 1371-1379

6. Trzaskowski M., Eley T.C., Davis O.S., Doherty S.J., Hanscombe K.B., Meaburn E.L., Haworth C.M. et al. (2013) First genome-wide association study on anxiety-related behaviours in childhood. PloS one *8*, e58676

7. Wood A.R., Esko T., Yang J., Vedantam S., Pers T.H., Gustafsson S., Chu A.Y. et al. (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. Nature genetics *46*, 1173-1186

8. Yang J., Lee S.H., Goddard M.E., and Visscher P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. The American Journal of Human Genetics *88*, 76-82

9. Lubke G.H., Hottenga J.J., Walters R., Laurin C., De Geus E.J., Willemsen G., Smit J.H. et al. (2012) Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. Biological psychiatry *72*, 707-709

10. Lubke G.H., Laurin C., Amin N., Hottenga J.J., Willemsen G., van Grootheest G., Abdellaoui A. et al. (2014) Genome-wide analyses of borderline personality features. Molecular psychiatry *19*, 923-929

11. Cross-Disorder Group of the Psychiatric Genomics Consortium (2013) Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nature genetics *45*, 984-994

12. Benke K.S., Nivard M.G., Velders F.P., Walters R.K., Pappa I., Scheet P.A., Xiao X. et al. (2014) A genome-wide association meta-analysis of preschool internalizing problems. Journal of the American Academy of Child & Adolescent Psychiatry *53*, 667-676

13. Wray N.R., Lee S.H., Mehta D., Vinkhuyzen A.A., Dudbridge F., and Middeldorp C.M. (2014) Research Review: Polygenic methods and their application to psychiatric traits. Journal of Child Psychology and Psychiatry *55*, 1068-1087

14. Trzaskowski M., Dale P.S., and Plomin R. (2013) No genetic influence for childhood behavior problems from DNA analysis. Journal of the American Academy of Child & Adolescent Psychiatry *52*, 1048-1056

15. Visscher P.M., Brown M.A., McCarthy M.I., and Yang J. (2012) Five years of GWAS discovery. The American Journal of Human Genetics *90*, 7-24

16. Golan D., Lander E.S., and Rosset S. (2014) Measuring missing heritability: Inferring the contribution of common variants. Proceedings of the National Academy of Sciences *111*, E5272-E5281

17. Wray N.R., Maier R. (2014) Genetic Basis of Complex Genetic Disease: The Contribution of Disease Heterogeneity to Missing Heritability. Current Epidemiology Reports *1*, 220-227

18. van Beijsterveldt C.E., Groen-Blokhuis M., Hottenga J.J., Franic S., Hudziak J.J., Lamb D., Huppertz C. et al. (2013) The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. Twin Research and Human Genetics *16*, 252-267

19. Willemsen G., Vink J.M., Abdellaoui A., den Braber A., van Beek J.H., Draisma H.H., van Dongen J. et al. (2013) The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. Twin Research and Human Genetics *16*, 271-281

20. Kan K.J., Dolan C.V., Nivard M.G., Middeldorp C.M., van Beijsterveldt C.E., Willemsen G., and Boomsma D.I. (2013) Genetic and environmental stability in attention problems across the lifespan: evidence from the Netherlands twin register. Journal of the American Academy of Child & Adolescent Psychiatry *52*, 12-25

21. Molenaar D., Dolan C.V. (2014) Testing systematic genotype by environment interactions using item level data. Behavior genetics *44*, 212-231

22. de Kort J.M., Dolan C.V., Kan K.J., van Beijsterveldt C.E., Bartels M., and Boomsma D.I. (2014) Can GE-Covariance Originating in Phenotype to Environment Transmission Account for the Flynn Effect? Journal of Intelligence *2*, 82-105

23. Buil A., Brown A.A., Lappalainen T., Viñuela A., Davies M.N., Zheng H.F., Richards J.B. et al. (2015) Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. Nature genetics *47*, 88-91

24. Nivard M.G., Dolan C.V., Kendler K.S., Kan K.J., Willemsen G., van Beijsterveldt C.E., Lindauer R.J. et al. (2014) Stability in symptoms of anxiety and depression as a function of genotype and environment: a longitudinal twin study from ages 3 to 63 years. Psychological medicine, 1-11

25. Power R.A., Verweij K.J.H., Zuhair M., Montgomery G.W., Henders A.K., Heath A.C., Madden P.A.F. et al. (2014) Genetic predisposition to schizophrenia associated with increased use of cannabis. Molecular psychiatry *19*, 1201-1204

26. de Zeeuw E.L., van Beijsterveldt C.E., Glasner T.J., Bartels M., Ehli E.A., Davies G.E., Hudziak J.J. et al. (2014) Polygenic scores associated with educational attainment in adults predict educational achievement and ADHD symptoms in children. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics *165*, 510-520

27. Zammit S., Moore T.H., Lingford-Hughes A., Barnes T.R., Jones P.B., Burke M., and Lewis G. (2008) Effects of cannabis use on outcomes of psychotic disorders: systematic review. The British Journal of Psychiatry *193*, 357-363

28. Smit F., Bolier L., and Cuijpers P. (2004) Cannabis use and the risk of later schizophrenia: a review. Addiction *99*, 425-430

29. Abdellaoui A., Ehli E.A., Hottenga J.J., Weber Z., Mbarek H., Willemsen G., van Beijsterveldt T. et al. (2015) CNV Concordance in 1,097 MZ Twin Pairs. Twin Research and Human Genetics, 1-12

30. Ehli E.A., Abdellaoui A., Hu Y., Hottenga J.J., Kattenberg M., van Beijsterveldt T., Bartels M. et al. (2012) De novo and inherited CNVs in MZ twin pairs selected for discordance and concordance on Attention Problems. European Journal of Human Genetics *20*, 1037-1043

31. Network T., Pathway Analysis Subgroup of the Psychiatric Genomics Consortium (2015) Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nature Neuroscience, 199-209

32. Kendler K.S. (2005) " A gene for...": the nature of gene action in psychiatric disorders. The American journal of psychiatry *162*, 1243

33. Lee S.H., Yang J., Goddard M.E., Visscher P.M., and Wray N.R. (2012) Estimation of pleiotropy between complex diseases using single-

nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics *28*, 2540-2542

34. Zhou X., Stephens M. (2014) Efficient multivariate linear mixed model algorithms for genome-wide association studies. Nature methods *11*, 407-409

35. Lee S.H., Yang J., Goddard M.E., Visscher P.M., and Wray N.R. (2012) Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics *28*, 2540-2542

36. Bulik-Sullivan B., Loh P.R., Finucane H., Ripke S., Yang J., Patterson N., Daly M.J. et al. (2014) LD score regression distinguishes confounding from polygenicity in genome-wide association studies. bioRxiv, 002931

37. Bulik-Sullivan B., Finucane H.K., Anttila V., Gusev A., Day F.R., Perry J.R., Patterson N. et al. (2015) An Atlas of Genetic Correlations across Human Diseases and Traits. Nature genetics, 014498

38. Finucane H.K., Bulik-Sullivan B., Gusev A., Trynka G., Reshef Y., Loh P.R., Anttilla V. et al. (2015) Partitioning heritability by functional category using GWAS summary statistics. bioRxiv, 014241

39. Maier R., Moser G., Chen G.B., Ripke S., Coryell W., Potash J.B., Scheftner W.A. et al. (2015) Joint Analysis of Psychiatric Disorders Increases Accuracy of Risk Prediction for Schizophrenia, Bipolar Disorder, and Major Depressive Disorder. The American Journal of Human Genetics

40. Kendler K.S., First M.B. (2010) Alternative futures for the DSM revision process: iteration v. paradigm shift. The British Journal of Psychiatry *197*, 263-265

41. Kendler K.S. (2009) An historical framework for psychiatric nosology. Psychological medicine *39*, 1935-1941

Appendix I: Statistical supplement to chapter 2: Stability in symptoms of anxiety and depression as a function of genotype and environment: A longitudinal twin study from age 3 to 63 years

Stability and change: Genetic simplex

A genetic simplex model (Boomsma and Molenaar, 1987) formed the basis to analyze stability and change in symptoms of anxiety and depression (SxAnxDep). Genetic simplex models include latent genetic and environmental factors that underlie the observed outcome variables (phenotypes). Applying the genetic simplex model results in a linear genetic decomposition of phenotypic variance into genetic and non-genetic sources. Such a decomposition is possible in a twin design using the fact that monozygotic (MZ) twins are (nearly) 100% genetically identical, while dizygotic (DZ) twins share on average 50% of their segregating genes (Boomsma *et al.* 2002; Middeldorp and Boomsma, 2009). The genetic simplex makes further use of the information captured by the MZ and DZ twin correlations across age (eTables 1-3), that is the correlation for SxAnxDep of twin 1 at for example age 3 and SxAnxDep of twin 2 at for example age 7. If these correlations are substantial (and positive), and if the twin correlations across ages are stronger in MZ twins than in DZ twins (eFigure 2), this suggest that phenotypic stability is (at least in part) due to genetic factors, or in other words the genetic factors at age t are correlated with the genetic factors at the subsequent age at which we measure SxAnxDep (t+1). In a similar vein, environmental factors may be correlated across age (the extent to which the MZ correlations across age differ from the within person correlations across age is an indication for the importance of environmental stability).

