Genetic determinants of eating disorders

Slof-Op 't Landt, M.C.T.

Chapter 5

Association study in eating disorders: TPH2 associates with anorexia nervosa and self-induced vomiting

This chapter was previously published:

Slof-Op 't Landt, M. C. T., Meulenbelt, I., Bartels, M., Suchiman, H. E. D., Middeldorp, C. M., Houwing-Duistermaat, J. J., van Trier, J., Onkenhout, E. J., Vink, J. M., Van Beijsterveldt, C. E. M., Brandys, M. K., Sanders, N., Zipfel, S., Herzog, W., Herpertz-Dahlmann, B., Warnke, A., Fleischhaker, C., Zeeck, A., de Zwaan, M., Herpertz, S., Ehrlich, S., van Elburg, A. A., Adan, R. A. H., Scherag, S., Hinney, A., Hebebrand, J., Boomsma, D. I., van Furth, E. F., & Slagboom, P. E. (2011).

Genes, Brian and Behavior, 10, 236-243.

Abstract

Twin studies suggest that genetic factors play a substantial role in anorexia nervosa (AN) and self-induced vomiting (SV), a key symptom that is shared among different types of eating disorders. We investigated the association of 25 single nucleotide polymorphisms (SNPs), capturing 71 to 91% of the common variance in candidate genes stathmin (*STMN1*) , serotonin receptor 1D (*HTR1D*), tryptophan hydroxylase 2 (*TPH2*) and brain-derived neurotrophic factor (*BDNF*), with AN and eating disorders characterized by SV.
First, allele frequencies of all SNPs were compared between a Dutch case group (182 AN, 149 eating disorders characterized by SV) and 607 controls. Associations rendering *p*-values <0.05 from this initial study were then tested for replication in a meta-analysis with two additional independent eating disorder case-control samples, together providing 887 AN participants, 306 participants with an eating disorder characterized SV, and 1914 controls. A significant effect for the minor C-allele of *TPH2* rs1473473 was observed for both AN (OR=1.30, 95% CI 1.08-1.57, $p<0.003$) and eating disorders characterized by SV (OR=1.52, 95% CI 1.28-2.04, $p<0.006$). In the combined case group a dominant effect was observed for rs1473473 (OR=1.38, 95% CI 1.16-1.64, $p<0.0003$). The meta-analysis revealed that the *TPH2* polymorphism rs1473473 was associated with a higher risk for AN, eating disorders characterized by SV and for the combined group.

Eating disorders (ED) are debilitating diseases with high chronicity and mortality rates (Crow et al., 2009; Steinhausen, 2002; Steinhausen & Weber, 2009). Genetic influences appear to be considerable for ED, with heritability estimates ranging from 28 to 83% in women (Bulik et al., 2006; Slof-Op 't Landt et al., 2005). A common and frequently occurring symptom in individuals with ED is self-induced vomiting (SV). This symptom was associated with greater clinical severity (Dalle Grave et al., 2009; Reba et al., 2005), and also appears to be heritable (8-72%) (Sullivan et al., 1998a; Wade et al., 2008b).

Despite the multitude of performed molecular genetic studies in ED, no specific genes have been definitively implicated as causal, although several promising candidate genes exist (Scherag et al., 2010; Slof-Op 't Landt et al., 2005). To retain adequate statistical power we selected four of these candidate genes to test for association in a case-control design. The selected genes were serotonin receptor 1D (*HTR1D*), tryptophan hydroxylase 2 (*TPH2*), stathmin (*STMN1*) and brain-derived neurotrophic factor (*BDNF*).

HTR1D and *TPH2* belong to the serotonin pathway. Serotonin is involved in a broad range of functions, including body weight regulation, eating behavior and mood (Lucki, 1998). Furthermore, the functional activity of the serotonin system appears to be altered in both current as well as recovered ED patients (Ehrlich et al., 2010; Kaye, 2008; Kaye et al., 2005a). *HTR1D* is located under the linkage peak for anorexia nervosa (AN) at 1p33-36 (Grice et al., 2002), and was significantly associated to AN in two independent studies (Bergen et al., 2003; Brown et al., 2006). *TPH2* encodes the rate-determining enzyme in the synthesis of serotonin tryptophan hydroxylase in the brain (Walther & Bader, 2003), and was previously associated with depression and anxiety (Barnett & Smoller, 2009; Kim et al., 2009; Tsai et al., 2009; Zhang et al., 2006).

STMN1 is also located under the linkage peak for restrictive AN (Grice et al., 2002), and was associated to fear processing and anxiety in both mice and humans (Brocke et al., 2010; Shumyatsky et al., 2005).

Finally, the involvement of *BDNF* in ED was reported by two large collaborative studies, that showed an association between AN and the functional Val-66-Met polymorphism (Ribases et al., 2004; Ribases et al., 2005). This finding was replicated by some but not all subsequent studies (Scherag et al., 2010).

