

Online Resource 1

Supplementary Materials belonging to:

The genetic architecture of liver enzyme levels: GGT, ALT and AST

Jenny H.D.A. van Beek ^{1,2,5}, Marleen H.M. de Moor ^{1,2}, Eco. J.C. de Geus ^{1,2,3}, Gitta H. Lubke ^{1,4}, Jacqueline M. Vink ^{1,3}, Gonneke Willemsen ^{1,2}, & Dorret I. Boomsma ^{1,2,3}

¹ Department of Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands

² EMGO+ Institute, VU University Medical Center, Amsterdam, the Netherlands

³ Neuroscience Campus Amsterdam, Amsterdam, the Netherlands

⁴ Department of Psychology, University of Notre Dame, Notre Dame, Indiana, USA

⁵ To whom correspondence should be addressed: J.H.D.A. van Beek, VU University Amsterdam, Department of Biological Psychology, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands. Telephone: +31 20 598 8787. Fax: +31 20 598 8832. E-mail: J.H.D.A.van.Beek@vu.nl.

Submitted for publication in Behavior Genetics

Supplementary Materials Online methods

Statistical analysis of the genetic architecture of liver enzyme levels

The genetic analysis of liver enzyme levels was performed with saturated models and structural equation models. In the model we have called the full saturated model, 16 familial correlations were estimated, six means, six beta coefficients for the age regression, and six standard deviations (SDs) (for twins, siblings, and parents, separate for both sexes). This model is referred to as model 1a. In a series of submodels, sex and age differences in means and variances were tested, as well as age and sex effects on the correlation structure. This was done by comparing the model fit results for the constrained models with the full model, based on the likelihood ratio test (Bentler and Bonett 1980) with a significance level (α) of .01. If model fit deteriorated significantly for the constrained model compared to the full model, the specific parameter constraints were not allowed (e.g. sex-invariant familial correlations), thus informing which effects were present (e.g. sex effects on the correlations).

Fixed effects of age were examined by (a) testing whether mean enzyme levels differed for twins, siblings and parents (within sex), (b) constraining betas for the age regression to zero to see if an effect of age was present, and if so (c) by constraining the betas to be equal within sex (equating betas for twins, siblings and parents) to test whether the age effect was similar for twins, siblings and parents who differ in mean age, and (d) by constraining the betas over sex to test whether the age effect was equal over sex. Fixed effects of sex on enzyme levels were tested by equating mean enzyme levels over sex. To test for fixed effects of age and sex on the variance, the standard deviations (SDs) were equated for twins, siblings, and parents within sex and over sex respectively. Model 1b included the constraints on the means, SDs and age

regression that were allowed based on the tests described above and provided information on the 16 familial correlations given in figures 2a-c.

The effect of age on the familial correlations was tested in models 2a-c, and that of sex on the familial correlations in models 3a-e. In model 2a it was tested whether twin-sibling correlations were similar to sibling-sibling correlations. Using this model as a reference, in model 2b it was tested whether dizygotic (DZ) twin correlations could be constrained to sibling correlations. When this was allowed, no evidence was present for a special twin environment. Model 2c tested whether correlations among first-degree relatives could be equated. Since parent-offspring pairs show larger within-pair age differences than offspring and DZ twin pairs, the absence of significant differences in these correlations would indicate that there were no age differences in the heritability. Model 3a and 3b tested for quantitative sex differences in the correlations. In model 3a it was tested whether the correlation among monozygotic twin pairs could be constrained to be equal. In model 3b this was tested for all first-degree relatives. Model 3c-e tested for qualitative sex differences in the correlations. In model 3c, male and female DZ twin and sibling correlations were constrained to be equal. Taking this model as a reference, with model 3d it was examined whether the correlations among opposite-sex DZ twin and sibling pairs could be equated to same-sex DZ twin and sibling pairs. In model 5e it was tested whether opposite-sex parent-offspring correlations could be constrained to be equal to same-sex parent-offspring correlations. The presence of spousal resemblance was tested in model 4. When significant spousal resemblance was detected (model 4), we explored whether this could be explained by cohabitation effects. This was tested by estimating the correlation between the absolute difference in enzyme levels for spouses and the duration of their relationship in Mx v3.2 (Neale et al. 2006) in two zygosity groups (MZ, DZ). If living together influences resemblance

between spouses, this correlation will be negative, i.e. smaller differences between spouses the longer they are together. This test was based on data of 727 parental pairs and 257 twin-spouse pairs with data on relationship duration.