In the genetic simplex model, correlations among latent genetic and environmental factors are specified through an autoregression process (see Figure 1 chapter 2).

Specifically, SxAnxDep at each age t was modeled as follows:

$SxAnxDep_t = A_t + E_t + m_t + res_t + \lambda_{Ct} C,$

where A_t , E_t and C denote the (zero-mean) additive genetic, nonshared (unique) environmental and shared environmental values, m_t the grand mean of SxAnxDep, and res_t the residual term. Subscript t denotes age t=1,…30, where t=1 is the first age bin for which we observed outcome data (age 3-4) and t=30 is the last age bin (61-63 years). Here, A and E are modeled as autoregressive processes and C is modeled as a common factor (see below).

The autoregressive patterns were modeled as:

$A_t = \beta_{At} * A_{t-1} + \zeta_{At}$,

$$E_t = \beta_{Et} * E_{t-1} + \zeta_{Et},$$

where coefficient β_{At} and β_{Et} denote the autoregression coefficients, reflecting that genetic and nonshared environmental effects present at the previous age (t-1) influence the next age (t) and are thus carried over. Terms ζ_{At} and ζ_{Et} denote additive genetic and nonshared effects that arise at age t, and which are called innovations. Effects of innovations are thus not present earlier than age t, but are carried over to subsequent ages.

To model the effects of shared environment in childhood at age 7, 10 and 12, we extended the AE simplex model by allowing the outcome measures at these ages to also load on a common, shared environmental factor: $SxAnxDep_t=\lambda_{Ct}$ C, where λ_{Ct} represents the factor loading at ages 7, 10 and 12 , and where C is a latent factor (see figure 1) with unit variance and $SxAnxDep_t$ is the phenotype at age t. At age 12, C influences maternal and self-ratings of SxAnxDep.

All genetic and environmental factors are assumed to be independent from each other, such that at each age:

$$Var(SxAnxDep_t) = Var(A_t) + Var(C_t) + Var(E_t) + Var(res_t),$$

where

$$Var(A_t) = \beta_{At}^2 * Var(A_{t-1}) + Var(\zeta_{At}),$$
$$Var(E_t) = \beta_{Et}^2 * Var(E_{t-1}) + Var(\zeta_{Et}),$$
$$Var(C_t) = \lambda_{Ct}^2 * Var(C) = \lambda_{Ct}^2$$

At each age, the residual variance $Var(res_t)$ in the phenotype was decomposed into genetic and environmental factors, according to a standard behavior genetic model:

$$res_t = res_{At} + res_{Et},$$

such that

$$Var(res_t) = Var(res_{At}) + Var(res_{Et})$$

This reflects the possibility that the age-specific phenotypic residuals include the effects of age-specific additive genetic, and nonshared environmental factors, i.e. these age-specific effects are not carried over to subsequent ages (and in this respect can be distinguished from innovations). Note that the component $Var(res_{Et})$ also includes variance due to measurement error.

We allowed for mean differences in SxAnxDep between males and females by modeling the observations (m_t) as a function of sex:

$$m_t = b_{0t} + b_{1t}*sex,$$

where b_{0t} denotes the intercept and b_{1t} the effect of dummy coded variable sex (0 = males, 1= females) at age t. The model, as implemented, did not allow for sex differences in heritabilities (Vink *et al.* 2012). To explore whether this was reasonable, the data were divided into 7 larger age bins (3 to 12, 12 to 19, 19 to 27, 27 to 35, 35 to 43, 43 to 53, and 53 to 63 years of age) and heritabilities were estimated for males and females separately each of the larger age bins. The larger bins did not show a difference in heritability between males and females but for age 35-43 (see figure 2B).

The complete model was implemented in Mplus version 6.11 (Muthén and Muthén, 1998-2010). All scripts may be obtained from the first author. Analyses were carried out using robust full information maximum likelihood, which allows for the presence of missing data and the analysis of non-normally distributed continuous outcome variables. Commonly, comparative model fits of models as the model above are evaluated using likelihood ratio tests (-2*log-likelihood difference between two models). However, given the oversensitivity for large sample sizes of this test (Tanguma, 2001) and the large sample size in the current study, model fit was assessed by Akaike's information criterion (AIC) and Bayesian information criterion (BIC).

Identification, trimming, and evaluation of the genetic simplex model

At the first age for which we have observed data, A_t and E_t obviously cannot be predicted by data from earlier ages, therefore at this age all variance in the latent additive genetic and latent environmental factor is due to innovation variance: $Var(A_1) = Var(\zeta_{A1})$, and $Var(E_1) = Var(\zeta_{E1})$.

Residual variances at the first and last age at which we have observed data are only identified when additional assumptions are made. The following constraints were added to the model: $Var(res_{A1})=Var(res_{A2})$, $Var(res_{E1})=Var(res_{E2})$, $Var(res_{A29})=Var(res_{A30})$, and $Var(res_{E29})=Var(res_{E30})$. To speed up computation and to avoid empirical identification problems, the variances of the residuals (both E

and A residuals) were constrained to be equal over three subsequent age bins: $Var(res_{A1})=Var(res_{A2})=Var(res_{A3})$, $Var(res_{A4})=Var(res_{A5})=Var(res_{A6})$, etc.

Based on the initial results of fitting the simplex model, and in order to arrive at an optimal, parsimonious model, we trimmed the model by allowing for further constraints. The autoregression parameters in the genetic simplex model (β_{At}) between the age of 18 and 63 all approached 1. Constraining these regression parameters to be 1 proved to be acceptable as judged by the AIC and BIC. We further observed that the innovations (ζ_{At}) of the genetic variables A from age 18 to age 63 were close to zero, and often did not reach significance ($\alpha=0.05$), given their standard error. Restricting these parameters to zero reduced the BIC slightly, but increased the AIC (see eTable 2). Effectively, these constraints reduced the genetic model equivalent to a single factor model from age 18 onwards, and rendered the model substantially more parsimonious, the results of which are reported in the main text.

From the final model, standardized coefficients of the contribution of genetic (and environmental) effects were derived for each age (t), including heritability coefficients (h^2).

$$h_t^2 = (Var(A_t)+ Var(res_{At}))/(Var(A_t) + Var(E_t) + Var(C_t) + Var(res_{At})+ Var(res_{Et}))$$

$$c_t^2 = Var(C_t)/(Var(A_t) + Var(E_t) + Var(C_t) + Var(res_{At}) + Var(res_{Et}))$$

$$e_t^2 = (Var(E_t)+ Var(res_{Et}))/(Var(A_t) + Var(E_t) + Var(C_t) + Var(res_{At})+ Var(res_{Et}))$$

Additional mean trend analyses.

Mean trends over age, and sex differences herein were analyzed by post-hoc analyses. Age trends were analyzed using weighted least squares (WLS) on the means. The trends fitted are visible as dashed lines in figure 2A. Model derived means of males and females are vectorized from age 3 to age 63 and added to a single variable (MV, mean vector). This mean vector is regressed on a sex dummy variable (S), a design variable (D1) that linearly increases between age 12 and 28 and a design variable that monotonically increases from age 30 to 63(D2). These variables are combined in the following regression equation:

$$MV = I + b_1*S + b_2*D1 + b_3*D2 + b_4* S*D1 + b_5 * S*D2$$

The parameter b_1 reflects the intercept difference between male and female means. This intercept, I, reflects the mean SxAnxDep score between the ages of 3 and 12. Parameter b_2 reflects the change in SxAnxDep between ages 12 and 28 for male participants. Parameter b_3 reflects the change in SxAnxDep for male participants between the ages of 30 and 63. Parameter b_4 reflects the change for female participants over and above the change of male participants between the ages of 12 and 28. Parameter b_5 reflects the change for female participants over and above the change of male participants between the ages of 30 and 63. The regression is weighted to reflect the differences in sample size used for the mean estimates at the different age points.