In general, consistent associations in the ED field are lacking, possibly due to small sample sizes and the limited number of polymorphisms assessed (Scherag et al., 2010; Slof-Op 't Landt et al., 2005). In the current study we selected 25 tagging SNPs across the four genes and tested them for association with AN (N=182) and participants (N=149) with an ED characterized by self-induced vomiting (SV). Replication occurred in a meta-analysis

with two additional independent ED case-control samples from Germany and the Netherlands together providing 887 AN cases, 306 SV cases and 1914 controls.

Methods and Materials

Participants

This study was approved by each national ethics committee. All participants (and if underage, their parents) gave written informed consent.

Three hundred and eighty nine female ED patients were recruited through ten specialist ED units throughout the Netherlands (the GenED study). All participants fulfilled DSM-IV (Diagnostic and Statistical Manual of Mental Disorders 4th edition) criteria for an ED, made by experienced clinicians based on a semi-structured interview at intake, and via the self-report eating disorder examination questionnaire (EDEQ; Fairburn & Beglin, 1994). For AN, criterion D - amenorrhea for three consecutive months- was discarded because some of the participants despite having AN continue to menstruate (for example, due to treatment with oral contraceptives). Of the 389 cases, 182 fulfilled the DSM-IV criteria (excluding criterion D) for AN. Based on the EDEQ (q14: Over the past 28 days, how many times have you made yourself sick (vomit) as a means of controlling your shape and weight?) and assessment interviews (current and past slef-induced vomiting), we defined a subgroup of ED cases (N=149) who reported regular self-induced vomiting (SV). Frequencies of mean rates of self-induced vomiting were 30% 2 to 8 times per month, 40% 8 to 20 times per month, and 30% more than 20 times per month. Participants with SV fulfilled the following DSM-IV diagnoses AN (N=64), boulimia nervosa (N=74) and ED not otherwise specified (N=11) (see Table 5.1). Thus the two groups were partly overlapping, with 64 participants belonging to both groups.

Random controls come from the population-based Netherlands Twin Registry (NTR), which was established in the late 1980s at the VU University in Amsterdam, the Netherlands. Data on the multiples (twins or triplets) and their families have been collected every two to three years in longitudinal survey studies (Boomsma et al., 2002). Subsamples of the multiples were invited to participate in experimental and laboratory studies and donate their DNA (Boomsma et al., 2006). For the current study, one woman per family served as control, yielding a control group of 607 unrelated women (Middeldorp et al., 2010).

For the meta-analysis additional sample collections were used from Essen (The EDNET and Essen study, Germany) and Utrecht (The Netherlands) (see Table 5.1). The EDNET and Essen sample consisted of 420 female participants with AN according to

DSM-IV criteria and 189 normal weight controls (75 men and 114 women; females with ED were excluded) (Muller et al., 2008). The Utrecht sample consisted of 481 female participants diagnosed with an eating disorder, 285 participants fulfilled DSM-IV criteria for AN and 157 participants reported regular self-induced vomiting. These two groups were partly overlapping, with 56 participants belonging to both groups. As a control population measured and imputed genotype data from the female control group of the GAIN GWA study were used (Boomsma et al., 2008). This group comprised 1118 unrelated female participants from the NTR who were at low liability for major depressive disorder. The GAIN control group was independent of the initial NTR control group.

Table 5.1 Cases and controls

				DSM-IV Eating disorder diagnosis				
Cases and controls	Total N	Age (SD)	Overlap AN & SV	AN-R	AN-BP	AN-P	BN	EDNOS
GenED								
NTR Controls	607	25.4 (13.6)						
GenED AN	182	28.7 (9.9)	64	108	35	39	-	-
GenED SV	149	28.9 (9.9)	64	-	29	35	74	11
EDNET-Essen								
EDNET-Essen Controls	189	24.6 (2.5)						
EDNET-Essen AN	420	21.4 (9.1)	-	152	N.A.	N.A.	-	-
Utrecht								
Gain GWA NTR Controls	1118	44.0 (13.7)						
Utrecht AN	285	22.9 (4.8)	56	213	N.A.	N.A.	-	-
Utrecht SV	157	23.8 (5.7)	56	-	N.A.	N.A.	37	63

AN= anorexia nervosa, AN-R= anorexia nervosa restriciting type, AN-BP= anorexia nervosa binge purging type, AN-P= anorexia nervosa purging type (without binge eating), SV= eating disorders characterized by self-induced vomiting, BN=bulimia nervosa, EDNOS=eating disorders not otherwise specified, N.A.=data not available

SNP selection and genotype measurements

Genomic DNA was isolated from buccal swabs for the case group from the GenED study and for part of the NTR control group (39%). For the EDNET-Essen, and the Utrecht samples, genomic DNA was isolated from blood samples.