Structural equation modeling was performed to estimate the sources of variance that could explain differences in liver enzyme levels. Factor models were fit to the data by which phenotypic values were regressed on latent factors specifying the genetic and environmental contributions (see figure 1). In the ACE model with shared environmental transmission, the variation in the phenotypic values for fathers, mothers, sons and daughters (depicted as PFA, PMO, PSO, PDA respectively) was regressed on three latent variables: an additive genetic factor (with factor loadings $a_{\text{♂}}$ and $a_{\text{♀}}$ for males and females respectively), and a shared environmental (factor loadings $c_{\text{♂}}$ and $c_{\text{♀}}$) and non-shared environmental factor (factor loadings $e_{\text{♂}}$ and $e_{\text{♀}}$). If there were sex differences in the correlation structure, the factor loadings were sex-specific (and thus the variance decomposition by $\Lambda\Psi\Lambda'$), otherwise, these were constrained to be equal over sex. MZ twin pairs were modeled to share their entire genetic material (that is, the additive genetic factors (A) correlate 1 within twin pairs), whereas DZ twin pairs/siblings, and parent and offspring are estimated to share half of their genes (A factors were modeled to correlate .5 among DZ twin, sibling and parent-offspring pairs). The latent additive genetic variance in the offspring generation of .5 reflects the segregation variance that emerges due to recombination. This within-family additive genetic variance emerges since parents pass alleles, not genotypes, to their offspring, resulting in new genetic variance for every generation. C factors reflecting shared environmental influences were correlated 1 among the offspring generation within each family. This model allowed for shared environmental transmission from parents to offspring, indicated by $t_{\text{FA,SO}}$, $t_{\text{MO,SO}}$, $t_{\text{FA,DA}}$, $t_{\text{MO,DA}}$. The presence of both genetic and shared environmental

transmission, gives rise to a correlation between the additive genetic and shared environmental factors (A-C correlation). This is represented by $R_{A,C}$. Spousal resemblance was modeled to run via the Δ -path that represents the correlations between the latent genetic and environmental factors influencing the phenotypes of the parents, that result from phenotypic assortment. The presence of phenotypic assortment will increase the additive genetic variance, represented by R_A .

If the familial correlations suggested qualitative sex differences, it could be tested whether these resulted from (a) a different expression of genes over sex, (b) differences in the environmental factors shared within the offspring generation or (c) a sex-specific transmission of shared environmental effects from parents to offspring. Whether liver enzyme levels among males and females were affected by different genes could be tested by estimating the genetic correlation between opposite-sex parent-offspring pairs and testing whether this correlation was significantly lower than .5. Qualitative sex differences in the environmental factors shared by the offspring generation, could be examined by estimating the correlation between the shared environmental factors ($r_{C,OS}$) in opposite-sex offspring pairs and testing whether this correlation differed significantly from 1. Qualitative sex differences in the shared environmental transmission paths could be explored by testing whether shared environmental transmission differed significantly for opposite-sex parent-offspring pairs and same-sex parent offspring pairs.

In case the familial correlations did not indicate evidence for qualitative sex differences, the tests to explore the qualitative sex differences in the genetic model (described above) were not performed. Thus if the familial correlations did not suggest qualitative sex differences, the genetic correlation for opposite-sex parent-offspring pairs was fixed to .5; the correlation between the shared environmental factors for opposite-sex twin/sibling pairs to 1; and the path

loadings for the shared environmental transmission for opposite-sex parent-offspring pairs equated to those for same-sex parent-offspring pairs (resulting in the estimation of one sex-invariant shared environmental transmission path).

The ACDE model differed from the ACE model with shared environmental transmission in that the shared environmental transmission paths from parents to offspring were set to zero (for reasons of model identification) and additionally, non-additive genetic influences (D) were estimated (with factor loadings $d_{\text{♂}}$ and $d_{\text{♀}}$). D factors correlated 1 among MZ twin pairs and .25 among DZ twins and sibling pairs. By dropping shared environmental transmission from the model, A-C correlation was no longer implicated and set to zero.