References

1. Boomsma, D., Busjahn, A. & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews. Genetics* 3, 872-882.

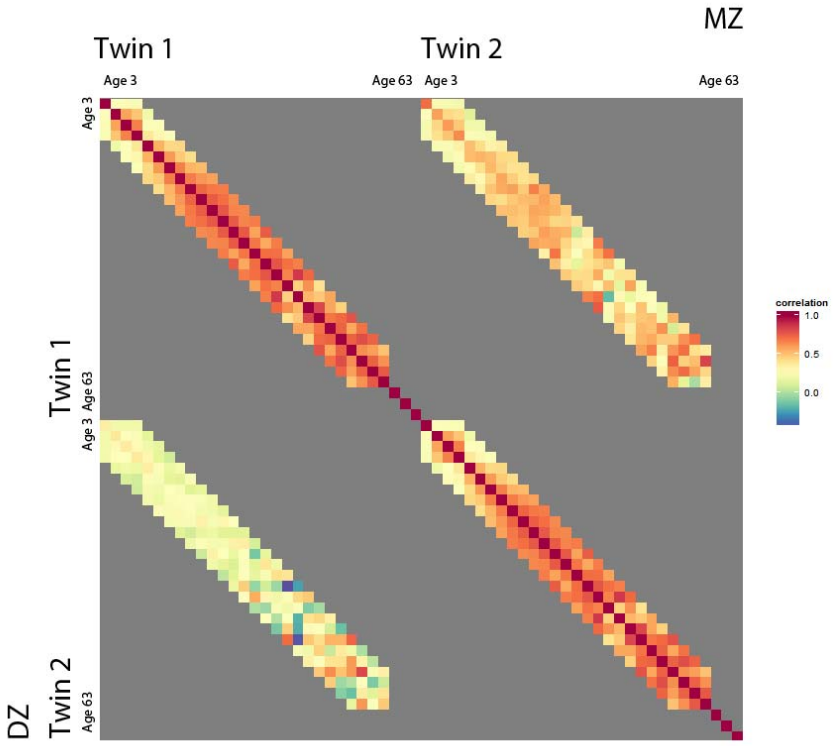
2. Boomsma, D.I. & Molenaar, P.C.M. (1987). The genetic analysis of repeated measures. I. Simplex models. *Behavior Genetics* 17(2), 111-123.

3. Middeldorp, C.M. & Boomsma, D.I. (2009). Genetics and Psychopathology. In *Handbook of Neuroscience for Behavioral Sciences*, (ed. G. G. Berntson and J. T. Cacioppo), pp. 1180-1202. Wiley: New York.

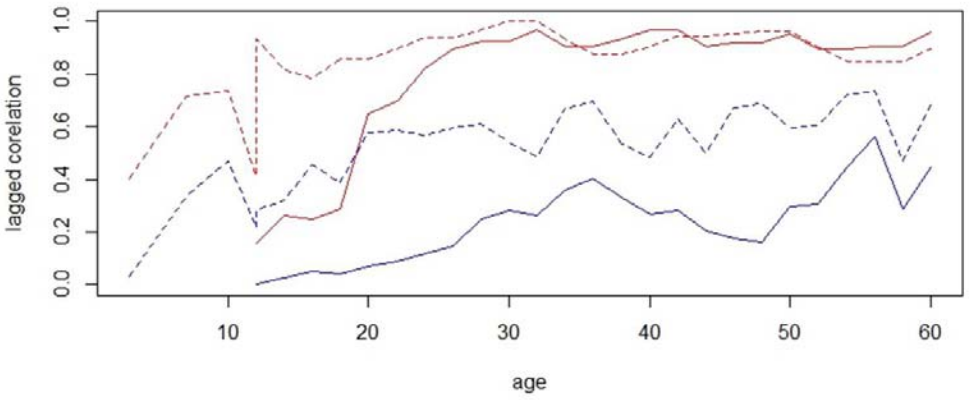
4. Muthén, L.K. & Muthén, B.O. (1998-2010). *Mplus User's Guide. Sixth Edition.* Muthén & Muthén: Los Angeles, CA.

5. Tanguma, J. (2001). Effects of Sample Size on the Distribution of Selected Fit Indices: A Graphical Approach. *Educational and Psychological Measurement* 61, 759-776

6. Vink J.M., Bartels M., van Beijsterveldt T.C., van Dongen J., van Beek J.H., Distel M.A., de Moor M.H., Smit D.J., Minica C.C., Ligthart L., Geels L.M., Abdellaoui A., Middeldorp C.M., Hottenga J.-J., Willemsen G., de Geus E.J., Boomsma D.I. (2012). Sex differences in genetic architecture of complex phenotypes? *PLoS One* 7(12), e47371



eFigure 1: Maximum Likelihood estimates of correlations within and across age for SxAnxDep. For each age, correlations are given with the 3 preceding and 3 subsequent ages. Diagonal: within-person longitudinal correlations. Above diagonal (right upper corner): MZ correlations. Below diagonal (left lower corner): DZ correlations. Within-person across-age correlations are higher than MZ twin across-age correlations, indicating a role for unique environment on stability. MZ across-age correlations are higher than DZ across-age correlations, indicating a role for genetic influences on stability.



eFigure 2: Pairwise genetic (A, red) and Environmental (E, blue) correlations between observations 2 years (dashed) and 10 years (solid) apart.

eTable 1: Number of subjects per number of repeated observations

Number of observations	1	2	3	4	5	6	7	8
Subjects	20987	9940	6588	7015	2812	1690	405	87

eTable 2: Model fitting results based on the longitudinal simplex model

Model:	Parameters	BIC	AIC	ΔBIC vs Full
Full AE simplex	199	564233.711	562629.486	-
A-transmission fixed at 1 in adulthood,	171	563989.625	562603.059	-244
A-innovations fixed at 0 in adulthood	156	563912.402	562654.819	-77

eTable 3: Twin correlations for monozygotic (MZ) and dizygotic (DZ) pairs at ages 3, 7, … , 62 years. The middle row gives the correlation estimate, and the first and third row give the upper and lower bound for the 95% confidence intervals.

MZ															
Age	**3**	**7**	**10**	**12**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**26**	**28**	**30**	**32**
Upper bound	.72	.60	.60	.65	.56	.59	.57	.61	.62	.55	.66	.64	.65	.56	.36
correlation	.70	.58	.58	.62	.48	.55	.53	.57	.57	.48	.60	.55	.56	.44	.20
lower bound	.69	.56	.55	.60	.40	.51	.48	.53	.51	.40	.52	.45	.45	.30	.02
Age	**34**	**36**	**38**	**40**	**42**	**44**	**46**	**48**	**50**	**52**	**54**	**56**	**58**	**60**	**62**
Upper bound	.60	.70	.65	.56	.64	.67	.69	.58	.83	.69	.77	.63	.59	.89	.67
correlation	.47	.58	.51	.34	.44	.52	.55	.37	.71	.51	.59	.36	.32	.76	.35
lower bound	.30	.43	.33	.08	.20	.33	.36	.11	.51	.25	.34	.02	-.02	.49	-.07
DZ															
Age	**3**	**7**	**10**	**12**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**26**	**28**	**30**	**32**
Upper bound	.36	.38	.37	.38	.40	.26	.30	.29	.29	.35	.27	.33	.25	.38	.34
correlation	.34	.36	.34	.36	.32	.22	.25	.23	.22	.27	.17	.21	.10	.22	.16
lower bound	.32	.34	.32	.33	.23	.17	.21	.18	.15	.19	.06	.09	-.05	.06	-.02
Age	**34**	**36**	**38**	**40**	**42**	**44**	**46**	**48**	**50**	**52**	**54**	**56**	**58**	**60**	**62**
Upper bound	.36	.48	.39	.34	.54	.62	.56	.69	.83	.58	.40	.59	.87	.70	.64
correlation	.13	.26	.12	.03	.27	.37	.31	.46	.62	.29	.05	.16	.57	.37	.19
lower bound	-.11	.02	-.17	-.29	-.05	.07	.01	.14	.25	-.06	-.31	-.33	-.04	-.08	-.36

Appendix II: Twin correlation and sampling strategy for AnxDep and AP as used in Chapter 7 Detection of gene –environment interaction in pedigree data using genome-wide genotypes.