Table 5.2 Selected SNPs per candidate gene

Gene	SNP	Remarks
STMN1	rs12037513	The 3 SNPs genotyped capture
	rs807055	11 out of 12 (91%) alleles of
	rs807062	STMN1 at r^2 >=0.8
HTR1D	rs605367	The 2 tagging SNPs (rs676643
	rs6300	and rs674386) genotyped
	rs676643	capture 9 out of 10 (90%)
	rs674386	alleles of HTR1D at r^2 >=0.8
TPH2	rs10748185	The 10 SNPs genotyped
	rs2129575	capture 108 out of 148 (72%)
	rs17110489	alleles of TPH2 at r^2 >=0.8
	rs7305115	
	rs1007023	
	rs4760820	
	rs1473473	
	rs3903502	
	rs12231356	
	rs4474484	
BDNF	rs7124442	The 8 SNPs genotyped capture
	rs6265	38 out of 53 (71%) alleles of
	rs11030107	BDNF at r^2 >=0.8
	rs7103873	
	rs11030123	
	rs17309930	
	rs2049048	
	rs1491851	

HTR1D SNPs were selected based on previous association studies in AN (Bergen et al., 2003; Brown et al., 2006). For *BDNF*, *STMN1*, and *TPH2* tagging SNPs were selected from HapMap Public Release #19 applying the efficient multimarker method with r^2>0.8 and minor allele frequency (MAF)>0.05 as implemented in the HapMap web browsers (http://www.hapmap.org; de Bakker et al., 2005). Two of the selected *HTR1D* SNPs

(rs676643 and rs674386) were also present as tagging SNPs in the HapMap database. In Table 5.2 the selected SNPs and coverage rate per candidate gene are listed.

Multiplex genotyping assays were designed using Assay Designer software (Sequenom, San Diego, CA). SNPs were genotyped by mass spectrometry (the homogeneous MassARRAY system; Sequenom, San Diego, CA) using standard conditions. PCR reactions were carried out in a final volume of 5 µl and contained standard reagents and 2.5 ng of genomic DNA. Genotypes were assigned by using Genotyper version 3 software (Sequenom, San Diego, CA).

Genotype call rates for each multiplex were checked within the cohorts. Samples with call rate <75% were excluded from further analyses in the datasets. Success rates of the SNPs ranged from 97.9 to 100% for the GenED case group, and from 87.3 to 100% for the NTR control group. Between 6 and 10% of the samples were genotyped in duplicate and checked for concordance. Duplicate genotyping error rates were 0.07% in the case group, 0.2% for the control sample.

For the GAIN GWA controls, genomic DNA was isolated from blood samples. Individual genotyping was conducted by Perlegen Sciences (Mountain View, CA, USA) using a set of four proprietary, high-density oligonucleotide arrays (Sullivan et al., 2009). SNPs were imputed by Abecasis' MACH (v1). For the imputed SNPs the average maximum posterior probability was calculated. This measure represents how much uncertainty there is for the imputation of each SNP, ranging from 0 (high uncertainty) to 1 (low uncertainty).

Statistical analyses

The χ^2 test for Hardy-Weinberg equilibrium (HWE) was calculated in the NTR controls using the HWE program of LINKUTIL (http://linkage.rockefeller.edu/ott/linkutil.htm). To investigate the association of the 25 SNPs from four candidate genes we applied a two-stepped approach. First, allele frequencies for all SNPs were compared between cases from the GenED study and controls from the NTR. SNPs that showed nominal significant association (p<0.05) with either AN or SV in the first step, were tested for replication in a meta-analysis with the two additional independent case-control samples (EDNET and Essen, and Utrecht).

Differences in allele frequencies were compared and tested for significance by Pearson's chi-square test with SPSS version 15 software (SPSS, Chicago, IL). For the meta-analysis, the fixed- and random-effects model of DerSimonian and Laird (DerSimonian & Laird, 1986) was used to estimate summary odds ratio's (ORs), as implemented in R (http://www.r-project.org/, package meta). The heterogeneity was

quantified using the I2 statistic for inconsistency (Higgins & Thompson, 2002) and its statistical significance was tested with the $\chi 2$ distributed Cochran Q statistic (Lau et al., 1997). I^2 describes the proportion of variation that is unlikely to be due to chance and is considered large for values over 50% (Higgins & Thompson, 2002). Two tailed *p*-values are reported for all analyses.