If the familial correlations suggested age differences, the genetic models allowed to examine whether these represented quantitative or qualitative age differences in the heritability. Qualitative age differences denote a different expression of genes over age and would be indicated if the genetic correlation between same-sex parent-offspring pairs was significantly lower than .5. Note that with data of twins, siblings and parents, simultaneous estimation of the genetic parent-offspring correlation, and additive genetic, non-additive genetic, shared environmental and non-shared environmental factors is not possible and one of these has to be set to zero to keep the model identified (a model estimating the genetic correlation between parents and offspring *and* cultural transmission effects from parents to offspring is also not identified). Quantitative age differences do not reflect a different expression of genes over age, but indicate that the heritability differs over age, for instance because of the presence of shared environmental effects or non-additive genetic effects that are shared within the offspring generation but not between parents and offspring.

Supplementary Materials Table S1

Supplementary Materials Table S1

Model fit statistics for tests of age and sex effects on familial correlations for GGT, ALT and AST

	Model	NP	-2LL	df	vs	$\Delta\chi^2$	Δdf	p	Effect on correlations
GGT	1a full saturated model	34	47724.962	8076					
	1b saturated model with constraints on means and SDs ^a	25	47742.568	8085	1a	17.606	9	.040	
	2a twin-sibling=sibling-sibling correlation (within sex)	22	47749.840	8088	1b	7.272	3	.064	equal twin-sibling and sibling-sibling correlations
	2b DZ twin=sibling correlation (within sex)	19	47752.697	8091	2a	2.857	3	.414	no special twin environment
	2c DZ twin=sibling=parent-offspring correlation (within sex)	15	47782.640	8095	1b	40.072	10	<.001	age differences in correlations first-degree relatives
	3a male=female MZ correlation	24	47742.661	8086	1b	.093	1	.761	no quantitative sex diff in MZ corr
	3b male=female=opposite-sex DZ twin/sibling/parent-offspring correlation	13	47789.155	8097	1b	46.587	12	<.001	quantitative sex differences in correlations first-degree relatives
	3c male=female DZ twin/sibling correlation	20	47768.211	8090	1b	25.643	5	<.001	quantitative sex differences in offspring generation
	3d male=female=opposite-sex DZ twin/sibling correlation	17	47779.352	8093	3c	11.141	3	.011	no qualitative sex differences in correlations offspring generation
	3e same-sex=opposite-sex par-off correlation	23	47742.847	8087	1b	.279	2	.870	no qualitative sex differences in parent-offspring correlations
4 spouse correlation=0	24	47748.226	8086	1b	5.658	1	.017	no spousal resemblance	
ALT	1a full saturated model	34	45446.979	7761					
	1b saturated model with constraints on means and SDs ^b	23	45466.019	7772	1a	19.040	11	.060	
	2a twin-sibling=sibling-sibling correlation (within sex)	20	45475.796	7775	1b	9.777	3	.021	equal twin-sibling and sibling-sibling correlations
	2b DZ twin=sibling correlation (within sex)	17	45477.593	7778	2a	1.797	3	.616	no special twin environment