In Chapter 7 the variance explained by SNPs and the additive genetic variance in anxious depression(AD) and Attention Problems(AP) was estimated. This required the biggest possible sample, as the technique used (GCTA) requires large samples to arrive at a precise estimate (Visscher PloS genetics; 2012). This required combining measurements collected in multiple surveys. Before arriving at a final phenotype in Chapter 7 we undertook the following exploration of the Anxious depressed and AP data available in the Netherlands Twin Register(NTR).

The Adult Netherlands Twin Registry (ANTR) has collected data in 10 Surveys between 1991 and 2013. The youth adult self report (YASR) was added to surveys in 1991, 1995, 1997, 2000, 2002, 2009, 2011 and 2013. The Young Netherlands Twin Registry(YNTR) has collected data on the YASR scale at age 14,16 and 18.

AP

Phenotype

Attention problems were measured using the (A)YSR scale. Different scales across surveys contained a different attention problems scale. The scale different both in number of items and item content. To account for the different number of items, we calculated a mean item score for each individual. AP measures were included in ANTR surveys in 1991, 1995,1997,2000,2009 and 2013. An AP

Twin correlations

We express the twin correlations per survey of the AP mean item score in Table 1 and Figure 1. The MZ correlations, both male and female consistently higher than the male, female and opposite sex pair DZ correlations. This indicates AP is a heritable.

Table 1 : Twin Correlations for AP mean items scores

			MZM	DZM	MZF	DZF	DOS
Survey 1	1991	cor	0.403	0.119	0.372	0.193	0.198
		N	275	230	371	289	464
Survey 3	1995	cor	0.444	0.111	0.437	0.23	0.017
		N	276	226	423	267	466
Survey 4	1997	cor	0.501	0.192	0.395	0.227	0.235
		N	218	145	402	254	319
Survey 5	2000	cor	0.395	0.079	0.443	0.076	0.19
		N	252	134	653	315	310
Survey 8	2009	cor	0.471	0.177	0.47	0.188	0.17
		N	337	169	1026	450	404
Survey 10	2013	cor	0.475	0.118	0.469	0.284	0.198
		N	233	109	665	232	278
YNTR DHBQ		cor	0.457	0.167	0.473	0.185	0.185
		N	928	764	1339	1033	1700

Table 2: number of genotyped participants for a given number of repeated observations of AP.

Number of obs	1	2	3	4	5	6
N	2362	1989	1338	482	303	144

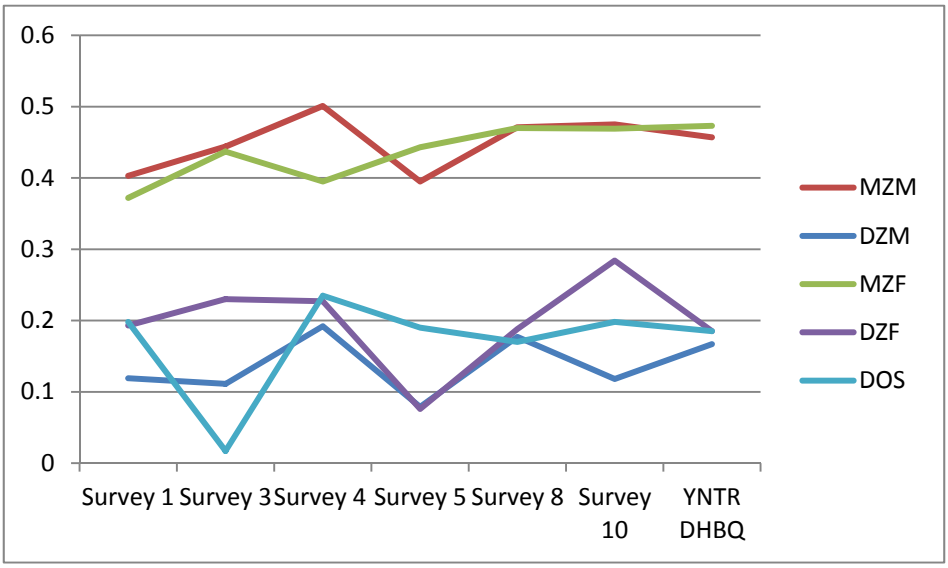


Figure 1: Plot of Twin correlation split over the (A)NTR surveys

GCTA: random sample versus most recent observation

Multiple repeated measures of AP are available per individual participant due to the longitudinal nature of the NTR. As one of the aim of analysis in Chapter 7 was gene by age interaction we sampled the most recently measured phenotype available from each participant. This sampling optimizes the variance of the age variable in the final data as observations on younger individuals are more abundant. However "Most recent observation" is not a random sample of the available data. We ran speed optimized version the intended model on 50 samples constructed of a random draw from the data point available for each participant. So if a participant has 3 data points available 1 is drawn at random, it another participant has 1 data point available, this data point is included. Table 2 gives the N for the AP data broken down to the number of repeated AP measurements available.

We compared the SNP heritability of the most recently obtained measure of AP for each participant to the SNP heritability of 50 sets where the observation for each individuals was randomly drawn from the measures available. The results are plotted in figure 2. For AP sampling the most recently available observation per participant did not result in dramatically different estimates of SNP heritability then randomly sampling an observation for each participant.

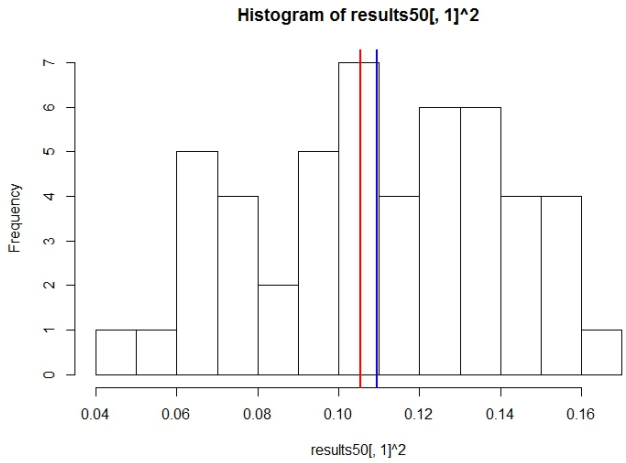


Figure 2: SNP heritability for AP for 50 randomly drawn univariate samples from the available longitudinal data and for the most recently obtained sample (red).

AnxDep.

Phenotype:

We specifically look at the Anxious Depression subscale of the YASR ,based on 13 items that are available in all version of the YASR used over the different ages across the YNTR and the ANTR.

Items used:

I feel lonely

I feel confused or in a fog

I cry a lot

I am afraid I might think or do something bad

I feel that I have to be perfect

I feel that no one loves me

I feel worthless or inferior

I am nervous or tense

I am too fearful or anxious

I feel too guilty

I am self-conscious or easily embarrassed

I am unhappy, sad or depressed

I worry a lot

Twin correlations per survey:

The correlations expressed below were calculated during data selection for the analysis in chapter 7. The analysis in chapter 7 does included all genotyped

individuals for which the phenotype was available. Chapter 7 therefore also includes parents of twins and siblings of twins. Chapter 7 also does not include all of the data points that contribute to these correlations as not all twins are genotyped. Below in table 3 the twin correlations for AnxDep per survey Note the oldest available measure in the YNTR is used to compute the twin correlations for the YNTR. The N here represent complete twin pairs for where for both twins a specific list is available.

Table 3: twin correlations per survey for AnxDep.