Power calculations were performed in Quanto version 1.2.4 (2009). Instead of adjusting *p*-values a priori for multiple testing, nominal *p*-values are provided in order to allow the reader to interpret the level of significance. The results from the final analyses were corrected for multiple testing by using an interface developed by Nyholt (2004), available at http://genepi.qimr.edu.au/general/daleN/SNPSpD/. Given the fact that the linkage disequilibrium (LD) structure among the SNPs was not independent, adjusting the *p*-value for the actual number of tests would be overly stringent and result in a loss of power. With this method the *p*-values were therefore adjusted for the estimated number of *independent* SNPs tested. Calculation of the number of independent SNPs (also called the effective number of SNPs; M_{eff}) was based on the number of eigenvalues of the *n* x *n* correlation matrix of allele frequencies of SNPs using equation 5 by Li and Ji (2005).

Results

SNP association analysis

In the NTR control group none of the SNPs revealed a departure from HWE ($p>0.01$). Depending on the MAF of the SNP, this initial study had adequate power (85% power at an alpha level of 0.05, log-additive or allelic model) to detect effects sizes ranging between 1.45 and 1.8 for AN and ranging between 1.48 and 1.85 for SV.

The results of the association analysis in the initial study (GenED cases and NTR controls) are presented in Table 5.3. A nominal significant association ($p<0.05$) was observed for *TPH2* rs1473473 in AN as well as SV. This SNP was followed-up in the meta-analysis.

Exploratory association analyses were performed in the restricting type AN subgroup (N=108) of the GenED study and the NTR controls. The results of these analyses are presented in supplementary Table S5.1. No significant association was observed for any of the 25 SNPs

Table 5.3 Minor allele frequencies (MAF) for each SNP in cases of the GenED study and NTR controls

				Control	AN			SV		
			DNA	(n=607)	(n=182)			(n=149)		
Gene	Position	SNP	change	MAF	MAF	X^2	*p*	MAF	X^2	*p*
STMN1	1.p36.11	rs12037513	A>G	0.35	0.32			0.33		
		rs807055	C>T	0.43	0.39			0.37	3.14	0.08
		rs807062	G>C	0.25	0.26			0.24		
HTR1D	1.p36.12	rs605367	T>C	0.31	0.33			0.33		
		rs6300	A>G	0.10	0.10			0.07		
		rs676643	G>A	0.16	0.15			0.16		
		rs674386	G>A	0.29	0.30			0.30		
TPH2	12.q21.1	rs10748185	G>A	0.49	0.45			0.46		
		rs2129575	G>T	0.26	0.25			0.24		
		rs17110489	T>C	0.26	0.27			0.24		
		rs7305115	G>A	0.41	0.41			0.41		
		rs1007023	T>G	0.12	0.15			0.16	3.38	0.07
		rs4760820	C>G	0.43	0.40			0.38	2.84	0.09
		rs1473473	T>C	0.14	0.18	4.26	0.04	0.19	4.82	0.03
		rs3903502	C>T	0.39	0.42			0.41		
		rs12231356	C>T	0.08	0.05	3.41	0.07	0.07		
		rs4474484	G>A	0.35	0.36			0.37		
BDNF	11p14.1	rs7124442	T>C	0.33	0.29			0.28		
		rs6265	C>T	0.19	0.19			0.20		
		rs11030107	A>G	0.27	0.23			0.24		
		rs7103873	G>C	0.46	0.49			0.48		
		rs11030123	G>A	0.11	0.10			0.10		
		rs17309930	C>A	0.20	0.20			0.18		
		rs2049048	G>A	0.16	0.13			0.17		
		rs1491851	C>T	0.46	0.45			0.46		

AN=anorexia nervosa, SV=eating disorders characterized by self-induced vomiting

Reported results are comparisons between allele frequencies (1df), *p*-values <0.1 are shown only.

Meta-analysis

The *TPH2* SNP rs1473473 was genotyped in the EDNET and Essen and the Utrecht case-control samples. In the GAIN GWA control group this SNP was imputed. The average maximum posterior probability, which represents how much uncertainty there is for the imputation of a SNP, was 0.99 for *TPH2* rs1437473. For the meta-analysis, genotype data was available for a total of 2,987 individuals (887 AN cases, 306 SV cases, 1914 controls) which provides adequate power (85% power at an alpha level of 0.05, log-additive or allelic

model, MAF of 0.16) to detect effect sizes higher than 1.25 for AN and higher than 1.4 for SV.

Table 5.4 Meta-analysis of *TPH2* SNP rs1473473 in anorexia nervosa and eating disorders characterized by self-induced vomiting

	SNP rs1473473				SNP rs1473473			
	AN				SV			
Study	OR	CIL	CIR	*p*	OR	CIL	CIR	*p*
GenED	1.39	1.02	1.92	0.040	1.46	1.04	2.04	0.029
Utrecht	1.25	0.98	1.60	0.067	1.24	0.91	1.70	0.176
Ednet-Essen	1.11	0.81	1.51	0.53	-	-	-	-
Meta-analysis	**1.25**	**1.06**	**1.47**	**0.009**	**1.34**	**1.06**	**1.69**	**0.013**

AN=anorexia nervosa, SV=eating disorders characterized by self-induced vomiting , CIL=lower 95% confidence interval, CIR=upper 95% confidence interval. Number of AN and SV cases per study: GenED AN N=182, SV N=149; Utrecht AN N=285, SV N=157; EDNET-Essen AN N=420, SV N=0.