	2c	DZ twin=sibling=parent-offspring corr (within sex)	13	45479.039	7782	1b	13.020	10	.223	no age differences in correlations first-degree relatives
	3a	male=female MZ correlation	22	45473.616	7773	1b	7.597	1	.006	quantitative sex diff in MZ corr
	3b	male=female opp-sex DZ twin/sibling/parent-offspring correlation	11	45511.921	7786	1b	18.037	12	.115	no quantitative sex differences in correlations first-degree relatives
	3c	male=female DZ twin/sibling correlation	18	45480.320	7777	1b	14.301	5	.014	no quantitative sex differences in correlations offspring generation
	3d	male=female=opp-sex DZ twin/sibling correlation	15	45482.886	7780	3c	2.566	3	.463	no qualitative sex differences in correlations offspring generation
	3e	same-sex=opposite-sex par-off correlation	21	45466.937	7774	1b	.918	2	.632	no qualitative sex differences in parent-offspring correlations
	4	spouse correlation=0	22	45486.658	7773	1b	20.640	1	<.001	significant spousal resemblance
AST	1a	full saturated model	34	38520.527	8034					
	1b	saturated model with constraints on means and SDs ^c	22	38549.677	8046	1a	29.150	12	.085	
	2a	twin-sibling=sibling-sibling correlation (within sex)	19	38554.675	8049	1b	4.998	3	.172	equal twin-sibling and sibling-sibling correlations
	2b	DZ twin=sibling correlation (within sex)	16	38563.456	8052	2a	8.781	3	.032	no special twin environment
	2c	DZ twin=sibling=parent-offspring correlation (within sex)	12	38570.088	8056	1b	20.412	10	.026	no age differences in correlations first-degree relatives
	3a	male=female MZ correlation	21	38550.371	8047	1b	.694	1	.405	no quantitative sex diff in MZ corr
	3b	male=female opposite-sex DZ twin/sibling/parent-offspring correlation	10	38571.054	8058	1b	21.378	12	.045	no quantitative sex differences in correlations first-degree relatives
	3c	male=female DZ twin/sibling correlation	17	38557.833	8051	1b	8.156	5	.148	no quantitative sex differences in correlations offspring generation
	3d	male=female=opposite-sex DZ twin/sibling correlation	14	38566.211	8054	3c	8.378	3	.039	no qualitative sex differences in correlations offspring generation
	3e	same-sex=opposite-sex par-off correlation	20	38550.182	8048	1b	.506	2	.777	no qualitative sex differences in parent-offspring generation
	4	spouse correlation=0	21	38578.384	8047	1b	28.708	1	<.001	significant spousal resemblance

NP=number of parameters in statistical model

-2LL=-2 log likelihood, fit function for model with *df* degrees of freedom

$\Delta\chi^2$ =model fit statistic; difference in -2LL of two nested models is approximately distributed as χ^2 with $df=\Delta df$ (the difference in NP between the models); p -value was regarded significant when $<.01$

vs=versus, model that was used as reference model to assess the fit of the constrained model

sib=sibling; par-off=parent-offspring; opp-sex=opposite-sex; corr=correlation; diff=differences

^a For GGT constraints included three SDs (male offspring, female offspring, parents), four means (male offspring, female offspring, fathers, mothers), and two betas for the age regression (fathers, others)

^b For ALT constraints included one sex-invariant SD, two sex-specific means, and four betas for age regression (male offspring, female offspring, fathers, mothers)

^c For AST constraints included sex-specific SDs, means, and betas for the age regression

Supplementary Materials Tables S2a-c

Supplementary Materials Table S2a

Parameter constraints (and number of degrees of freedom) associated with tests on saturated models

Model comparison	Test	Phenotype	Δdf	Constraints
1b vs 1a	constraints on means and SDs	GGT	9	mean father=mean mother ($\Delta df=1$) mean male twin=male sibling ($\Delta df=1$) mean female twin=female sibling ($\Delta df=1$) SD male twin=male sibling ($\Delta df=1$) SD female twin=female sibling ($\Delta df=1$) beta age male twin=male sib=female twin=female sib=mother ($\Delta df=4$) <i>Number of parameters saturated model with constraints (model 1b): 25</i> - 4 SDs (male twins/siblings, female twin/siblings, fathers, mothers) - 3 means (male twins/siblings, female twin/siblings, parents) - 2 betas age regression (fathers, others) - 16 familial correlations (see text)
		ALT	11	mean male twin=male sibling=father ($\Delta df=2$) mean female twin=female sibling=mother ($\Delta df=2$) SD male=female twin/sibling/parent ($\Delta df=5$) beta age male twin=male sibling ($\Delta df=1$) beta age female twin=female sibling ($\Delta df=1$) <i>Number of parameters saturated model with constraints (model 1b): 23</i> - 1 SDs (sex-invariant) - 2 means (males, females) - 4 betas age regression (male twin/sib, female twin/sib, fathers, mothers) - 16 familial correlations (see text)
		AST	12	mean male twin=male sibling=father ($\Delta df=2$) mean female twin=female sibling=mother ($\Delta df=2$)