		MZ-males	DZ-males	MZ-females	DZ-females	Opposite sex dyzygotic pairs
Survey 1	cor	0.388	0.093	0.46	0.151	0.205
1991	N	273	229	370	286	463
Survey 3	cor	0.445	0.13	0.566	0.25	0.167
1995	N	273	225	421	265	458
Survey 4	cor	0.602	0.081	0.494	0.343	0.309
1997	N	217	144	400	251	318
Survey 5	cor	0.438	0.23	0.452	0.125	0.267
2000	N	250	132	642	303	303
Survey 6	cor	0.336	0.194	0.52	0.247	0.132
2002	N	232	102	615	295	288
Survey 8	cor	0.43	0.162	0.439	0.248	0.182
2009	N	385	203	1141	511	517
Survey 9	cor	0.387	0.079	0.417	0.262	0.197
2011	N	288	135	781	307	332
Survey 10	cor	0.409	0.098	0.425	0.181	0.129
2013	N	229	109	660	229	269
YNTR Survey	cor	0.408	0.245	0.491	0.22	0.229
	N	866	740	1294	998	1635

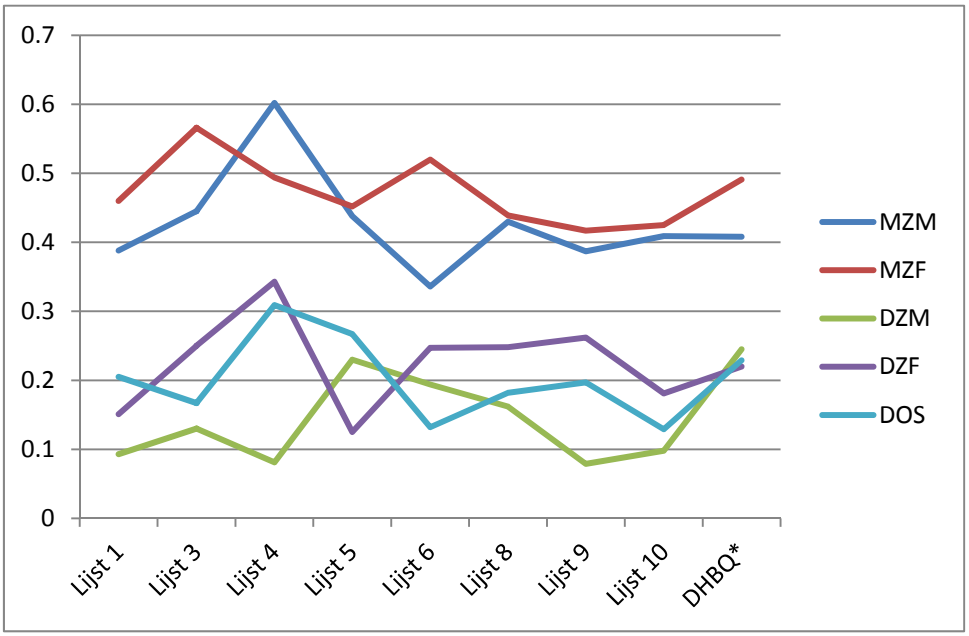


Figure 3: AnxDep twin correlations per survey.

GCTA random sampling versus most recent observation:

As the analysis in Chapter 7 requires a single phenotype where for some individuals multiple phenotypes are available due to the longitudinal nature of the NTR. In Figure 4 a histogram of the estimated SNP heritability in the 50 random samples. The mean of 50 repetitions (blue line) and our preferred sample (red)taking the most recent observation available are indicated in figure 2. For AnxDep sampling the most recently available observation per participant did not result in dramatically different estimates of SNP heritability then randomly sampling an observation for each participant.

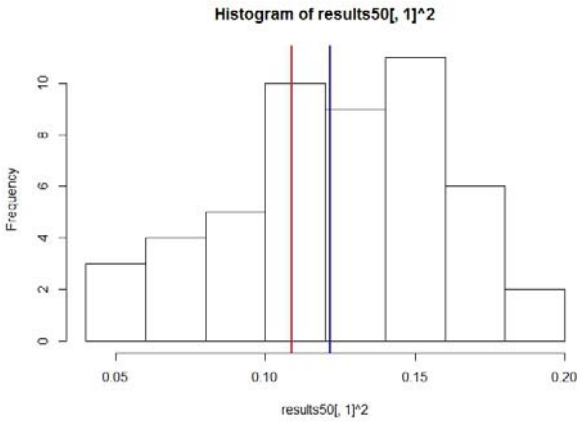


Figure 2: SNP heritability for Anxious Depression from randomly drawn univariate samples from the available repeated measures.

Apendix III: Genome wide association analysis of anxiety disorders

Background

NTR and NESDA take part in an international consortium (PI J Hettema) in which a meta-analysis will be done on the results of GWA studies focusing on identifying genetic variants for anxiety disorders. I performed the genome-wide association (GWA) study on "any anxiety disorder" which indicates the lifetime presence of generalized anxiety disorder, social phobia, panic disorder and/or agoraphobia. In this appendix I describe the analyses and result for the Dutch studies. The two parent projects that supplied information for this GWAS are large-scale longitudinal studies, the Netherlands Study of Depression and Anxiety (NESDA) and the Netherlands Twin Register (NTR)[1-3].

Participant and phenotype information

Recruitment of participants for NESDA took place from 09/2004-02/2007, and ascertainment was from outpatient specialist mental health facilities, primary care practices and the general population. Additional cases were from the NTR. Inclusion criteria were age 18-65 years, self-reported northwestern European ancestry a lifetime diagnosis of the following DSM-IV anxiety disorders as diagnosed via the Composite International Diagnostic Interview (CIDI) (version 2.1) [4] during one of the NESDA or NTR assessments: generalized anxiety disorder, social phobia, panic disorder and/or agoraphobia. Persons who were not fluent in Dutch and those with a primary diagnosis of a psychotic disorder, obsessive compulsive disorder, bipolar disorder, or severe substance use dependence were excluded.

Control subjects were mainly from the NTR. Longitudinal phenotyping includes assessment of depressive symptoms (via multiple instruments), anxiety, neuroticism, and personality measures. Inclusion for controls required a low score on the trait version of the STAI (State-Trait Anxiety Inventory) or on a composite measure of neuroticism, anxiety and depression [5-8]. A subsample of the NTR controls were also screened via a CIDI interview. A subset of controls were from NESDA and had no lifetime diagnosis of depressive or anxiety disorder as assessed by the CIDI.

There were 1521 cases and 2970 controls wit genotype and phenotype information.

NESDA and NTR studies were approved by the Central Ethics Committee on Research Involving Human Subjects of the VU University Medical Center, Amsterdam, an Institutional Review Board certified by the US Office of Human Research Protections (IRB number IRB-2991 under Federal wide Assurance-3703; IRB/institute codes, NESDA 03-183; NTR 03-180). All subjects provided written informed consent.

Genotyping, Imputation and Genome-Wide Association analysis

Whole blood and/or buccal DNA samples were collected for various projects done by the NTR and for the NESDA study (see [1-3;9]). Genotyping was performed on multiple platforms: Affymetrix 6.0, Affymetrix Perlegen 5.0, Illumina 370, Illumina 660, Illumina Omni 1M. Genotype calls were made with the platform specific software i.e. Birdseed, genotyper and Beadstudio.

QC was then performed within and between the different platforms and all genotypes were lifted over to build HG37 of the human genome. Genotypes that did not properly map to HG37 were removed as well as SNPs with a minor allele frequency below 1%, and allele frequency difference with the reference set above 20% , HWE < 0.00001 or a call rate below 95%.

IBD was calculated for all pairs and was compared to the expected pedigree structure. Samples where IBD did not match the expected pedigree were removed. Cross-platform concordance was calculated for samples that were genotyped on multiple platforms, samples that showed a concordance below 99% were removed.

Imputation was done in a two stage approach. First, the genotype platform specific SNPs were imputed using the MaCH software suite. Next, the 1000 genomes reference set SNPs were imputed using Minimach.

The genome-wide association analyses were performed on the dosage genotype data that were transformed into a single additive dosage score per SNP and imported into R. Subjects from non-Western European ancestry were not included in the analyses. Three logistic regression analyses were run, all including 3 principal components correcting for population stratification within the Netherlands and a dichotomous variable coding for study of origin as covariates. 1) without any additional covariates, 2) with sex and age included as covariates, 3) with sex, age and a SNP by sex interaction term as covariate. Minimal QC was performed before uploading the GWA results. SNPs were selected with an R^2 > 0.30, a MAF of 0.01 and HWE > $10e^{-5}$.