Table 5.4 shows ORs, their 95% CI and *p*-values within the individual case-control samples and the subsequent meta-analyses. For the minor C-allele (frequency 0.16) of *TPH2* SNP rs1473473 a significant association was observed in the meta-analyses with both AN and SV. We observed an OR of 1.25 (95% CI 1.06-1.47, $p<0.009$) for AN, and an OR of 1.34 (95% CI 1.06-1.69, $p<0.013$) for SV. There was no significant evidence for heterogeneity of the effect in the AN or SV analyses ($p=0.58$, $I^2= 0\%$ and $p=0.50$, $I^2=0\%$).

Table 5.5 Genotype counts *TPH2* rs1473473 for the three case-control samples

	AN			SV			Control		
	Genotype (n)			Genotype (n)			Genotype (n)		
Case-control sample	11	12	22	11	12	22	11	12	22
GenED	123	52	7	95	52	2	447	125	18
Utrecht	187	90	8	95	49	3	789	300	29
EDNET-Essen	266	128	16	-	-	-	130	50	9

The OR for the combined group of AN and/or SV cases (n=1073) was 1.24 (95% CI 1.06-1.44, $p<0.006$). We could not observe significant evidence for heterogeneity of the effect ($p=0.38$, $I^2= 0\%$) between the different case-control samples. Based on the genotype frequencies of the *TPH2* SNP rs1473473 (presented in Table 5.5) we expected a dominant effect to be underlying the association. Therefore, we evaluated the association with this SNP in the combined case-group under a dominant genotypic model.

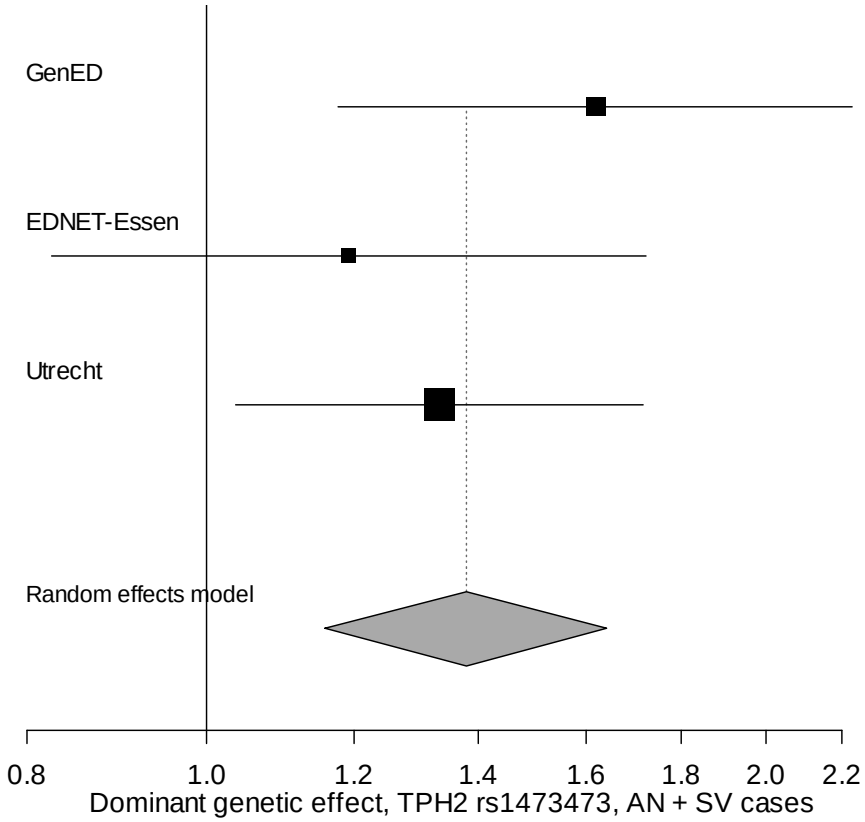


Figure 5.1 Random effect plot of the association between *TPH2* rs1473473 and the combined AN/SV cases under a dominant genotypic model.
Results: GenED: OR=1.62 (95% CI 1.18-2.23); EDNET-Essen: OR=1.19 (95% CI 0.83-1.72); Utrecht OR=1.33 (95% CI 1.04-1.72); Random effect model total: OR=1.38 (95% CI 1.16-1.64, p<0.0003)