SD male twin=male sibling=father ($\Delta df=2$)
SD female twin=female sibling=mother ($\Delta df=2$)
beta age male twin=male sibling=father ($\Delta df=2$)
beta age female twin=female sibling=mother ($\Delta df=2$)
Number of parameters saturated model with constraints (model 1b): 22
- 2 SDs (males, females)
- 2 means (males, females)
- 2 betas age regression (males, females)
- 16 familial correlations (see text)

vs=versus; sib=sibling

Δdf = the difference in degrees of freedom between the tested model and the full model (is equal to the difference in the number of parameters of the two models)

Supplementary Materials Table S2b

Parameter constraints (and number of degrees of freedom) associated with tests on correlations

Model comparison	Test	Pheno-type	Δdf	Constraints
2a vs 1b	twin-sibling=sibling-sibling correlation (within sex)	ALL	3	male twin-brother=brother-brother correlation ($\Delta df=1$) female twin-sister=sister-sister correlation ($\Delta df=1$) opposite-sex twin-sibling=brother-sister correlation ($\Delta df=1$)
2b vs 2a	DZ twin=sibling correlation (within sex)	ALL	3	male DZ twin=male twin/brother-brother correlation ($\Delta df=1$) female DZ twin=female twin/sister-sister correlation ($\Delta df=1$) opposite-sex DZ twin=opposite-sex twin/sibling-sibling corr ($\Delta df=1$)
2c vs 1b	DZ twin=sibling=parent-offspring correlation (within sex)	ALL	10	male DZ twin=male twin-brother=brother-brother=father-son correlation ($\Delta df=3$) female DZ twin=female twin-sister=sister-sister=mother-daughter corr ($\Delta df=3$) opposite-sex DZ twin=opposite-sex twin-sibling=brother-sister=mother-son=father-daughter correlation ($\Delta df=4$)
3a vs 1b	male=female MZ correlation	ALL	1	male MZ twin=female MZ twin correlation
3b vs 1b	male=female opposite-sex DZ twin/sibling/ parent-offspring correlation	ALL	12	male DZ twin=male twin-brother=brother-brother=father-son= female DZ twin=female twin-sister=sister-sister=mother-daughter= opposite-sex DZ twin=opposite-sex twin-sibling=brother-sister= mother-son=father-daughter correlation ($\Delta df=12$)
3d vs 3c	male=female=opposite-sex DZ twin/sibling correlation	ALL	3	same-sex DZ twin/sibling=opposite-sex DZ twin correlation ($\Delta df=1$) same-sex DZ twin/sibling=opposite-sex twin-sibling correlation ($\Delta df=1$) same-sex DZ twin/sibling=opposite-sex sibling-sibling correlation ($\Delta df=1$)
3e vs 1b	same-sex=opposite-sex par-off correlation	ALL	2	father-son=mother-son correlation ($\Delta df=1$) mother-daughter=mother-son correlation ($\Delta df=1$)
4 vs 1b	spouse correlation=0	ALL	1	father-mother correlation=0 ($\Delta df=1$)

vs=versus; sib=sibling

Δdf = the difference in degrees of freedom between the tested model and the full model (is equal to the difference in the number of parameters (NP) of the two models)

Supplementary Materials Table S2c

Parameter constraints (and number of degrees of freedom) associated with tests in ACDE model/ ACE model with shared environmental transmission

Model comparison	Test	Phenotype	Δdf	Constraints
ACE with shared environmental transmission vs 1b	Does the ACE model with shared environmental transmission fit data	GGT	8	<i>Number of parameters saturated model with constraints (model 1b): 25</i> <i>Number of parameters ACE model with shared environmental transmission: 17</i> - 4 SDs (male twins/siblings, female twin/siblings, fathers, mothers) - 3 means (male twins/siblings, female twin/siblings, parents) - 2 betas age regression (fathers, others) - 2 additive genetic (a) path-loadings (males, females) - 2 shared environmental (c) path-loadings (males, females) - 1 shared environmental transmission (t) path (sex-invariant) - 2 non-shared individual-specific environmental (e) path loadings - 1 covariance between additive genetic and shared environmental effects ($r_{A,C}$)
		ALT	6	----- <i>Number of parameters saturated model with constraints (model 1b): 23</i> <i>Number of parameters ACE model with shared environmental transmission: 17</i> - 1 SD (sex-invariant) - 2 means (males, females) - 4 betas age regression (male twin/sib, female twin/sib, fathers, mothers) - 2 additive genetic (a) path-loadings (males, females) - 2 shared environmental (c) path-loadings (males, females) - 1 shared environmental transmission (t) path (sex-invariant) - 2 non-shared individual-specific environmental (e) path loadings - 1 covariance between additive genetic and shared environmental effects ($r_{A,C}$) - 1 covariance among parents (delta path, Δ , to model phenotypic assortment) - 1 additive genetic variance (R_A) (due to phenotypic assortment)
		AST	9	----- <i>Number of parameters saturated model with constraints (model 1b): 22</i>