Results

No genome wide signals were observed in the NTR sample Below in figure 1 the Manhattan plot of the model that includes the effect of the SNP, and the main effects of sex and age. The genomic inflation after initial filtering in the NTR/NESDA sample is reasonable (Lambda = 1.021) indicating incusion of principal components did control for batch effects and stratification.

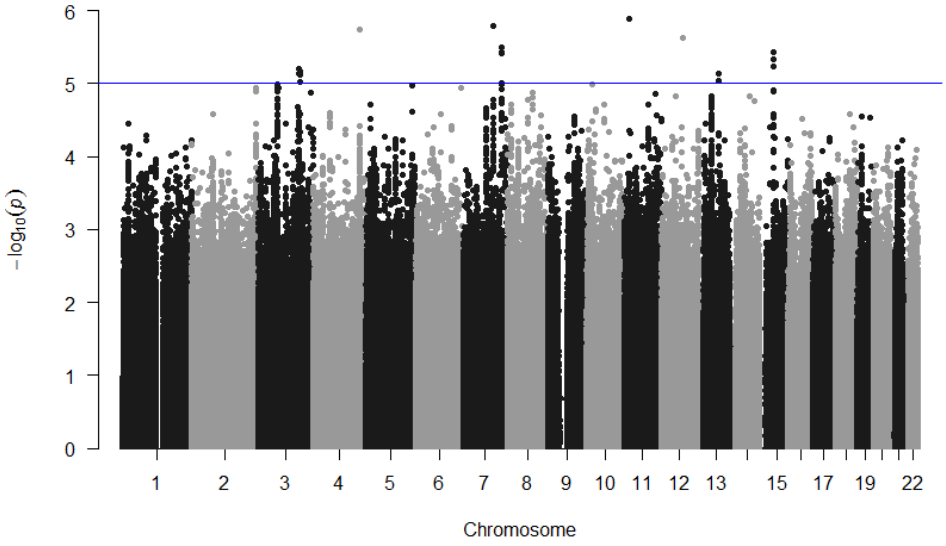


Figure 1: Manhattan plot for the NTR NESDA anxiety disorder case contol GWAS.

References

1. Boomsma DI, Sullivan PF, de Geus EJ et al. (2008) Genome-wide association of major depression: description of samples for the GAIN major depressive disorder study: NTR and NESDA Biobank Projects. European Journal of Human Genetics *16,* 335-342.

2. Penninx BW, Beekman AT, Smit JH et al. (2008) The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. International Journal of pscyhiatric research methods*. 17,* 121-140.

3. Willemsen G, Vink JM, Abdellaoui A et al. (2013) The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. Twin Reseach and Human Genetics , 16 271-281.

4. World Health Organization. (1992) Composite International Diagnostic Interview (version 2.1). Geneva: WHO.

5. Achenbach TM, Rescorla LA. (2003) Manual for the ASEBA Adult Forms & Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families, 2003.

6. Beck AT, Rial WY, Rickels K. (1974) Short form of depression inventory: cross-validation. Psychology Repeports *34,* 1184-1186.

7. Van der Ploeg H, Defares P.B., Spielberger C.D. (1979) *Zelfbeoordelingsvragenslijst STAI, versie DY-1 en DY-2.* Lisse: Swets & Zeitlinger, 1979.

8. Wilde GJS. (1970) Neurotische labiliteit gemeten volgens de vragenlijstmethode (The questionnaire method as a means of measuring neurotic instability). Amsterdam: Van Rossen, 1970.

9. Scheet P, Ehli EA, Xiao X et al. (2012) Twins, Tissue, and Time: An Assessment of SNPs and CNVs. Twin Research and Human Genetics 1-9.

Nederlandse Samenvating

In **hoofdstuk 1** van dit proefschrift wordende methoden uiteengezet die ten grondslag liggen aan gedragsgenetisch onderzoek. In gedragsgenetisch onderzoek toetst met in hoeverre individuele verschillen in fenotypes worden bepaald door individuele verschillen in genotypes. Ieder complex menselijk gedrag, of aandoening, kan doorgaans worden onderzocht als fenotype. De oudste vorm van gedragsgenetisch onderzoek geschiedt door middel van tweelingstudies, die de mogelijkheid bieden om te onderzoeken in hoeverre individuele verschillen tussen mensen (de populatie variantie) worden verklaard door genetische factoren en door de omgeving. Tweelingstudies bieden de mogelijkheid om de erfelijkheid van een fenoptype te onderzoeken omdat eeneiige tweelingen hun volledige DNA delen, en twee-eiige tweelingen is dit bij benadering de helft van het DNA. Als eeneiige tweelingen voor een bepaald fenotype meer op elkaar lijken dan twee-eiige tweelingen, dan is dit een aanwijzing dat individuele verschillen in dit fenotype door individuele verschillen in het genotype worden bepaald. Genetische factoren kunnen worden onderscheiden in “additief genetische effecten” en “non-additief genetische” of “dominantie effecten”. Bij effecten vanuit de omgeving kan een onderscheid worden gemaakt tussen omgevingsinvloeden die uniek zijn voor ieder kind binnen een tweelingpaar, en gedeelde omgevingsinvloeden, die in hun effecten hetzelfde zijn voor beide kinderen van een tweelingpaar. Op het moment dat tweelingonderzoek aantoont dat individuele verschillen deels verklaard worden door genetische factoren is vervolgens de vraag welke genen van invloed zijn.

Sinds de jaren ‘70 bestaat de mogelijkheid om door middel van genetische “linkage” (koppeling) de invloed van individuele genen of een individueel stuk genoom op gedrag te onderzoeken. Over het hele genoom delen twee-eiige tweelingen de helft van hun DNA, maar voor een bepaalde locus op het genoom zijn ze identiek, ze delen de helft, of ze zijn uniek. In koppelingsonderzoek wordt eerst bepaald of twee-eiige tweelingen, of broers en zussen, voor een bepaald stuk DNA identiek zijn, de helft van het DNA delen, of uniek zijn. Vervolgens word gekeken of de genetische overeenkomst op een bepaald stuk DNA tussen paren samengaat met overeenkomsten in gedrag, d.w.z. of de paren die het stuk genoom delen meer op elkaar lijken dan de paren die op dat stuk niets delen. Vanaf de jaren ‘90 word in de gedragsgenetica genoombreed koppelingsonderzoek uitgevoerd. Hierbij wordt op enkele honderden tot duizenden plekken over het hele genoom tegelijk geanalyseerd of verschillen in

het genoom samenhangen met verschillen in gedrag. Doordat men voor dit onderzoek broers en zussen gebruikt kan men met slechts enkele honderden tot duizenden metingen uitspraken doen over het hele genoom. Broers en zussen delen immers vaak lange aaneengesloten stukken genoom van een of beide ouders. Een eting van een enkel stukje DNA kan worden gegeneraliseerd over een langer stuk DNA, waar wel de zekerheid van de generalisatie afneemt naarmate de afstand met het gemeten stuk DNA toeneemt.

In 2000 werd de volledige sequentie van het menselijke genoom blootgelegd door het "Human Genome Project" wat de mogelijkheid bied om op een nog veel fijnere schaal genetisch onderzoek uit te voeren. In een genoombrede associatiestudie worden enkele honderdduizenden tot tientallen miljoenen "single nucleotide polymorphisms" (SNPs) onderzocht. De hoge dichtheid van de metingen betekent dat men deze associatiestudies ook kan uitvoeren in ongerelateerde individuen. Men in voor associatie niet meer afhankelijk van directe verwantschap tussen proefpersonen.

In de eerste genoombrede associatiestudies was het aantal significante associaties tussen gen en gedrag beperkt. Dit leidde tot de vraag of de veel voorkomende genetische varianten die in een dergelijke studie worden geanalyseerd wel van belang zijn bij dit soort fenotypes. Polygenetische analyses, waarbij het effect van al de genetische varianten gezamenlijk wordt onderzocht, liet zien dat deze SNPs wel degelijk een signaal herbergen, maar dat de effecten van de afzonderlijke SNPs zeer klein zijn. Dat betekent dat grote steekproeven nodig zijn om deze effecten te detecteren. Samenwerking tussen laboratoria over de hele wereld heeft inmiddels geleid tot de identificatie van honderden associaties tussen SNPs en menselijke fenotypes, zoals terug is te vinden in de GWAS catalogus (https://www.genome.gov/26525384). Het betreft bevindingen voor een groot aantal fenotypes zoals BMI, lichaamslengte, lipiden, hart- en vaatziekten, breinvolumes, maar ook psychiatrische ziekten zoals schizofrenie of bipolaire stoornis, en gedragsmaten zoals roken en neuroticisme.