Figure 5.1 represents the results of this association. Homo- and/or heterozygous carriers of the minor allele of rs1473473 had an increased probability of either AN or SV (OR=1.38, 95% CI 1.16-1.64, p<0.0003). Again, no evidence for heterogeneity was observed (p=0.44, I^2= 0%). As there is a general tendency for initial studies to overestimate effect sizes, we tested sensitivity of the association by excluding the discovery sample (GenED cases and NTR controls). Under the dominant genotypic model, carriers of the minor allele of rs1473473 had an OR of 1.29 (95%CI 1.05-1.59, p<0.018) among the two replication case-control samples. In supplementary Figure S5.1, the LD-plot between *TPH2* rs1473473 and the nine other selected *TPH2* tagging SNPs is depicted.

Because the LD structure among the SNPs was not completely independent, adjusting the *p*-value for the actual number of tests would be overly stringent and result in a loss of power. By using the interface developed by Nyholt (2004), the number of independent SNPs in our study was estimated to be 23.5. This led to an experiment-wide significance threshold of $p<0.002$. Thus the observed dominant effect of rs1473473 in the final analysis in the combined AN-SV group remained significant after adjustment for multiple testing. However, the observed effects in the separate AN and SV analyses did not remain significant. In this case, the method by Nyholt (2004) was still conservative since not all 25 SNPs were measured in the additional EDNET and Essen, and Utrecht case-control samples.

Discussion

This is the first study to report that *TPH2* SNP rs1473473 is significantly associated to AN and ED characterized by SV. When the two ED case groups are combined, a dominant genotypic model for rs1473473 shows that carriers of the minor allele of rs1473473 had a higher risk of AN or SV (OR=1.38, 95% CI 1.16-1.64, $p<0.0003$). This SNP tags an LD block that spans across part of the *TPH2* gene, and is ended by a recombination hotspot on one side. Therefore it is highly likely that this SNP is in LD with a functionally relevant variant(s) in the *TPH2* gene. The *TPH2* gene encodes the main rate-limiting enzyme in the synthesis of serotonin in the brain (Zill et al., 2007). Serotonin is involved in satiety, anxious and obsessional behavior, mood, and impulse control, features all linked to ED (Kaye, 2008; Lucki, 1998). In long-term recovered ED participants elevated 5-hydroxyindoleacetic acid levels in cerebrospinal fluid were detected (Kaye, 2008; Kaye et al., 2005c). This is the major metabolite of serotonin in the brain and body and is thought to reflect extracellular serotonin concentrations. This finding thus could be indicative of an 'overactive' serotonin system in ED, which in turn could be caused by an increased function of the *TPH2* gene. *TPH2* was also one of 182 candidate genes that were tested for association by comparing in total 5151 SNPs between 1085 AN cases and 677 controls (Pinheiro et al., 2010). After accounting for multiple testing, there were no statistically significant associations for any individual SNP (including *TPH2*). Rs1473473 is not in LD with known *TPH2* mutations (Haavik et al., 2008). *TPH2* SNPs in LD with rs1473473 however, have been associated with a suicidal mental condition in Finnish men (Zhou et al., 2005), with antidepressant response in depressive patients (Peters et al., 2004), and with allelic mRNA expression imbalance in sections of the human pons (Lim et al., 2007),

indicating that genetic variation at this locus may contribute to mental conditions and could influence gene function.

To retain adequate statistical power, the current study only covered a selection of candidate genes for ED. To replicate previous results in ED, we selected genes for which association was observed and confirmed in studies with an adequate sample size. Both *HTR1D* and *BDNF* fulfilled these criteria, although we acknowledge that inclusion of the gene encoding the opioid delta receptor (*OPRD1*) would also have been appropriate (Bergen et al., 2003; Brown et al., 2006; Ribases et al., 2004; Ribases et al., 2005). Because of previous inconsistent results the Serotonin Receptor 2A and the Serotonin Transporter genes were not included in our selection (for a review see Bulik et al., 2007b; Slof-Op 't Landt et al., 2005). Besides replication of previous results, the current study also aimed to evaluate the involvement of two unexplored candidate genes for ED. Like *HTR1D* and *OPRD1*, *STMN1* was located under the linkage peak of restrictive AN (1p33-36) (Grice et al., 2002). Because the associations with *HTR1D* and *OPRD1* only explained part of the linkage, it was expected that additional candidate genes could underlie the linkage peak (Bergen et al., 2003). *TPH2* was selected because of the link between serotonin and ED. The role of *TPH2* in the synthesis of serotonin (Zill et al., 2007), makes it a plausible candidate gene for ED. Thus far no other genes have been analysed in the GenED study.