				<hr/> <i>Number of parameters ACE model with shared environmental transmission: 13</i> - 2 SDs (males, females) - 2 means (males, females) - 2 betas age regression (males, females) - 1 additive genetic (a) path-loading (sex-invariant) - 1 shared environmental (c) path-loading (sex-invariant) - 1 shared environmental transmission (t) path (sex-invariant) - 1 non-shared individual-specific environmental (e) path loading (sex-invariant) - 1 covariance between additive genetic and shared environmental effects ($r_{A,C}$) - 1 covariance among parents (delta path, Δ , to model phenotypic assortment) - 1 additive genetic variance (R_A) (due to presence phenotypic assortment) <hr/>
ACDE vs 1b	Does the ACDE model fit data	GGT	8	<hr/> <i>Number of parameters saturated model with constraints (model 1b): 25</i> <i>Number of parameters ACDE model: 17</i> - 3 SDs (male offspring, female offspring, parents) - 4 means (male offspring, female offspring, fathers, mothers) - 4 betas age regression (male twin/sibling, female twin/sibling, fathers, mothers) - 2 additive genetic (a) path-loadings (males, females) - 2 non-additive genetic (d) path-loadings (males, females) - 2 shared environmental (c) path-loadings (males, females) - 2 non-shared individual-specific environmental (e) path loadings <hr/>
		ALT	6	<hr/> <i>Number of parameters saturated model with constraints (model 1b): 23</i> <i>Number of parameters ACE model with CT: 17</i> - 1 SD (sex-invariant) - 2 means (males, females) - 4 betas age regression (male twin/sibling, female twin/sibling, fathers, mothers) - 2 additive genetic (a) path-loadings (males, females) - 2 non-additive genetic (d) path-loadings (males, females) - 2 shared environmental (c) path-loadings (males, females) - 2 non-shared individual-specific environmental (e) path loadings - 1 covariance among parents (delta path, Δ , to model phenotypic assortment) - 1 additive genetic variance (R_A) (due to phenotypic assortment) <hr/>

		AST	10	<hr/> <i>Number of parameters saturated model with constraints (model 1b): 22</i> <i>Number of parameters ACE model with CT: 12</i> - 2 SDs (males, females) - 2 means (males, females) - 2 betas age regression (males, females) - 1 additive genetic (a) path-loading (sex-invariant) - 1 non-additive genetic (d) path-loading (sex-invariant) - 1 shared environmental (c) path-loading (sex-invariant) - 1 non-shared individual-specific environmental (e) path loading (sex-invariant) - 1 covariance among parents (delta path, Δ , to model phenotypic assortment) - 1 additive genetic variance (R_A) (due to presence phenotypic assortment)
ADE vs ACDE	drop C to zero	ALL both sexes	2	c-path males=c-path females=0 ($\Delta df=2$)
ACE vs ACDE	drop D to zero	ALL both sexes	2	d-path males=d-path females=0 ($\Delta df=2$)
AE vs ACDE	drop A+D to zero	ALL both sexes	5	a-path males=a-path females=0 ($\Delta df=2$) d-path males=d-path females=0 ($\Delta df=2$) additive genetic variance (R_A)=1 ($\Delta df=1$)

vs=versus; sib=sibling

Δdf = the difference in degrees of freedom between the tested model and the full model (is equal to the difference in the number of parameters (NP) of the two models)

References

Bentler PM, Bonett DG (1980) Significance tests and goodness of fit in the analysis of covariance structures. *Psychol Bull* 88:588-606

Neale, MC, Boker, SM, Xie, G, & Maes, HH (2006) *Mx: Statistical modeling* (7 ed).
Department of Psychiatry, VCU, Richmond