Hoofdstuk 2 beschrijft een studie waarin met een longitudinaal tweelingmodel wordt gekeken hoe de genetische invloeden op symptomen van angst en depressie veranderen met leeftijd. Dit wordt ook wel "gen leeftijd interactie" genoemd. Hiervoor is een longitudinaal simplex model gebruikt, waarin de variatie op iedere leeftijd wordt gesplitst in genetische variatie en

omgevingsvariatie. De genetische en omgevingsvariaties worden verder opgedeeld in een component die op een eerdere leeftijd al invloed had op symptomen van angst en depressie en een invloed “innovatie” die nieuw is op deze leeftijd. Deze studie laat zien dat de genetische invloed op angst en depressie gedurende de levensloop toeneemt. Tot de leeftijd van 18 jaar is er dus steeds sprake van nieuwe genetisch effecten en deze toename stopt na het 18e levensjaar. De omgevingsinvloeden op symptomen van angst en depressie nemen ook toe tijdens de jeugd. Deze toename van omgevingsinvloeden is (tijdens de jeugd) groter dan die van de genetische invloeden. Dit leidt ertoe dat de erfelijkheid, gedefinieerd als de genetische variatie gedeeld door totale variatie, afneemt van de vroege jeugd tot de volwassenheid. Er is geen evidentie dat nieuwe genetische effecten op volwassen leeftijd invloed uitoefenen op symptomen van angst en depressie, terwijl er wel evidentie is voor nieuwe omgevingsinvloeden op volwassen leeftijd.

Hoofdstuk 3 beschrijft een studie waarin op populatieniveau groeicurven worden geschat voor de op DSM-IV gebaseerde internaliserende en externaliserende stoornissen. Deze groeicurven beschrijven de in deze populatie meest voorkomende ontwikkelingstrajecten voor externaliserend gedrag en internaliserend gedrag tussen leeftijd 7 en 18. Uiteindelijk beschrijft het model waarin er 5 discrete ontwikkelingstrajecten bestaan voor zowel externaliserend als internaliserend gedrag deze data het beste. Voor de internaliserende stoornissen is er evidentie voor trajecten met een “blijvend zeer lage symptoom score”, een “lage symptoom score”, “stijgende symptoom score”, “afnemende symptoom score” en een “in de adolescentie toenemende symptoom score”. Voor externaliserende stoornissen zijn er ook de eerste vier trajecten. In plaats van een traject van de “in de adolescentie toenemende scores” is er evidentie voor een traject met “blijvend hoge externaliserende scores”. Covariaten gemeten rond de zwangerschap, zoals bijvoorbeeld het roken van de moeder, verhogen de kans om in een hoge of stijgende externaliserende klasse terecht te komen, ook mannen zitten vaker in een stijgende of hoge externaliserende klasse. Vrouwen hebben een verhoogde de kans om in een hoog scorende internaliserend traject te worden geclassificeerd. Na de internaliserende en externaliserende trajecten te hebben vast gesteld, onderzoeken we comorbiditeit tussen internaliserende en externaliserende trajecten. Hier kijken we of bepaalde internaliserende trajecten vaker dan verwacht samenhangen met bepaalde externaliserende trajecten.

Vergelijkbare trajecten komen vaker dan verwacht samen voor: kinderen die in toenemnde mate internaliserende problemen hebben gaat vaak samen met toenemende externaliserende problemen, hetzelfde geldt voor afnemende problemen. Het internaliserende traject voor individuen die in de adolescentie een stijgende score laten zien blijkt onafhankelijk van de hoge en stijgende externaliserende trajecten. Deze studie laat zien dat eventuele comorbiditeit moet worden overwogen als een kind of adolescent wordt gediagnosticeerd met een externaliserende of internaliserende stoornis, niet alleen ten tijde van het stellen van de diagnose, maar ook later in het traject.

Hoofdstuk 4 beschrijft een genoombrede associatie (genome wide association: GWA) studie naar internaliserende problemen op 3 jarige leeftijd, uitgevoerd in 4596 kinderen. De meta-analyse is gebaseerd op analyses in drie cohorten: het Nederlandse Tweeling Register, "generation R" (een Rotterdamse populatiestudie) en "The Western Australian Pregnancy Cohort" (RAINE, eenAustralische populatiestudie). Zoals op grond van het aantal proefpersonen verwacht kan worden laat de meta-analyse geen genoombrede significante associaties zien. Wel zijn er suggestieve signalen in het PCSK2 gen dat zich bevindt op chromosoom 20p12.1. Overige analyses laten zien dat alle SNPs gezamenlijk tussen de 13% en 43% van de individuele verschillen in internaliserend gedrag van 3 jarigen verklaren. Verder blijken SNPs die in eerdere studies zijn geassocieerd met psychiatrische fenotypes in kinderen en volwassenen ook een collectief effect te hebben op internaliserend gedrag in 3 jarigen. Hoewel deze studie geen individuele varianten identificeert, is wel evidentie vergaard dat de genetische invloeden op internaliserende problemen bij 3 jarigen zeer poly-genetisch zijn, dat wil zeggen dat er veel genen zijn met ieder een zeer klein effect. Dit is in lijn met bevindingen uit grotere meta-analyses van psychiatrische fenotypes bij volwassenen.

In **Hoofdstuk 5** is gekeken naar de overlaptussen het genetisch risico op schizofrenie en "DSM-IV gebaseerde probleemscores" bij Nederlandse en Britse kinderen. Op basis van een genoombrede meta-analyse naar schizofrenie bij meer dan 30.000 schizofreniepatiënten en meer dan 40.000 gezonde controle-individuen werd een lineaire combinatie van SNPs gecreëerd. Deze combinatie van SNP's voorspelt schizofrenie in onafhankelijke cohorten (genetisch profiel score: polygenic risk score: PRS). Dezelfde PRS wordt gebruikt om symptomen te

voorspellen voor DSM-IV depressie, angststoornissen, externaliserende stoornissen en ADHD op leeftijd 7, 10, 13 en 15 in Britse en Nederlandse kinderen. Meta-analyse van de resultaten laat zien dat de PRS voor schizofrenie samenhangt met angst op leeftijd 10 en in mindere mate met angst op leeftijd 7 en 13 en depressie op leeftijd 7 en 10. Post-hoc toetsen laten zien dat er een sterkere samenhang is tussen PRS en internaliserende problemen in kinderen dan tussen PRS en externaliserende problemen in kinderen.

In Hoofdstuk 6 wordt een associatie tussen het catenin delta 2 gen (CTNND2) en angst die gevonden is in ratten, gerepliceerd in mensen. Data van het internationale "Psychiatric Genetics Consortium"(PGC) en de "Nederlandse studie van angst en depressie" (NESDA) zijn hiervoor geanalyseerd. Hieruit blijkt dat de varianten in het CTNND2 gen, collectief, meer met schizofrenie en depressie zijn geassocieerd dan verwacht op basis van kans. Voor angst en bipolaire stoornissen werd geen associatie aangetoond met CTNND2. Deze studie laat zien dat het waardevol kan zijnom bevindingen uit proefdierstudies te repliceren.