A note concerning our study populations is the fact that the EDNET and Essen control population was limited in size and consisted of both men and women. However, no difference in the allele frequency of rs1473473 between sexes was observed, in either the German controls or the GAIN GWA control group (Boomsma et al., 2008). So it is unlikely that this has interfered with our results. Another remark with regard to the German sample is the lack of information regarding self-induced vomiting. Finally, the NTR control group were random controls, not selected based on for example liability to psychiatric disorders or social economic status. Due to the low prevalence of ED in the general population we do not think that this has affected our results.

Another concern is the issue of multiple testing. We acknowledge that if we correct for multiple testing in the GenED study, the association with rs1473473 does not remain significant. However, if we perform permutation analysis in this study the global *p*-value for the association between the *TPH2* gene and SV and AN is still trend significant ($p<0.10$). Therefore we do think that the decision to follow-up the association of *TPH2* SNP rs1473473 in the additional cohorts was justified.

The reported association between the functional *BDNF* Val-66-Met polymorphism (rs6265) and AN was not replicated in this study (Ribases et al., 2004; Ribases et al., 2005). However, this result is in line with several other studies which also could not confirm this

association (Dardennes et al., 2006; de Krom et al., 2006; Dmitrzak-Weglarz et al., 2007; Friedel et al., 2005; Koizumi et al., 2004; Mercader et al., 2007).

Previously two studies have reported significant association between *HTR1D* SNPs (including rs6300 and rs674386) and AN (Bergen et al., 2003; Brown et al., 2006). We did not detect any allele frequency differences between controls and AN cases in the four SNPs that were examined. Considering the strength of the previous association and the allele frequency, we should have had sufficient power to detect an effect of rs6300. For rs674386 on the other hand, statistical power was lower (60%), and the association may have been missed due to this reason.

No consistent associations were observed for the other positional candidate gene, *STMN1*. Despite its position under the linkage peak for AN, it might not be involved in ED. However, since linkage was observed in the restrictive subtype of AN, it is also possible that an effect of this gene is only apparent in this specific eating disorder subgroup. The exploratory analyses in restrictive AN (N=108) of the GenED study and the NTR controls (supplementary Table S5.1), also did not reveal an association with *STMN1*. This exploratory study had adequate power (85% power at an alpha level of 0.05, log-additive or allelic model) to detect effects sizes around 1.6 for restricting AN. Thus, the association may have been missed due to limited statistical power.

For the first time candidate genes in ED characterized by self-induced vomiting were evaluated. We selected this phenotype because there is no *a priori* reason to believe that the DSM diagnostic schema represent more 'genetic' syndromes than underlying core behaviors or traits. A distinctive eating disorder symptom that is shared among different types of ED is self-induced vomiting. Prevalences of vomiting within clinical samples ranged between 31 and 39% for AN (Ben-Tovim et al., 1989; Garner et al., 1993), and even over 90% in BN (Ben-Tovim et al., 1989). The reliability of the measurement of this behavior and the heritability of self-induced vomiting has also been demonstrated (Sullivan et al., 1998a; Wade et al., 2008b). Other symptoms that are shared among ED are binge eating and the undue influence of weight and shape on self-evaluation. Binge eating has a substantial heritability but is less reliably measured (Bulik et al., 1998; Reichborn-Kjennerud et al., 2003; Sullivan et al., 1998a; Wade et al., 2000a; Wade et al., 2008b). The undue influence of body weight appears to be more environmentally mediated (Reichborn-Kjennerud et al., 2004a; Wade & Bulik, 2007).

Many genetic studies in AN have been performed, mainly in small populations measuring only one or a few SNPs (Bulik et al., 2007b). In the current study we used a large population of AN cases. We selected 25 SNPs to capture the majority of the common variation within four candidate genes (*STMN1*, *HTR1D*, *TPH2* and *BDNF*). Our two-step

