A Genome-wide Association Meta-analysis of Preschool Internalizing Problems

Kelly S. Benke, PhD, Michel G. Nivard, BA, Fleur P. Velders, PhD, Raymond K. Walters, PhD, Irene Pappa, MD, Paul A. Scheet, PhD, Xiangjun Xiao, MS, Erik A. Ehli, PhD, Lyle J. Palmer, PhD, Andrew J.O. Whitehouse, PhD, Frank C. Verhulst, MD, PhD, Vincent W. Jaddoe, MD, PhD, Fernando Rivadeneira, MD, PhD, Maria M. Groen-Blokhuis, MD, Catharina E.M. van Beijsterveldt, PhD, Gareth E. Davies, PhD, James J. Hudziak, MD, Gitta H. Lubke, PhD, Dorret I. Boomsma, PhD, Craig E. Pennell, MD, PhD, Henning Tiemeier, PhD, MD, Christel M. Middeldorp, MD, PhD, for the Early Genetics and Lifecourse Epidemiology Consortium

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 3 cohorts (total N = 4,596 children) in which INT was assessed with the same instrument, the Child Behavior Checklist (CBCL). **Method:** First, genome-wide association (GWA) results were used for density estimation and genome-wide complex trait analysis (GCTA) to calculate the variance explained by all SNPs. Next, a fixed-effect inverse variance meta-analysis of the 3 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 ($\lambda = 1.26$, $p < .03$). **Conclusion:** 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. J. Am. Acad. Child Adolesc. Psychiatry, 2014;53(6):667–676. **Key Words:** GWA study, internalizing problems, pcsk2, variance explained, GCTA

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]

Supplemental material cited in this article is available online.

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 years continue to have an effect later in life, even into adulthood (M.G. Nivard, C.V. Dolan, K.S. Kendler, K.J. Kan, G. Willemsen, C.E.M. Van Beijsterveldt , R.J.L. Lindauer, J.H.D.A. Van Beek, L.M. Geels, M. Bartels,

C.M. Middeldorp, D.I. Boomsma, unpublished material, April 2013).[5]

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 1 genome-wide association (GWA) study that analyzed anxiety-related behaviors in children (N = 2,810) 7 years old.[12] None of the effects of the top 10 single nucleotide polymorphisms (SNPs) (p values between 8.7×10^{-7} and 1.2×10^{-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.[13] 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 (standard error [SE] = 0.11) and 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 3 cohorts with a total of 4,596 children in which INT was measured with the same instrument. Each cohort carried out a genome-wide association study (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 3 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 years. 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 in childhood, and internalizing symptoms predict a range of disorders in adulthood, including disruptive disorders and schizophrenia.[1,14-16] This could well be due to overlapping genetic risk factors. Furthermore, 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.

METHOD

Participants

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

Generation R. The Generation R study (www.generationr.nl) is a prospective population-based cohort of 9,745 children born in Rotterdam, the Netherlands, whose due dates were between April 2002 and January 2006.[18,19] Data from a total of 7,893 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

TABLE 1 Description of Cohorts, Internalizing Problem (INT) Scores and Measure, and Estimates of Variance Explained by All Single Nucleotide Polymorphisms (SNPs) Obtained With Density Estimation Method (DE) and With Genetic Complex Trait Analysis (GCTA)

Characteristic	Generation R		NTR		Raine	
N (% girls)	2,037	(49)	1,475	(50)	1,084	(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 version	CBCL/1½ - 5		CBCL/2-3		CBCL/2-3	
Website	www.generationr.nl		www.tweelingenregister.org/en/		www.rainestudy.org.au	
Explained variance by all SNPs 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)

Note: CBCL = Child Behavior Checklist; NTR = Netherlands Twin Registry.