In **Hoofdstuk 7** wordt een nieuwe statistische techniek beschreven om gen-omgevingsinteracties te onderzoeken en te kwantificeren. Deze techniek maakt gebruik een genetische relatie matrix (GRM). Deze matrix bevat voor een grote groep personen hun genetische verwantschap met alle andere personen in de studie voor wie SNPs gemeten zijn. Voor ongerelateerde personen is de verwachte verwantschap geschaald naar 0, voor broers en zussen is de verwachte verwantschap 0.5, maar rondom deze verachtingen is een zekere mate van spreiding. Sommige ongerelateerde personen delen meer SNPs dan andere, en de vraag is of ze ook fenotypisch meer op elkaar lijken. De verwantschappen in de GRM worden bepaald op basis van gemeten SNPs. De GRM matrix wordt vervolgens gesplitst. Er is 1 matrix die alle informatie bevat en een tweede die slechts de verwantschap informatie bevat voor mensen met een paarsgewijze verwantschap boven de 0.05, dus van mensen die aan elkaar gerelateerd zijn. Door vervolgens de variantie in het fenotype te modelleren als een functie van beide matrices, kan men bepalen welk deel van de variantie in het fenotype verklaard wordt door de gemeten SNPs, en welk deel door alle genetische effecten. Ik heb vervolgens een methode ontwikkeld, die moderatie van de genetische variantie toestaat door een continue verdeelde variabele. Uit de analyses van BMI, lichaamslengte, symptomen van angst en depressie en attentieproblemen, blijkt dat de erfelijkheid van BMI en attentieproblemen

varieert met leeftijd, die van lichaamslengte met geboorte jaar. Bij BMI komt dit door veranderingen in de genetische invloeden en de omgevingsinvloeden op het fenotype, bij lichaamslengte en attentie problemen komt dit door een verandering in de omgevingsvariantie. Er is geen moderatie van de genetische effecten van angst en depressie.

Hoofdstuk 8 van dit proefschrift beschrijft een methode om de genetische covariantie tussen meerdere fenotypes vast te stellen op basis van gemeten SNPs en verwantschapsinformatie. Hier wordt zoals in hoofdstuk 7 een GRM gebruikt. Voor het schatten van de genetische covariantie tussen meerder fenotypen wordt de som van twee fenotypen gebruikt. De variantie van de som van 2 variabelen is immers de variantie van de eerste variabele plus de variantie van de tweede variabele plus twee maal de covariantie tussen de twee variabelen. Door de variantie van de som van twee variabelen in genetische en omgevingscomponenten te ontbinden en de twee afzonderlijke variabelen in een genetische en omgevingscomponent te ontbinden kan men de genetische en omgevingscovariantie tussen de twee variabelen bepalen. Met deze methode is een computationeel intensieve stap in multivariaat modeleren te voorkomen. De hier voorgestelde techniek is toegepast om de genetische covariantie te schatten tussen de 12 neuroticisme items en de 12 extraversie items van de NEO persoonlijkheidsvragenlijst. Deze analyses laten zien dat de covariantie structuur tussen deze items, gebaseerd op gemeten genetische variatie, een duidelijk onderscheid maakt tussen de neuroticisme items en de extraversie items. Hiermee is de op basis van tweelingstudies voorspelde covariantie structuur teruggevonden in moleculaire genetische data.

Hoofdstuk 9 en **Hoofdstuk 10** betreffen de Engelstalige samenvatting, en discussie van het proefschrift. In de discussie van dit proefschrift wordt een breder kader van de humane genetica geschetst.

Acknowledgments

I would like to thank all participants in the NTR studies, who contributed their time to fill out the longitudinal surveys and donated biological samples for zygosity assessment and genotyping. I am very grateful to the twins, and their extended families without whom this research would not have been possible. I am also extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. Finally, I am very grateful to the participants in the "Netherlands Study of Anxiety and Depression" for their participation, time and effort.

Funding for this project was provided by the Netherlands Scientific Organization (NWO) (912-100-20): "Genetic influences on stability and change in psychopathology from childhood to young adulthood", and the European Research Council (ERC-230374): "Genetics of Mental Illness".

The work in this thesis relies on data collected in the Netherlands Twin Register (NTR). We therefore wish to acknowledge all funding that contributed to the data collected in the NTR, including funding by the Netherlands Organization for Scientific Research (NWO) and MagW/ZonMW, the CMSB: Center for Medical Systems Biology and BBMRI –NL : Biobanking and Biomolecular Resources Research Infrastructure; the VU University Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA).

The Avera Institute *for Human Genetics*, Sioux Falls, South Dakota (USA) and the Rutgers University Cell and DNA Repository contributed significantly to the NTR zygosity assessment and genotyping through two Grand Opportunity grants: Integration of genomics and transcriptomics in normal twins and major depression (NIMH 1RC2MH089951-01) and Developmental trajectories of psychopathology (NIMH 1RC2 MH089995). Computing is supported by BiG Grid, the Dutch e-Science Grid, which is financially supported by the Netherlands Organisation for Scientific Research (NWO).

The ALSPAC study was funded by: The UK Medical Research Council and the Wellcome Trust (Grant ref: 092731) and the University of Bristol provide core support for ALSPAC.

Dankwoord

Als onderzoeker binnen de afdeling biologische psychologie ben ik veel dank verschuldigd aan de tweelingen en hun familie die jaar in jaar uit deelnemen aan het onderzoek van het Nederlandse tweelingen register. Dit proefschrift had niet tot stand kunnen komen zonder hun herhaalde deelname aan onderzoek de afgelopen 27 jaar.

Ik zou graag op deze plaats dank betuigen aan een aantal mensen die mij gedurende mijn promotie gesteund hebben. Allereerst wil ik mijn promotor Dorret Boomsma en co-promotoren Christel Middeldorp en Conor Dolan bedanken. Bedankt, voor de wijze lessen, de vele kansen en de prettige samenwerking. Daarnaast bedank ik ze graag voor hun geduld met mijn diselectie.

Naast mijn directe begeleiders bedank ik graag Gonneke Willemsen en Toos van Beijsterveldt voor de uitstekende begeleiding en hulp bij het samenstellen van fenotypische datasets, maar ook wat betreft het meedenken over analyses en analyse technieken. Ook bedank ik graag Jouke-Jan Hottenga, zonder wie ik niet ver was gekomen op het gebied van statistische genetica. Ik bedank graag Gitta Lubke voor de prettige samenwerking, de altijd relevante feedback en voor het statistische onderwijs op Notre Dame university dat ik kosteloos via Skype mocht volgen. Verder wil ik Meike Bartels en Jacqueline Vink bedanken voor de excellente begeleiding en Marleen de Moor die altijd bereid was om mee te denken aan statistische modellen.

I would further like to thank the members of the reading committee; Dale Nyholt, Ramon Lindauer, Hein Putter, Jeffrey Glennon, Denny Borsboom and Marleen de Moor who took the time to read and evaluate my dissertation.

I would like to thank Marcus Munafo, Beate St. Pourcaine and David Evans for the collaboration and the excellent supervision during my stay at the University of Bristol. Ghazaleh, John, Dylan, Alex and Bridget thanks for making me feel right at home in Bristol. I would further like to thank John Hewitt and the faculty of the international twin workshop for letting me tag along as junior faculty during the 2012 and 2014 workshops.

Verder wil ik mijn academische kamer genoten Abdel en Dirk bedanken, de informele discussies in de kamer hebben tot substantiële verbetering van menig

paper geleid. Dan wil ik verder Maria bedanken die vooral in het begin van mijn promotie traject geduldig iedere vraag beantwoorden, ook als ik tot 5 maal binnen en uur met een vraag binnenkwam, of zonder vraag. Verder wil ik Sanja bedanken, je stond altijd klaar om me met wat dan ook te helpen, of me terecht te wijzen, of meestal beide. Ik wil Jenny bedanken voor de verrijkende samenwerking met betrekking tot de epi-genetische data, een samenwerking die hopelijk in de toekomst kan worden voortgezet. Ik wil Harmen bedanken voor de samenwerking in het onderwijs en onderzoek. Ik wil Wouter, Maarten, Kees-Jan, Klaas-Jan, Bart, Janneke, Hamdi, Karin, Irina en Camelia bedanken voor het meedenken over statistische modellen, artikelen en/of software. Verder bedank ik mijn teamgenoot Laura.

Ik wil verder de (oud) collega's en gasten van biologische-psychologie en NESDA; Anouk, Dennis, Elsje, Lannie, Niels, Rene, Melanie, Charlotte, Nienke, Evelien, Jorien, Suzanne, Fiona, Quinta, Mike, Jenny, Rene, Matthijs, Lot, Diane, Ineke, Bochao, Nuno, Charles, Raymond, Jouri, Rik, Dan, Michiel, Michelle, Corina, Stephanie, Brenda, Natascha en Eco de Geus bedanken voor een fantastisch (academisch) werk klimaat.

Dan rest mij nog Renske, en mijn ouders, Rietje en Guillaume Nivard te bedanken voor de steun, het geduld en een luisterend oor.