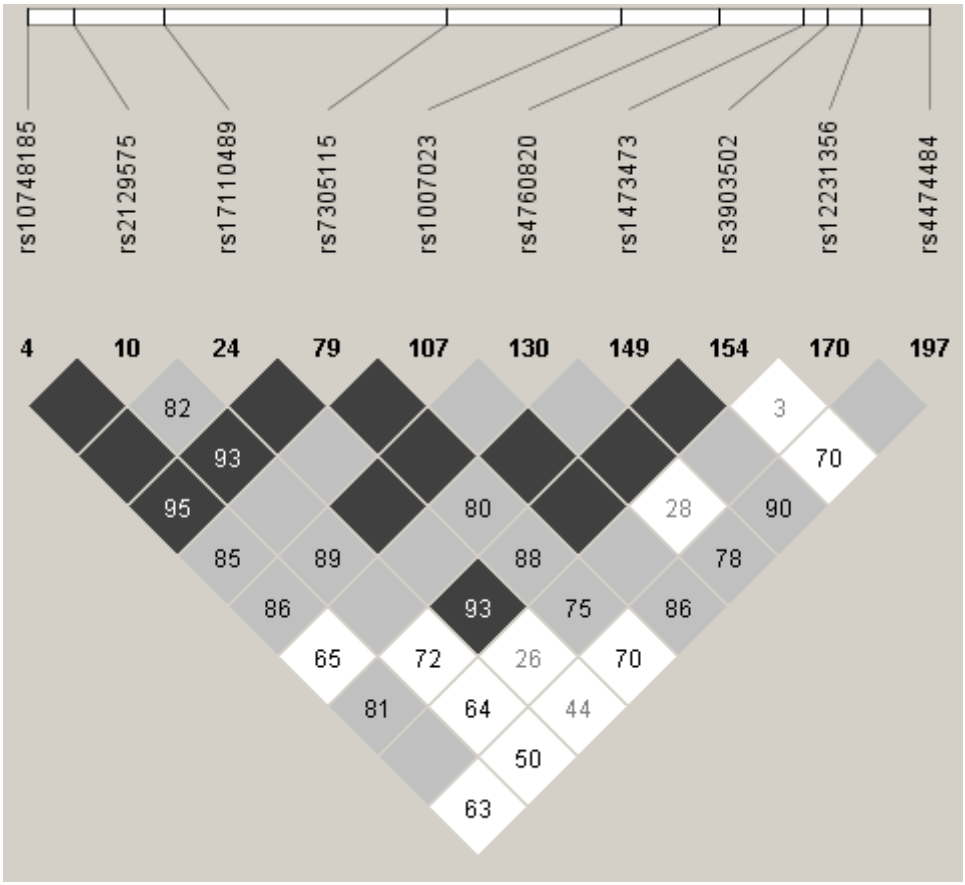
approach gave us the opportunity to explore association with all 25 SNPs in the first step, and to evaluate the initial findings in two additional independent case-control samples. This approach has led to a robust association of the *TPH2* SNP rs1473473. The minor allele of this SNP was associated with a higher risk for AN, SV and for the combined group. It is interesting that the same SNP was associated with both types of ED. Although there was overlap between the two types of ED, 13% of the 887 AN cases also belonged to the SV group, the effect of rs1473473 is also present in the independent AN and SV groups. It has been hypothesized that AN, BN, and also subthreshold forms of ED share at least some risk and liability factors (Kaye, 2008; Strober et al., 2000). In a Swedish twin study, approximately half of the genetic factors contributed to liability to both AN and BN (Bulik et al., 2010). Our current finding is consistent with this hypothesis. For future studies we aim to establish the effect of genetic variation at the *TPH2* gene on behaviors underlying different types of ED, like perfectionism, impulsivity or obsessive-compulsiveness (Kaye, 2008).

Supplementary table S5.1 Minor allele frequencies (MAF) for each SNP in restricting AN cases of the GenED study and NTR controls

				Control (n=607)	Restricting AN (n=108)		
Gene	Position	SNP	DNA change	MAF	MAF	X^2	*p*
STMN1	1.p36.11	rs12037513	A>G	0.35	0.31		
		rs807055	C>T	0.43	0.42		
		rs807062	G>C	0.25	0.29		
HTR1D	1.p36.12	rs605367	T>C	0.31	0.33		
		rs6300	A>G	0.10	0.11		
		rs676643	G>A	0.16	0.17		
		rs674386	G>A	0.29	0.29		
TPH2	12.q21.1	rs10748185	G>A	0.49	0.44		
		rs2129575	G>T	0.26	0.27		
		rs17110489	T>C	0.26	0.27		
		rs7305115	G>A	0.41	0.41		
		rs1007023	T>G	0.12	0.14		
		rs4760820	C>G	0.43	0.42		
		rs1473473	T>C	0.14	0.18	2.77	0.09
		rs3903502	C>T	0.39	0.42		
		rs12231356	C>T	0.08	0.06		
		rs4474484	G>A	0.35	0.37		
BDNF	11p14.1	rs7124442	T>C	0.33	0.29		
		rs6265	C>T	0.19	0.22		
		rs11030107	A>G	0.27	0.24		
		rs7103873	G>C	0.46	0.45		
		rs11030123	G>A	0.11	0.11		
		rs17309930	C>A	0.20	0.19		
		rs2049048	G>A	0.16	0.12		
		rs1491851	C>T	0.46	0.44		

AN=anorexia nervosa.

Reported results are comparisons between allele frequencies (1df), *p*-values <0.1 are shown only.



Supplementary figure S5.1 Linkage disequilibrium (LD) plot for the ten *TPH2* tagging SNPs, based on HapMap. D' values are presented (if absent D'=1), color scheme dark grey: strong evidence of LD ; light gray: uninformative; white: strong evidence of recombination.